ESSAYS IN PUBLIC ECONOMICS

Multidimensional Inequalities, Fairness Preferences, and Policy Instruments

Daniel Weishaar

Dissertation

Munich, 2025

Essays in Public Economics: Multidimensional Inequalities, Fairness Preferences, and Policy Instruments

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2025

vorgelegt von
Daniel Weishaar

Referent:	Prof. Dr. Andreas Peichl
Korreferent:	Prof. Dr. Emanuel Hansen
Promotionsabschlussberatung:	16.07.2025

Datum der mündlichen Prüfung: 01. Juli 2025

Namen der Berichterstatter:
Prof. Dr. Andreas Peichl
Prof. Dr. Emanuel Hansen
Prof. Dr. Dr. h.c. Clemens Fuest

Research is formalized curiosity. It is poking and prying with a purpose. It is a seeking that he who wishes may know the cosmic secrets of the world and they that dwell therein.

Zora Neale Hurston, *Dust Tracks on a Road*

Acknowledgements

This thesis is the result of a journey filled with intellectual rewards, joy, and personal growth, but also moments of frustration and despair. I am deeply grateful to all those who made this journey special and supported me in overcoming its challenges.

First, I would like to express my sincere gratitude to my supervisors, Andreas Peichl and Emanuel Hansen. Together, they formed an excellent and complementary supervision team, equipping me with the knowledge, networks, opportunities, and guidance that enabled my research to thrive over the past six years. Andreas played a pivotal role in my decision to pursue a PhD, sparking my curiosity in a wide range of economic topics and involving me in research projects early on. I greatly appreciate my in-depth discussions with Emanuel, who always took the time to think through problems in meticulous detail.

The Center for Economic Studies (CES) at LMU Munich was the institutional core of my PhD. I am therefore grateful to Clemens Fuest for maintaining CES as a hub for intellectual exchange and high-quality research in public economics, notably through the visitor program and seminar series, as well as for providing generous financial support for data collection and research stays. Furthermore, I gratefully acknowledge financial support from the Fritz Thyssen Foundation and the Joachim Herz Foundation, which enabled research stays at UC Berkeley and Columbia University. I benefited tremendously from the exchange with researchers there, and I am particularly indebted to Emmanuel Saez and Wojciech Kopczuk for hosting me at these institutions. The whole administrative team at CES also deserves special thanks for their patience and support in navigating me through travel reimbursement forms and academic bureaucracy in general.

This thesis would not be what it is today without my brilliant coauthors and friends, both at LMU and around the world. I learned so much from Ingvild Almås, Felix Bierbrauer,

Maximilian Blömer, Pierre Boyer, Paul Hufe, Niels Johannesen, and Jakob Miethe. It was a genuinely rewarding experience to work with all of you, discussing tax theory, identification strategies, the measurement of fairness preferences, microsimulation models, and so much more. Jakob and Paul, thank you for being such inspiring role models and for your open ears and advice during difficult times.

I was fortunate to be in great company on the PhD journey with amazing colleagues at CES, among others Mehmet Ayaz, Amelie Grosenick, Leonie Koch, Jonas Löbbing, Manuel Pannier, Dominik Sachs, Paul Schüle, and Karl Schulz. Mehmet and Paul, thank you for the numerous coffee breaks, jokes, mutual encouragement, and fantastic joint conference experiences. I am very happy that we went through this journey together.

As research has shown, the foundation for a person's achievement is often shaped very early in life and by the support one receives from the family. I would thus like to thank my parents, Regina and Christoph, for their love, security, and encouragement to explore the world while always providing me with a safe place to return to at any time. I am also grateful to Brigitte and René for their understanding, sincere interest, and genuine care during my PhD, and for always making me feel at home in Bavaria.

This thesis is dedicated to my wife and *Lieblingsmensch* — Olivia. Your unwavering support, endless patience, and unconditional love have been my anchor through every high and low of this journey. Thank you for your inspiration, for your irrevocable belief in me, and for bringing me back to reality when the abstract problems of research and academia were overwhelming me once again. I couldn't have done it without you.

Contents

Preface **1**

1 Homes Incorporated: Offshore Ownership of Real Estate in the U.K. **7**
1.1 Introduction . . . 8
1.2 Data . . . 12
1.3 Descriptive Evidence . . . 16
1.4 Evidence on Causes . . . 25
1.5 Evidence on Consequences . . . 30
1.6 Conclusion . . . 41

2 Pareto-Improvements, Welfare Trade-Offs, and the Taxation of Couples **43**
2.1 Introduction . . . 44
2.2 Theory . . . 50
2.3 Empirical Analysis . . . 67
2.4 Concluding Remarks . . . 78

3 The Political Economy of Joint Taxation **79**
3.1 Introduction . . . 80
3.2 Related Literature . . . 84
3.3 The Model . . . 85
3.4 Marriage Penalties and Bonuses in the US since the 1960s . . . 91
3.5 Reforms towards Individual Taxation . . . 97
3.6 Empirical Analysis of Reforms towards Individual Taxation . . . 101
3.7 Concluding Remarks . . . 107

4 Self-Interest, Stated Preferences, and the Taxation of Couples **109**
4.1 Introduction . . . 110
4.2 Political Background and Data . . . 115
4.3 Material Self-Interest . . . 118
4.4 Stated Preferences . . . 123

4.5 Conclusion . . . 143

5 Just Cheap Talk? Investigating Fairness Preferences in Hypothetical Scenarios 145

5.1 Introduction . . . 146
5.2 Survey Tool and Data Collection . . . 149
5.3 Effects of the "Real-Person Treatment" . . . 156
5.4 Stability of Fairness Preferences . . . 163
5.5 Nature of Fairness Preferences . . . 168
5.6 Conclusion . . . 174

6 Work-Life Balance (Policies) — Evidence from Germany 175

6.1 Introduction . . . 176
6.2 Data . . . 180
6.3 Stylized Facts . . . 182
6.4 Policy Analysis . . . 193
6.5 Welfare Effects . . . 203
6.6 Conclusion . . . 219

A Appendix to Chapter 1 223

A.1 Additional Results . . . 224
A.2 Combining Data Sources . . . 231
A.3 Long-run Dynamics of the Offshore Market Share . . . 241
A.4 Taxation of Capital Gains . . . 246
A.5 Public Ownership Registers . . . 249

B Appendix to Chapter 2 251

B.1 Appendix for the Theory Part . . . 252
B.2 Appendix for the Empirical Part . . . 297

C Appendix to Chapter 3 339

C.1 Appendix: Theoretical Analysis . . . 340
C.2 Appendix: From Theory to Empirics . . . 349
C.3 Supplementary Material: Reforms *in* the System . . . 384
C.4 Supplementary Material: Reforms *of* the System . . . 402
C.5 Supplementary Material: Alternative Sample Restriction . . . 414
C.6 Supplementary Material: Narratives about Tax Reforms . . . 420

D Appendix to Chapter 4 429

D.1 Supplementary Material . . . 430
D.2 Questionnaire (German) . . . 454

D.3 Questionnaire (English Translation) 463

E Appendix to Chapter 5 **473**

E.1 Survey Tool and Data Collection 474

E.2 Effects of the "Real-Person Treatment" 478

E.3 Nature of Fairness Preferences 488

F Appendix to Chapter 6 **493**

F.1 Institutional Regulations 494

F.2 Supplementary Tables and Graphs 499

F.3 Supplementary Derivations 511

Bibliography **519**

List of Figures

1.1 Aggregate Property Investments by Foreign Corporations 19
1.2 Offshore Market Share by Price Segment 20
1.3 Geography of Offshore Market Shares 21
1.4 Offshore Market Share over Time 23
1.5 Beneficial Ownership by Home Country 24
1.6 The Tax Motive 28
1.7 The Secrecy Motive 30
1.8 Offshore Ownership and Property Prices around Brexit 33
1.9 Offshore Capital Flight after the Brexit Referendum 38

2.1 A Simple Tax Reform 54
2.2 Conditional Revenue Functions 55
2.3 A Change in Jointness 61
2.4 Intensive Margin Responses with Discontinuities in Marginal Tax Rates . 64
2.5 Conditional Revenue Functions, Deciles (2019) 71
2.6 Welfare, Reducing Jointness at the Median (2019) 74
2.7 Welfare Implications of Decreasing Jointness (2019) 75
2.8 Welfare Effects of Reforms at the Bottom (2019) 76

3.1 Demographic Change over Time 92
3.2 The Splitting Function σ over Time 93
3.3 Distribution of Marriage Bonuses and Penalties in USD over Time 94
3.4 Political Feasibility 95
3.5 Reform towards Individual Taxation: Political Economy 103
3.6 Reform towards Individual Taxation: Share of Winners over Time 104
3.7 Median Share of Primary and Male Earner 104
3.8 Reform towards Individual Taxation: Welfare (2019) 105
3.9 Swedish Reform to Individual Taxation (1971) Applied to US (2019) . . . 106
3.10 UK Reform to Individual Taxation (1990) Applied to US (2019) 107

4.1 Reform towards Individual Taxation, Political Feasibility 121

4.2 Reform towards Individual Taxation, Welfare . . . 122
4.3 Approval of Reforms . . . 124
4.4 Approval of Reforms by Married Couples . . . 126
4.5 Approval of Reforms by Married Couples, Heterogeneity . . . 128
4.6 Approval of Reforms, Heterogeneity . . . 129
4.7 General Principles, by Married Couple Type . . . 131
4.8 General Principles, by Political Party Preference . . . 132
4.9 General Principles, Correlations . . . 134
4.10 Example of Question . . . 135
4.11 (Preferred) Tax Progressivity, by Vignette Household Type . . . 137
4.12 (Preferred) Marriage Bonus . . . 138
4.13 (Preferred) Tax Benefit for Children . . . 139
4.14 Preferred Jointness . . . 141
4.15 Approval of Reforms, Changes through Information . . . 142

5.1 Survey Structure . . . 150
5.2 Fairness Preference Elicitation, Exemplary Question . . . 151
5.3 Distribution of Allocations, "Real-Person Treatment" . . . 159
5.4 Stability of Fairness Preferences, Correlation . . . 165
5.5 Stability of Fairness Preferences, "Real-Person Treatment" . . . 166
5.6 Gini Coefficient . . . 169
5.7 Fairness Preferences, Importance of Profile Characteristics . . . 172

6.1 Popularity of "Work Life Balance", Google Trends . . . 177
6.2 Desired Change in Working Hours . . . 183
6.3 Actual and Desired Working Hours . . . 184
6.4 Desire for Work-Life Balance across Ability (2019) . . . 186
6.5 (Preferred) Part-Time Share across Ability (2019) . . . 187
6.6 Reasons for Part-Time Employment . . . 189
6.7 Care-Related Part-Time Work, Restrictions (2020) . . . 189
6.8 Reasons for Part-Time Work by Education (2020) . . . 190
6.9 Temporary Part-Time Entitlement Policy (2019), Eligibility . . . 195
6.10 Work-Life Balance Policy and Part-Time Share . . . 196
6.11 Effect of Work-Life Balance Policy . . . 199
6.12 Work-Life Balance Policy and the Nature of Part-Time Employment . . . 203
6.13 Empirical Welfare Effects of WLB Policy . . . 217
6.14 Marginal Welfare Effect, Robustness . . . 218

A.1 Validating Long-Run Time Series . . . 224

A.2 Robustness of Imputation Approach . . . 224
A.3 Foreign Corporate Ownership of Residential Real Estate . . . 225
A.4 Tax Haven Share, Real Estate vs. Bank Deposits . . . 226
A.5 Beneficial Ownership, Real Estate vs. Banking . . . 227
A.6 Tax Motive, Flows . . . 228
A.7 The Tax Motive, Robustness . . . 229
A.8 The Secrecy Motive, Robustness . . . 230
A.9 Distribution of Offshore Ownership Shares in London . . . 231

B.1 Summing over all k . . . 259
B.2 A Couple's Discrete Choice Problem (I) . . . 266
B.3 A Couple's Discrete Choice Problem (II) . . . 271
B.4 Comparison of CPS and SOI Data (1974), Couple Types . . . 298
B.5 Filing Status According to SOI Data . . . 302
B.6 Married Couples Filing Jointly and Separately (2019) . . . 303
B.7 Cond. Revenue Functions, Secondary Earners, PE Q1 . . . 307
B.8 Cond. Revenue Functions, Secondary Earners, PE Q2 . . . 308
B.9 Cond. Revenue Functions, Secondary Earners, PE Q3 . . . 308
B.10 Cond. Revenue Functions, Secondary Earners, PE Q4 . . . 309
B.11 Cond. Revenue Functions, Secondary Earners, PE Q5 . . . 309
B.12 Cond. Revenue Functions, Secondary Earners, PE Q6 . . . 310
B.13 Cond. Revenue Functions, Secondary Earners, PE Q7 . . . 310
B.14 Cond. Revenue Functions, Secondary Earners, PE Q8 . . . 311
B.15 Cond. Revenue Functions, Secondary Earners, PE Q9 . . . 311
B.16 Cond. Revenue Functions, Secondary Earners, PE Q10 . . . 312
B.17 Cond. Revenue Functions, Primary Earners, SE Q1 . . . 312
B.18 Cond. Revenue Functions, Primary Earners, SE Q2 . . . 313
B.19 Cond. Revenue Functions, Primary Earners, SE Q3 . . . 313
B.20 Cond. Revenue Functions, Primary Earners, SE Q4 . . . 314
B.21 Cond. Revenue Functions, Primary Earners, SE Q5 . . . 314
B.22 Cond. Revenue Functions, Primary Earners, SE Q6 . . . 315
B.23 Cond. Revenue Functions, Primary Earners, SE Q7 . . . 315
B.24 Cond. Revenue Functions, Primary Earners, SE Q8 . . . 316
B.25 Cond. Revenue Functions, Primary Earners, SE Q9 . . . 316
B.26 Cond. Revenue Functions, Primary Earners, SE Q10 . . . 318
B.27 Cond. Revenue Functions, Deciles (2019), No Extensive Margin . . . 319
B.28 Cond. Revenue Functions, Deciles (2019), Low Elasticity . . . 320
B.29 Cond. Revenue Functions, Deciles (2019), High Elasticity . . . 321
B.30 Cond. Revenue functions, Quintiles (2019) . . . 322

B.31 Cond. Revenue Functions, Vingtiles (2019) . . . 323
B.32 Cond. Revenue Functions, Tax Brackets (2019) . . . 324
B.33 Welfare, Reducing Jointness at $B_1(6)$, $B_2(6)$ . . . 328
B.34 Welfare, Reducing Jointness at $B_1(6)$, $B_2(10)$ . . . 329
B.35 Decreasing Jointness, Varying Elasticities . . . 330
B.36 Decreasing Jointness, No Extensive Margin . . . 331
B.37 Conditional Reforms at Bottom (2019), Welfare Weights . . . 332
B.38 Conditional Reforms at Bottom (2019), No Extensive Margin . . . 332
B.39 Conditional Reforms at Bottom (2019), Different Elasticities . . . 333
B.40 Optimal Top Tax Rates over Time . . . 334
B.41 Pareto Coefficient, CPS Data . . . 337

C.1 Comparison of CPS and SOI Data (1974), Couple Types . . . 350
C.2 Filing Status According to SOI Data . . . 352
C.3 Married Couples Filing Jointly and Separately (2019) . . . 352
C.4 Upper Limit of Tax Bracket, Single/Couple . . . 358
C.5 Demographic Change, by State . . . 360
C.6 Demographic Change, by Income Decile . . . 361
C.7 Absolute Marriage Bonuses and Penalties over Time . . . 362
C.8 Illustration of σ-Function . . . 362
C.9 Average Tax Rates . . . 363
C.10 Heterogeneity Analysis of the Empirical Splitting Function σ . . . 364
C.11 Cumulative Distribution Function . . . 372
C.12 Probability Density Function . . . 373
C.13 Effective Marginal Tax Rates . . . 374
C.14 Average Elasticities of Couples . . . 375
C.15 CDF, Single Earner and Dual Earner Couples . . . 376
C.16 PDF, Single Earner and Dual Earner Couples . . . 377
C.17 Participation Tax Rates . . . 378
C.18 Participation Elasticities . . . 379
C.19 Cumulative Distribution Function, Primary and Secondary Earners . . . 380
C.20 Probability Density Function, Primary and Secondary Earners . . . 381
C.21 Average Marginal Tax Rates by Primary Earnings . . . 382
C.22 Average Marginal Tax Rates by Secondary Earnings . . . 383
C.23 Mechanical Effect of Tax Reform . . . 386
C.24 Change of σ . . . 387
C.25 Change of Marriage Bonuses and Penalties . . . 388
C.26 Change of Marriage Bonuses and Penalties (cont.) . . . 389
C.27 Political Feasibility . . . 391

C.28 Political Feasibility, Lump-Sum Adjustment Per-Capita 394
C.29 Choice of Lump-Sum Adjustment, Political Economy 395
C.30 Welfare, Baseline and Alternative Measure 401
C.31 Reform towards Individual Taxation, Political Economy 403
C.32 Reform towards Individual Taxation, Welfare (Middle) 404
C.33 Reform towards Individual Taxation, Welfare (Bottom) 405
C.34 Reconciling Rawlsian and Feminist Welfare (2019) 406
C.35 Hypothetical Reforms towards Individual Taxation, US 407
C.36 Sweden's Tax Reform of 1971 410
C.37 United Kingdom's Tax Reform of 1990 413
C.38 Demographic Change, Alternative Sample 415
C.39 The Splitting Function σ over Time 415
C.40 Political Feasibility, Alternative Sample 416
C.41 Reform towards Individual Taxation: Political Economy, Alternative Sample 417
C.42 Reform towards Individual Taxation: Share of Winners over Time, Alternative Sample 418
C.43 Median Share of Primary and Male Earner, Alternative Sample 418
C.44 Reform towards Individual Taxation: Welfare (2019), Alternative Sample 419
C.45 Narratives for TRA69 421
C.46 Narratives for ERTA81 422
C.47 Narratives for TCJA17 422

D.1 Median Income Share of Primary Earner by Deciles 432
D.2 Approval Rates, by Household Type 435
D.3 Approval Rates, by Sex 435
D.4 Approval Rates, by Age 436
D.5 Approval Rates, by Education 436
D.6 Approval Rates, by Type of Married Couple 437
D.7 Limit Reform to Newly Married Couples, Approval Rates 437
D.8 Approval Rates, by Political Party Preference 438
D.9 General Principles, by Household Type 439
D.10 General Principles, by Earner Type 440
D.11 General Principles, by Sex 441
D.12 General Principles, by Approval of Individual Taxation 442
D.13 General Principles, by Approval of Realsplitting 443
D.14 General Principles, by Approval of Family Splitting 444
D.15 General Principles, by Approval to Limit Marriage Benefit to Families with Children 445
D.16 Tax on Additional Income 446

D.17 Vignette Evaluations, Tax Liability, Self-Interested Behavior 447
D.18 Vignette Evaluations, Marginal Tax Rates, Self-Interested Behavior 448
D.19 Support for Reforms - Changes through Information, Details 449

E.1 Distribution Task, Chosen Characteristics 474
E.2 Survey Completion Time 475
E.3 Guess of the Real Person ID, Shares 479
E.4 Gini Coefficient, Treatment Group 488
E.5 Gini Coefficient and Respondent Demographics 490
E.6 Fairness Preferences, Importance of Profile Characteristics, Treatment Group 491

F.1 Desired Change in Working Hours, by Education 503
F.2 Desired Change in Working Hours, by Civil Status 504
F.3 Desired Change in Working Hours, by Sex 505
F.4 Desired Change in Working Hours, by Age 505
F.5 Desire for Work-Life Balance across Education, by Age 506
F.6 Desired Working Hours across Education (2019) 506
F.7 Reason for Part-Time Work, by Sex 507
F.8 Reason for Part-Time Work, by Age 507
F.9 Convexity of Earnings, Illustration 508
F.10 Work-Life Balance Policy and Part-Time Employment, Microcensus . . . 508
F.11 Work-Life Balance Policy and the Nature of Part-Time Employment, by Sex 509
F.12 Effect of Work-Life Balance Policy, Establishment-Level 510

List of Tables

1.1 Data Sources and Aggregates . . . 16
1.2 Aggregate Ownership Statistics (Dec. 2019) . . . 18
1.3 Offshore Ownership and Property Prices, Robustness . . . 35
1.4 Offshore Ownership and Property Prices, Structural Parameter . . . 37
1.5 Offshore Ownership and Vacancies . . . 40

2.1 Pareto Coefficients and Optimal Tax Rates (2018) . . . 77

3.1 Assumptions about Labor Supply Elasticities . . . 102

4.1 Characteristics and Levels in Vignettes . . . 135
4.2 MTR Correlations, Status Quo vs. Preferences . . . 140

5.1 Fairness Preference Elicitation, Characteristics and Expressions . . . 152
5.2 Demographics, Comparison to ACS . . . 155
5.3 Balancing Tests . . . 157
5.4 Allocations, "Real-Person Treatment" . . . 158
5.5 Allocations, "Real-Person Treatment", by Demographic Subgroups . . . 161
5.6 Allocations, "Real-Person Treatment", Alternative Samples . . . 162
5.7 Allocations, "Real-Person Treatment", by Survey Length Subgroups . . . 162
5.8 Preferences, Response Time (Min.), "Real-Person Treatment" . . . 163
5.9 Stability of Fairness Preferences, Correlation, Robustness . . . 167
5.10 Preference Types . . . 171

6.1 Hourly Wages and Working Hours . . . 192
6.2 Work-Life Balance Policies in Selected Countries . . . 194
6.3 Effect of Work-Life Balance Policy, Difference-in-Difference Estimates . . . 200
6.4 Effect of Work-Life Balance Policy, Robustness . . . 201
6.5 Effect of Work-Life Balance Policy, Heterogeneity . . . 202
6.6 Welfare Effect of Work-Life Balance Policy, Parameters . . . 216

A.1 Capital Gains Tax Rates for Foreigners (April 2019) . . . 247

B.1 TAXSIM Variables and CPS Application . . . 301
B.2 Assumptions about Labor Supply Elasticities . . . 307
B.3 Income Deciles (2019) . . . 317
B.4 Taxable Income Brackets (2019) . . . 318
B.5 Welfare Weights . . . 326
B.6 Jointness Reforms, Selected Reform Brackets (2019) . . . 326
B.7 Pareto Coefficients and Optimal Tax Rates, Single Earner Couples (2018) 335
B.8 Pareto Coefficients and Optimal Tax Rates, Dual Earner Couples (2018) . 335
B.9 Pareto Coefficients, CPS (2018) . . . 336
B.10 Optimal Top Tax Rates, CPS (2018) . . . 336

C.1 TAXSIM Variables and CPS Application . . . 353
C.2 Tax Treatment of Couples around the World . . . 355
C.3 US Tax Treatment of Singles and Couples . . . 357
C.4 Overview of US Reforms . . . 358
C.5 Assumptions about Labor Supply Elasticities . . . 390
C.6 Political Economy of Past Reforms (Part 1) . . . 392
C.7 Political Economy of Past Reforms (Part 2) . . . 393
C.8 Welfare Weights for Reforms *in* the System . . . 396
C.9 Welfare Implications of Past Reforms (Part 1) . . . 398
C.10 Welfare Implications of Past Reforms (Part 2) . . . 399
C.11 Welfare Implications of Past Reforms (Part 3) . . . 400
C.12 Welfare Weights for Reforms *of* the System . . . 402
C.13 Tax Reforms in OECD Countries . . . 408
C.14 Sweden, Distribution of Married Couples, 1969 . . . 409
C.15 Text Data Source Types . . . 421
C.16 Underlying Text Data for TRA69 (1967–1975) . . . 423
C.17 Underlying Text Data for ERTA81 (1978–1985) . . . 423
C.18 Underlying Text Data for TCJA17 (2016–2019) . . . 426

D.1 Selected Demographics, SOEP, 2017 . . . 430
D.2 Selected Demographics, forsa Survey, 2021 . . . 431
D.3 Welfare Weights . . . 432
D.4 Statements from Party Proposals . . . 433
D.5 Statements from Party Proposals (cont.) . . . 434
D.6 Aggregate Fiscal Effects, Change in Billion Euro – Taxation Reform including Lump Sum Transfer . . . 450
D.7 Disposable Household Income – Individual Taxation and Lump Sum Transfer 451
D.8 Disposable Household Income – Real Tax Splitting and Lump Sum Transfer 452

D.9 Disposable Household Income – Family Tax Splitting and Lump Sum Transfer . . . 453

E.1 Characteristics of the Real Person . . . 474
E.2 Demographics . . . 476
E.2 Demographics (cont.) . . . 477
E.3 Joint Balancing Tests . . . 478
E.4 Allocated Shares, "Real-Person Treatment" . . . 478
E.5 Distribution of Allocations, Kolmogorov–Smirnov Test . . . 479
E.6 Balancing Tests, Alternative Sample Restrictions . . . 480
E.7 Distribution of Allocations, Kolmogorov–Smirnov Test, Alternative Samples 481
E.8 Allocations, "Real-Person Treatment", Male . . . 481
E.9 Allocations, "Real-Person Treatment", Female . . . 482
E.10 Allocations, "Real-Person Treatment", Age below 45 . . . 482
E.11 Allocations, "Real-Person Treatment", Age 45 and above . . . 482
E.12 Allocations, "Real-Person Treatment", White / Caucasian Ethnicity . . . 483
E.13 Allocations, "Real-Person Treatment", Non White / Non Caucasian Ethnicity . . . 483
E.14 Allocations, "Real-Person Treatment", High School Education and Below 483
E.15 Allocations, "Real-Person Treatment", More Than High School Education 484
E.16 Allocations, "Real-Person Treatment", (Self-) Employed . . . 484
E.17 Allocations, "Real-Person Treatment", Unemployed/Student/Retiree . . . 484
E.18 Allocations, "Real-Person Treatment", Working Less Than 35 Hours/Week 485
E.19 Allocations, "Real-Person Treatment", Working 35 Hours/Week or More . 485
E.20 Allocations, "Real-Person Treatment", Earnings Less Than 40000 . . . 485
E.21 Allocations, "Real-Person Treatment", Earnings 40000 or More . . . 486
E.22 Allocations, "Real-Person Treatment", 5 Questions . . . 486
E.23 Allocations, "Real-Person Treatment", 7 Questions . . . 486
E.24 Allocations, "Real-Person Treatment", 9 Questions . . . 487
E.25 Preferences, Response Time (Min.), "Real-Person Treatment", No Limit . 487
E.26 Preference Types, Treatment Group . . . 488
E.27 Fairness Preferences, Regression Table . . . 489

F.1 Part-Time Regulations Around the World . . . 494
F.2 Part-Time Regulations Around the World (cont.) . . . 495
F.3 Entitlement to Part-Time Work in Germany over Time . . . 496
F.4 General Entitlements to Part-Time Work, Germany . . . 497
F.5 Specific Entitlements to Part-Time Work, Germany . . . 498
F.6 Summary Statistics, Mikrozensus (2020) . . . 499

F.7 Summary Statistics, BHP (2018) 500
F.8 Summary Statistics, SIAB (2018) 501
F.9 Hourly Wages and Working Hours, Summary Statistics (2014) 502
F.10 Hourly Wages and Working Hours, Heterogeneity 503
F.11 Elasticity Estimate, Sensitivity 518

Preface

This dissertation contributes to the field of public economics by examining economic policies and distributional outcomes across different domains, ranging from housing markets (Chapter 1) and the taxation of couples (Chapters 2–4) to fairness preferences (Chapters 4–5) and work-life balance (Chapter 6). Before summarizing the six self-contained essays, I briefly outline their substantive and methodological connections.

On the substantive side, the overarching theme of this thesis is the multidimensional nature of inequalities, fairness preferences, and policy instruments. Chapter 1 evaluates inequalities in the housing market, with a particular focus on the dimension of foreign corporate ownership structures, which play a crucial role for investment motives and market outcomes. Chapters 2 and 3 develop comprehensive theoretical and empirical tools to analyze the consequences of tax reforms in a setting characterized by multidimensional heterogeneity and multiple tax instruments. Building on these tools, one key application—the taxation of couples—is examined. In particular, these chapters assess the efficiency, welfare, and political economy properties of tax reforms when tax policy can differentiate between primary and secondary earners as well as between single households and married couples. Chapters 4 and 5 present evidence on the multidimensional nature of fairness preferences. Chapter 4 investigates citizens' views regarding concrete tax reforms, particularly the transition from joint to individual taxation, and the horizontal and vertical equity concerns embedded in different forms of household taxation. Chapter 5 takes a broader perspective by analyzing fairness preferences concerning different dimensions of the income-generating process and by showing that hypothetical surveys used to elicit these preferences provide valid and stable results that are in line with an incentivized experiment. Finally, Chapter 6 examines the state and evolution of work-life balance, i.e., the reconciliation of multiple life domains like market work, leisure activities, and family obligations. The chapter also discusses policies that facilitate work-life

balance, and shows how such policies can be evaluated from a welfare perspective, among others with reference to the relevant behavioral responses that are also estimated in this chapter.

On the methodological side, the different chapters share common features, too. Both Chapter 1 and Chapter 6 explore quasi-experimental research designs to estimate causal effects. In particular, Chapter 1 uses geographic variation in the degree of foreign ownership to assess the impact of foreign investments on house prices. Chapter 6 uses firm-level variation in employees' eligibility for a reduction in working hours to estimate individual responses to work-life balance policies. For the evaluation of tax reforms, chapters 2–4 rely on microsimulation models. These models allow researchers to assess the properties of the status quo tax system at the individual level, which is necessary to connect theoretical insights to the empirical assessment of tax reforms. For the measurement of fairness preferences, both Chapters 4 and 5 employ factorial surveys. In such surveys, respondents are asked to assess situations that vary randomly in multiple characteristics. For instance, in Chapter 4, respondents indicate their preferred tax liability for households that differ in characteristics such as the number of children, marital status, and earnings. Finally, Chapters 2, 3, and 6 use theoretical insights to derive sufficient statistics for an empirical evaluation of tax, and work-life balance policy reforms, respectively.

Chapter 1. In *Homes Incorporated: Offshore Ownership of Real Estate in the U.K.*, coauthored with Niels Johannesen and Jakob Miethe (Johannesen et al., 2024a), I explore the phenomenon of foreign corporate ownership in the U.K. housing market. The chapter examines the causes and consequences of offshore property investments, particularly their impact on house prices and vacancies. For this purpose, the chapter combines property level price information from the universe of residential real estate transactions in England and Wales with data on corporate ownership from the land register and beneficial ownership from leaked data. The chapter presents three main findings. First, offshore ownership is economically significant, accounting for over 1.5% of the overall market and around 15% of top-end properties. Second, both tax and disclosure rules shape patterns of offshore ownership, indicating that taxation and secrecy are key motives for beneficial owners. Third, lower offshore demand following the Brexit referendum led to declines in prices and vacancy rates, demonstrating that offshore ownership inflates prices and reduces the utilization of the housing stock.

Chapter 2. The second chapter, *Pareto-Improvements, Welfare Trade-Offs, and the Taxation of Couples*, coauthored with Felix Bierbrauer, Pierre Boyer, and Andreas Peichl (Bierbrauer et al., 2024), develops a theoretical framework for analyzing tax reforms in set-

tings where taxpayers exhibit multidimensional heterogeneity and face multiple economic decisions, each subject to fixed and variable costs. The chapter derives conditions under which Pareto- or welfare-improving tax reforms exist and applies this framework to the taxation of couples. A detailed analysis of behavioral responses to taxation in this setting reveals key insights into tax system design. Specifically, the chapter establishes sufficient statistics for the existence of Pareto- or welfare-improving tax reforms and applies them to U.S. data. The empirical findings indicate that tax rates on secondary earners are inefficiently high when their earnings approach those of primary earners. Furthermore, reducing the degree of jointness in taxation is not Pareto-improving, and its welfare implications hinge on a trade-off between poverty alleviation and gender balance. At the top, optimal tax rates for primary and secondary earners are substantially different, even if behavioral responses to taxation are assumed to be symmetric.

Chapter 3. Subsequently, I analyze *The Political Economy of Joint Taxation* in Chapter 3, which is based on research with the same coauthors (Bierbrauer et al., 2023b). While optimal tax theory suggests that secondary earners should face lower marginal tax rates due to higher labor supply elasticities, many tax systems—including the U.S. system—continue to tax couples jointly, imposing equal marginal tax rates on primary and secondary earnings. The chapter develops a framework to assess whether political economy constraints can explain the persistence of joint taxation and applies the corresponding sufficient statistics to historical U.S. data since the 1960s. The chapter presents three core findings. First, the support for reforms towards individual taxation has increased over time and recently reached the majority threshold. Second, while such a reform is supported by a social welfare function that prioritizes earnings equality within couples, it is rejected under a Rawlsian welfare objective, as single-earner couples—the main losers of the reform—are over-represented at the bottom of the income distribution. Third, replicating reform types inspired by past transitions to individual taxation in Sweden and the United Kingdom, I find that support depends strongly on whether the reform applies uniformly, is targeted to specific segments of the distribution, or is combined with general changes in tax progressivity.

Chapter 4. The fourth chapter, *Self-Interest, Stated Preferences and the Taxation of Couples*, extends the analysis of Chapter 3 to Germany and combines political economy analyses routed in utility-maximizing behavior and material self-interest with insights from stated preferences. It is based on coauthored work with Felix Bierbrauer, Maximilian Blömer, and Andreas Peichl, and shows that the tax treatment of couples in Germany is highly controversial. Based on material self-interest, the support for a reform towards

individual taxation is slightly below the majority threshold. Since the race is so close, views on just taxation or on the status of marriage can actually make a difference. To analyze the views of citizens with respect to these issues, the chapter presents results from a survey among a representative sample of around 8,000 individuals of the German population. The survey elicited both preferences towards specific reform options and attitudes towards general principles of taxation. It explained the policies in general terms, but also integrated a factorial survey in which respondents were presented with concrete trade-offs and the consequences of reforms for different household types. According to the preferences stated in the survey, support for a reform towards individual taxation is lower than suggested by the analysis based on self-interest. The divergence can be explained with reference to different social welfare concerns and genuine principles regarding fair taxation. However, even though respondents object to a complete reform towards individual taxation, I find that respondents are on average in favor of reducing the preferential tax treatment of couples, especially at the upper parts of the income distribution.

Chapter 5. In the fifth chapter, *Just Cheap Talk? Investigating Fairness Preferences in Hypothetical Scenarios*, coauthored with Paul Hufe (Hufe and Weishaar, 2025), I explore the reliability of survey-based methods for measuring fairness preferences, e.g., such as the tools used in Chapter 4. In particular, the chapter examines whether a non-incentivized survey captures the same distributional preferences as an impartial spectator design, where choices may apply to a real person. In contrast to prior studies, the survey design involves high stakes, with respondents determining a real person's monthly earnings, ranging from \$500 to \$5,700. Results indicate that the non-incentivized survey module yields nearly identical results compared to the incentivized experiment and recovers fairness preferences that are stable over time. Beyond its methodological contribution, the chapter also contains a substantive analysis of fairness views. The results show that most respondents adopt intermediate fairness positions, with fewer exhibiting strictly egalitarian or libertarian preferences. These findings suggest that high-stake incentives do not significantly impact the measurement of fairness preferences and that non-incentivized survey questions covering realistic scenarios offer valuable insights into the nature of these preferences.

Chapter 6. The single-authored sixth chapter, *Work-Life Balance (Policies) — Evidence from Germany*, examines labor market policies that promote work-life balance (WLB), particularly those enabling flexible reductions of working hours. Using administrative data from the German microcensus and social security records, I first document rising demand for WLB and significant part-time hourly wage penalties. Second, I an-

alyze a recent policy expansion that allowed employees to temporarily reduce working hours while maintaining their hourly wage. Exploiting eligibility criteria based on firm size, I find that the policy caused part-time employment to increase by 0.4 percentage points compared to the control group, driven primarily by direct transitions from full-time to part-time work, with no effects on part-time entry or full-time return. Third, I develop a model that interprets WLB policies as an insurance device against life shocks that prevent full-time employment, such as childcare-related time constraints. By linearizing the convex earnings schedule, these policies facilitate consumption smoothing during such shocks but also introduce moral hazard. Combining the model's sufficient statistics with the empirical estimates, I demonstrate that Germany's recent WLB policy expansion was welfare-improving under a plausible range of parameter values.

Together, these essays offer novel insights into critical questions in public economics, in particular those related to multidimensional inequalities, fairness preferences, and both tax and non-tax policy instruments. The essays further illustrate how theoretical insights and various empirical methods can be used as complementary tools, advancing academic knowledge, but also providing robust and actionable guidance for policymakers.

Chapter 1

Homes Incorporated: Offshore Ownership of Real Estate in the U.K.

This chapter is based on co-authored work with Niels Johannesen and Jakob Miethe. A previous version has been published as a Oxford CBT Working Paper. See Johannesen et al. (2024a) for the full reference.

1.1 INTRODUCTION

Anecdotal evidence suggests that top-end real estate in large cities like London, New York and Miami is increasingly held through opaque offshore structures. For instance, an investigative news story documents that around three quarters of the apartments in One Hyde Park, an exclusive residential development on one of London's most fashionable addresses, belong to corporations in offshore tax havens like the Cayman Islands, the British Virgin Islands and the Isle of Man with no trace of the beneficial owners in the land register (Shaxson, 2013).

Ownership of real estate through offshore shell corporations is a cause of concern for several reasons. First, it may play a role in money laundering and tax evasion. While governments have found new ways to pierce the offshore veil in the realm of financial investments (Alstadsæter et al., 2023; Boas et al., 2024; Johannesen et al., 2024b), no such instrument exists for property investments. Second, it may facilitate avoidance of taxes triggered by property transactions such as capital gains taxes and stamp duties. These taxes are almost impossible to enforce when properties change hands through the sale of shares in offshore shell corporations with no change of ownership recorded in the land registry. Third, the broader effects on real estate prices and urban development may be socially undesirable. The construction of exclusive properties for foreign investors motivated by tax and secrecy may contribute to increasing real estate prices (Knoll et al., 2017), crowd out scarce housing for the middle-class, and create ghost neighborhoods with large contingents of absentee residents.

In this paper, we present descriptive and causal analyses of offshore ownership in the residential property market in the United Kingdom. In the first descriptive part, we document the scale of offshore ownership and analyze its geography and long-run trends. In the second part on causes, we provide evidence on the role of tax and secrecy in shaping offshore ownership. In the third part on consequences, we estimate the real effects of offshore ownership on property prices as well as vacancy rates.

Our analysis combines micro-data from multiple sources. First, the land register serves to delimit the set of properties with corporate owners. Second, a comprehensive database of transaction prices allows us to estimate the value of these properties as well as their share of the market. Combining these two sources is far from trivial as it requires unbundling transactions and land titles to the level of single properties using natural language processing techniques. Third, a global firm database serves us to unwind corporate ownership structures and identify U.K. properties that are controlled indirectly through offshore corporations. Fourth, a database with information from offshore data leaks (e.g. Panama

Papers and Paradise Papers) allows us to learn about the beneficial owners behind the offshore corporations controlling U.K. properties. Fifth, administrative data on vacant dwellings allows us to measure the utilization of the housing stock at the district level.

The descriptive analysis suggests that corporations in offshore tax havens – most commonly the British Virgin Islands, Jersey and Guernsey – own residential properties in the United Kingdom worth around £80 billion. This is four times more than the combined holdings of corporations in other foreign countries such as the United States, Japan and France. The importance of offshore ownership varies strongly across market segments: offshore corporations own around 1.5% of the overall residential market but their market share is almost ten times larger in the most expensive market segments.[1] Today's high levels of offshore property ownership reflects a steeply increasing long-run trend, in particular a spectacular offshore boom in the period 2005-2015. The beneficial owners of the properties are mostly overseas: almost two thirds in Africa, Asia and the Middle East; much fewer in North America and neighboring European countries; and around 15% are round-tripping investors resident in the United Kingdom.[2]

Comparing these patterns to offshore ownership in financial markets yields three interesting insights. First, while offshore ownership is generally less pervasive in residential property markets than in financial markets, the opposite is true at the top of the property price distribution: a larger share of the money invested in exclusive U.K. properties comes from offshore tax havens than money sitting in U.K. bank accounts. Second, the offshore boom is a much more recent phenomenon in property markets than in financial markets: While offshore ownership of funds in U.K. bank accounts has been roughly stable since the financial crisis, offshore ownership of U.K. properties more than doubled in the period 2005-2015 and continues to grow although at a slower pace. Third, money invested in residential properties through offshore structures is much more likely to originate in developing countries than financial investments through offshore accounts.

The next set of results concerns the causal determinants of offshore ownership, in particular the role of taxation and secrecy for the beneficial owners. We exploit two sources of policy variation for identification. First, a reform that tightened the capital gains taxation of foreign corporations with U.K. properties, but temporarily created a loophole for

[1]This finding resonates with the well-documented pattern that offshore financial assets are highly concentrated in the hands of the wealthiest (Alstadsæter et al., 2019; Guyton et al., 2021; Johannesen et al., 2024b; Leenders et al., 2023; Londoño-Vélez and Ávila-Mahecha, 2021; Londoño-Vélez and Tortarolo, 2022).

[2]Existing studies have identified significant round-tripping, i.e. domestic investors buying domestic assets through offshore corporate vehicles, in the context of financial investment (Damgaard et al., 2024; Hanlon et al., 2015).

corporations registered in Luxembourg. Second, a government announcement mandating some but not all offshore tax havens associated with the United Kingdom to set up public corporate registers with information about beneficial ownership. Our results document that both events caused strong and immediate behavioral responses with offshore ownership migrating to jurisdictions with relatively low taxation and low transparency.

The third set of results concerns the real effects of offshore ownership in local property markets. Focusing on London where offshore ownership is particularly widespread, we use the surprising "leave" outcome of the referendum on European Union membership in 2016 as a natural experiment that triggered a sharp increase in property sales by offshore owners. In a model that uses repeat sales for identification and allows price dynamics to vary flexibly across price segments, we show that local areas with a high ex ante share of offshore ownership experienced larger price decreases after the referendum. Consistent with a causal interpretation, local areas with different offshore ownership were on parallel trends prior to the referendum and controls for potential confounders, such as the share of foreign residents and the share of total corporate ownership in the local area, have little impact on the estimates.

We develop an instrumental variable framework to retrieve a structural parameter that captures the sensitivity of property prices to offshore demand. The results suggest that a drop in property demand from offshore corporations equivalent to 1% of total demand reduces property prices by around 1.3%.[3] The large elasticity estimate suggests that property prices are highly sensitive to offshore demand. This may reflect that offshore property demand mostly constitutes new investment flows into the property market with a large potential for shifting prices.[4]

Finally, we study the effect of offshore ownership on the utilization of the housing stock. While the price effects of offshore ownership have mixed consequences for nationals – existing home owners win and potential buyers lose when offshore demand drives prices up – there is a clear efficiency case against unoccupied dwellings. The lack of property-level data on utilization makes identification challenging; however, our district-level analysis suggests that offshore ownership is associated with significantly higher vacancy rates. One set of estimates suggests that around 20% of dwellings owned by offshore corporations are unoccupied.

[3]Note that our elasticity estimate is not directly comparable to the existing much smaller estimates of the elasticity of property prices with respect to foreign demand (Cvijanović and Spaenjers, 2021; Gorback and Keys, 2020) as a 1% in total demand is typically much larger than a 1% increase in foreign demand.

[4]A recent influential paper in asset pricing finds an elasticity of stock prices with respect to new investment flows into the stock market of around 5 (Gabaix and Koijen, 2021).

Our paper contributes to the broad descriptive literature on household wealth held through offshore tax havens (Alstadsæter et al., 2019; Zucman, 2013). While this literature focuses almost exclusively on financial wealth, we add to the scant evidence on offshore ownership in property markets. By focusing on properties held through corporations in offshore tax havens, our object of interest is distinct from recent work that measures all ownership by foreign nationals, including local residents, in the Dubai property market (Alstadsæter et al., 2022). We complement existing descriptive analysis of offshore ownership in the United Kingdom (e.g. Bomare, 2019; Sá, 2014) in two ways. First, by integrating comprehensive price data in the analysis, we are able to compute the share of all residential real estate assets owned through offshore tax havens. Second, by exploiting historical transaction data, we can track this key metric back to the 1990s and document the rise of offshore ownership in the property market over several decades. We are not aware of existing papers that produce long time series of offshore property ownership. Our efforts to pierce the veil of offshore secrecy and assign real estate assets held through opaque ownership structures to their ultimate owners relates to similar efforts for corporate profits (Tørsløv et al., 2023); portfolio investment (Coppola et al., 2021; Zucman, 2013); bank deposits (Alstadsæter et al., 2018; Collin, 2021); and direct investment (Damgaard et al., 2024).

We also contribute to a small existing literature studying the causal determinants of offshore property ownership. Two studies report conflicting evidence on the role of secrecy. One of them shows that increased transparency in property investments – i.e., U.S. rules requiring disclosure of the ultimate owners when corporations make all-cash acquisitions of real estate – is associated with *no change* in behavior (Collin et al., 2022). The other one shows that increased transparency in financial investments – i.e., global rules requiring banks to share account information with the home countries of their foreign customers – is associated with *more* corporate investment in properties (Bomare and Le Guern Herry, 2022; De Simone et al., 2020). Our finding that offshore ownership shifts away from offshore tax havens that improve transparency toward those that do not is consistent with a secrecy motive and resonates with earlier evidence on shifting of financial assets across offshore tax havens in response to enforcement efforts (Casi et al., 2020; Johannesen, 2014; Johannesen and Zucman, 2014; Menkhoff and Miethe, 2019). Our finding that offshore ownership in property markets are highly responsive to tax incentives does not, to the best of our knowledge, have antecedents in the literature.

Finally, we contribute by showing that ownership through offshore tax havens has real consequences in housing markets. We show that a negative shock to offshore demand for U.K. property, triggered by the outcome of the Brexit referendum, was associated with a significant drop in prices and retrieve the structural elasticity. This relates to earlier

work showing that positive shocks to foreign demand for property, triggered by shocks in the investor countries (Badarinza and Ramadorai, 2018; Gorback and Keys, 2020; Sá, forthcoming), can drive up prices. Our analysis differs from this earlier literature by focusing on ownership through offshore tax havens, as opposed to ownership by foreigners per se. A priori, offshore ownership may be different from other foreign ownership in several dimensions. Our analysis contributes to the literature by suggesting it is associated with a lower utilization of the housing stock, presumably because the underlying motives are tax and secrecy rather than residence.

The paper proceeds in the following way. Section 1.2 describes the data. Section 1.3 provides descriptive evidence on offshore ownership. Sections 1.4 and 1.5 report results on the causes and consequences of offshore ownership respectively. Section 1.6 concludes.

1.2 DATA

This section describes the data sources we use in the paper and explains how we combine them to obtain a database of offshore ownership of real estate in England and Wales. In brief, we obtain comprehensive information about corporate acquisitions of property from the land register (section 1.2.1); add information about prices and transaction dates from an administrative database of real estate transactions in the residential market (section 1.2.2); and add information about holding companies and ultimate owners from a global firm database and from various offshore leaks (section 1.2.3). We provide an overview of the database (section 1.2.4) and report technical details about how the different data sources are combined in the Appendix A.2.

1.2.1 Corporate Ownership

Our main data source is two publicly available datasets from HM Land Registry with comprehensive administrative information about land and property in England and Wales held by legal entities. The Commercial and Corporate Ownership Data (CCOD) covers titles held by entities incorporated in the UK whereas the Overseas Company Ownership Data (OCOD) covers titles held by entities incorporated outside of the United Kingdom. Both datasets are organized as snapshots taken at regular intervals: The CCOD is available quarterly from March 2014 and monthly from December 2017 whereas OCOD is available quarterly from October 2015 and monthly from December 2017. Both datasets contain

information about properties (e.g. address, tenure), about owners (e.g. company name, country of incorporation) and the date on which the proprietor was added to the land register.

We complement these relatively recent snapshots of the land register with information from two alternative sources that cover a longer time horizon. First, through a Freedom of Information (FOI) request, we have obtained data on all purchases of real estate in England and Wales by foreign corporations since 1990. Second, we use information from the current affairs magazine Private Eye (PE) that has collected property and ownership information about property titles in England and Wales acquired by foreign companies between 1999 and 2014. No similar data on purchases by domestic corporations is available.[5]

By harmonizing property addresses and company names across these four sources using natural language processing techniques, we are able to combine them to a single database of corporate acquisitions of property in England and Wales. For each acquisition, we record the property, the corporation and the month in which the corporation is added to the land register as the owner of the property. For domestic corporations, the only data source is CCOD, which means that our database includes all acquisitions after March 2014, but only acquisitions before March 2014 to the extent that the corporation remained the registered owner by that date. For foreign corporations, we rely on OCOD as well as FOI and PE, which means that our database includes all acquisitions since 1990.

1.2.2 Transaction Prices and Dates

We combine the database of corporate acquisitions with the Price Paid Data (PPD) from the HM Land Registry, which covers all residential property transactions in the United Kingdom since January 1995. This step is complicated by the fact that both land titles in CCOD/OCOD and transactions in PPD not rarely lump together multiple properties. We overcome this challenge by developing an algorithm that breaks down all three data sources to the level of single properties with unique addresses and merge them at this level (see Appendix A.2 for details).

Combining ownership data with transaction data from PPD allows us to add important information for the subset of acquisitions that involves residential properties. First, we obtain the transaction prices. While the most recent transaction price is often available directly for the properties in the CCOD/OCOD snapshots, this is not the case for prop-

[5]HM Land Registry confirmed this in a response to a second Freedom of Information request.

erties that were acquired by corporations and sold again prior to the first snapshots. For instance, if an offshore corporation acquires a property in July 2010 and sells it in May 2012, the property is not in the CCOD/OCOD snapshots; however, we can identify the acquisition in the FOI data and retrieve the transaction prices in the PPD data. Second, we obtain the transaction month. This does not generally coincide with the month in which the new owner is recorded in the land register because of administrative delays. For instance, if an offshore corporation acquires a property in July 2010, it is typically added to the land register in October or November the same year, but we can determine the actual transaction month in the PPD data. Third, we obtain the month in which corporate ownership ends. Again, this can often be inferred approximately from CCOD/OCOD, as the month when a corporate owner drops out of the land registry, but the PPD allows us to determine the timing more precisely, as the month in which a property previously acquired by a corporation is transacted, and - perhaps more importantly - we can also apply this approach to the period not covered by the CCOD/OCOD snapshots. For instance, if an offshore corporation acquires a property in July 2010 and sells it in May 2012, the property is not in the CCOD/OCOD snapshots. However, we observe that ownership is transferred to a foreign corporation in the FOI data and we observe that the property is sold to another owner in May 2012 in the PPD data. If a property appears in both datasets, we classify it as residential and register all of its ownership changes.

1.2.3 Upstream Owners

We take two distinct approaches to tracing the owners behind the corporations recorded as proprietors of real estate in the land register. The first approach uses ORBIS, a proprietary dataset with accounting information and ownership links for a global sample of firms. For each of the domestic corporate owners identified in the CCOD dataset, we search for upstream ownership links in Orbis. This allows us to identify properties whose direct owner is a domestic corporation, which is in turn controlled by a holding company in a tax haven. Such structures allow the ultimate owners to benefit fully from the advantages offered by tax haven ownership in terms of secrecy and tax advantages while not exposing this directly in the land register.

The second approach uses leaked information from corporate service providers and corporate registers in offshore tax havens published by the International Consortium of Investigative Journalists (ICIJ). The leaks pierce the secrecy emblematic for tax havens by identifying the individuals who are the ultimate beneficial owners of the corporations in the leaked files. For each tax haven corporation in OCOD and for each upstream tax haven owner of non-haven corporations identified in ORBIS, we thus search for infor-

mation about the ultimate beneficial owners in all of the five available leaks: Pandora Papers, Paradise Papers, Bahamas Leaks, Panama Papers and Offshore Leaks. This results in a sample of properties held through offshore corporations, for which we are able to determine the home country of the individuals who are the ultimate beneficial owners. To be clear, in most of the analysis, we are interested in offshore ownership per se and do not restrict the sample to properties, for which we can identify the ultimate beneficial owners.

1.2.4 Data Overview

Table 1.1 provides an overview of the data sources and how we use them. As shown in Panel A, the various datasets deriving from the land register identify more than 4.2 million properties that are currently owned by corporations or have been owned by corporations at some point since 1990. While most of the corporate owners are registered in the United Kingdom, almost 240,000 properties are owned by corporations registered elsewhere.[6] As shown in Panel B, around 390,000 properties are controlled directly or indirectly through offshore tax havens.[7] For around 220,000 of the 240,000 properties identified in OCOD/-FOI, the foreign corporate owner is registered in an offshore tax haven. Additionally, for around 170,000 properties identified in CCOD, the domestic corporate owner is itself owned by a corporation in an offshore tax haven identified in ORBIS. As shown in Panel C, for around 16,000 properties owned through offshore tax havens, we are able to retrieve information about the beneficial owners by searching the ICIJ leaks. The Panama Papers contribute most of the links to beneficial owners, but the more recent Paradise Papers and Pandora Papers also make sizable contributions.

Finally, in each stage, a significant share of the properties we identify are residential, which allows us to obtain transaction prices from the PPD data (Column 2). Specifically, we identify around 940,000 residential properties with corporate owners across the various datasets (Panel A). Around 100,000 residential properties are owned directly or indirectly

[6]The subcategories in each panel do not sum to the total as the same property may appear in multiple datasets. For instance, if a property first has a domestic corporate owner and then a foreign corporate owner, it appears in both CCOD and OCOD.

[7]Our list of offshore tax havens combines the lists from Johannesen and Zucman (2014) and Gravelle (2015): Andorra, Anguilla, Antigua-Barbuda, Aruba, Austria, Bahamas, Bahrain, Barbados, Belgium, Belize, Bermuda, British Virgin Islands, Cayman Islands, Chile, Cook Islands, Costa Rica, Curacao, Cyprus, Dominica, Gibraltar, Grenada, Guernsey, Hong Kong, Ireland, Isle of Man, Jersey, Jordan, Lebanon, Liberia, Liechtenstein, Luxembourg, Macao, Malaysia, Maldives, Malta, Marshall Islands, Mauritius, Monaco, Montserrat, Nauru, Niue, Panama, Samoa, San Marino, Seychelles, Singapore, Sint Maarten, St. Kitts-Nevis, St. Lucia, St. Vincent-Grenadines, Switzerland, Tonga, Trinidad-Tobago, Turks-Caicos Islands, U.S. Virgin Islands, Uruguay, Vanuatu.

Table 1.1. – Data Sources and Aggregates

	(1) All	(2) Residential
Panel A: Corporate owners	**4,224,072**	**941,577**
- Direct owner is domestic corporation (CCOD)	4,059,065	885,209
- Direct owner is foreign corporation (OCOD + FOI)	239,049	79,117
Panel B: Offshore corporate owners	**389,155**	**98,076**
- Direct owner is tax haven corporation (OCOD + FOI)	221,224	75,013
- Indirect owner is tax haven corporation (ORBIS)	167,931	23,063
Panel C: Beneficial owner observed in leaks	**16,639**	**7,635**
- Pandora Papers	1,729	728
- Paradise Papers	3,753	1,945
- Panama Papers	8,439	3,623
- Offshore Leaks	1,035	568
- Bahamas Leaks	2,647	1,204

Note: This table shows the number of unique post-able addresses across all available years after the data preparation described in the main text. If a property appears in multiple datasets (for example because it changed hands from a domestic to a foreign corporate owner), it is counted for both datasets meaning the components do not necessarily add to the aggregates in bold font. While Column 1 includes all properties, Column 2 includes properties that are residential as they match with the price data (PPD). Panel A shows the number of addresses we identify in the corporate ownership data. Panel B shows the number of properties for which we observe an offshore corporate owner as the immediate investor (OCOD + FOI) and as indirect investors (ORBIS). Panel C shows the number of properties for which we can identify a natural person as the beneficial owner based on a match with the ICIJ data leaks.

through offshore tax havens (Panel B) and around 7,600 of them can be tied to beneficial owners (Panel C). Most of the analysis below requires price information and therefore uses only the sample of residential properties.

1.3 DESCRIPTIVE EVIDENCE

This section provides descriptive evidence on offshore ownership of residential properties in England and Wales. Our key metric is the *offshore market share*, which we define as the share of the property value owned directly or indirectly through corporations in offshore tax havens. Our starting point for estimating offshore market shares is the database created in the previous section.

Below, we present estimates of the offshore market share in the overall residential property market (section 1.3.1); break it down by jurisdictions (section 1.3.2); show how it varies cross-sectionally across price segments and local areas (section 1.3.3); and illustrate how it has evolved over time (section 1.3.4). Finally, we present evidence on the beneficial ownership of properties held through offshore corporations (section 1.3.5); and compare to patterns of beneficial ownership in financial markets (section 1.3.6).

1.3.1 Offshore Market Shares

To estimate offshore market shares in recent years where the land register snapshots are available, we face two empirical challenges. First, we need the values of all properties. As we only observe the market price of properties directly at the time they are transacted, we impute property values at all other times by applying changes in the district-level house price index to the most recent transaction prices. Second, we need to delimit the set of properties in the market of interest. While the land register snapshots only cover corporate-owned properties, the PPD covers all residential properties in England and Wales that have been transacted at least once since 1995. Hence, we can estimate the denominator of the offshore market share by summing imputed values over the set of properties in the market and the numerator by summing imputed values over the subset of these properties with offshore owners.

Table 1.2 shows the relevant magnitudes for the overall residential property market in England and Wales at the end of 2019. We estimate that the aggregate value of all properties in the market is around £5,000 billion.[8] The estimated value of the properties with offshore corporate owners is around £80 billion, with the bulk of the offshore ownership being direct and only a smaller fraction indirect through domestic corporations. These estimates imply a modest offshore market share in the overall residential market just above 1.5%.

By comparison, the aggregate value of residential properties with foreign non-haven corporate ownership is less than £20 billion. It follows that the foreign corporations owning properties in England and Wales are overwhelmingly registered in offshore jurisdictions. This striking pattern suggests that secrecy and tax evasion may be important motives for these investments, a hypothesis that we explore below. Finally, the value of the properties with domestic corporate ownership and no foreign corporations among the upstream owners is around £260 billion, around three times more than properties with foreign corporate owners. To be clear, the individuals who are the ultimate owners of these properties may be domestic as well as foreign, as we show below in the analysis of ultimate beneficial owners.

[8]Figure A.1 in the Appendix shows that our estimates of the aggregate market value are consistent with national accounts, both in terms of levels and time series dynamics (left panel).

Table 1.2. – Aggregate Ownership Statistics (Dec. 2019)

(A) Residential properties in England and Wales (£ bn.)	**4,920.60**
(B) Owned directly or indirectly by offshore corp. (£ bn.)	**79.28**
- of which: offshore corporation is immediate owner	68.92
- of which: offshore corporation is ultimate holding company	10.37
Offshore market share (= B/A)	*1.61%*
(C) Owned directly or indirectly by other foreign corp. (£ bn.)	**18.19**
- of which: foreign corporation is immediate owner	4.56
- of which: foreign corporation is ultimate holding company	13.63
Other foreign corporate market share (= C/A)	*0.37%*
(D) Owned exclusively by domestic corp. (£ bn.)	**264.43**
Domestic corporate market share (= D/A)	*5.37%*

Note: This table shows estimates of the aggregate value of residential properties in England and Wales by ownership structure as of December 2019. The total residential real estate stock value (A) includes properties owned by corporations in offshore tax havens (B), other foreign corporations (C), and domestic corporations (D). Foreign-owned properties are further differentiated by direct foreign ownership, observed in the land register, and indirect foreign ownership identified via ORBIS. For details on the ownership assignment and the estimation of market shares, see Appendix A.2 and Appendix A.3.

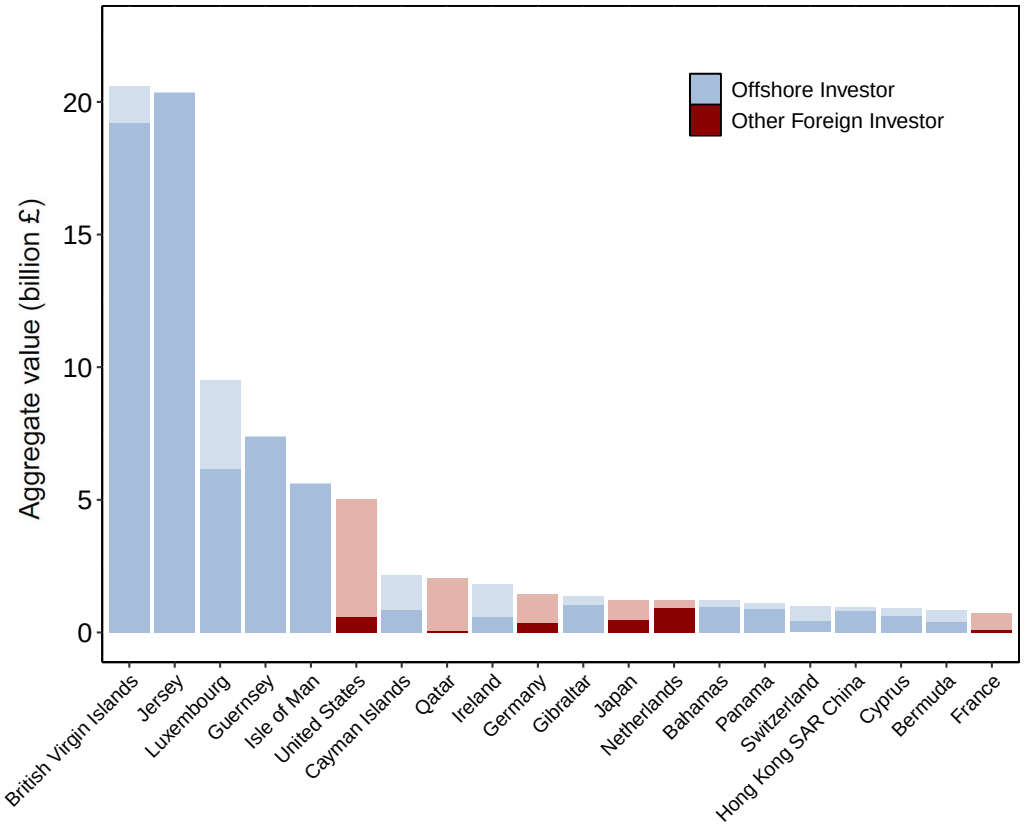
1.3.2 Geography of Foreign Corporate Ownership

Having shown above that the bulk of foreign corporate owners are registered in tax havens, we provide more detail by breaking down foreign corporate ownership by jurisdictions in Figure 1.1. The top owners are the British Virgin Islands and Jersey whose corporations jointly own residential property in England and Wales of around £40 billion. The vast majority of those properties are held directly (dark shading) whereas properties held indirectly through U.K. corporations contribute much less (light shading). The next jurisdictions are Luxembourg, Guernsey and the Isle of Man whose corporations together control properties worth around £20 billion. Hence, all jurisdictions in the top-5 are small tax havens with a legal infrastructure that makes them attractive places to establish shell corporations. Only six countries in the top-20 are not tax havens, e.g. United States, Japan and Germany.

1.3.3 Heterogeneity Across Price Segments and Local Areas

We explore the heterogeneity in the offshore market share across price segments and local areas. As shown in Figure 1.2, the offshore market share in the overall market of just above 1.5% conceals a strong positive price gradient. While less than 0.5% of the residential

Figure 1.1. – Aggregate Property Investments by Foreign Corporations

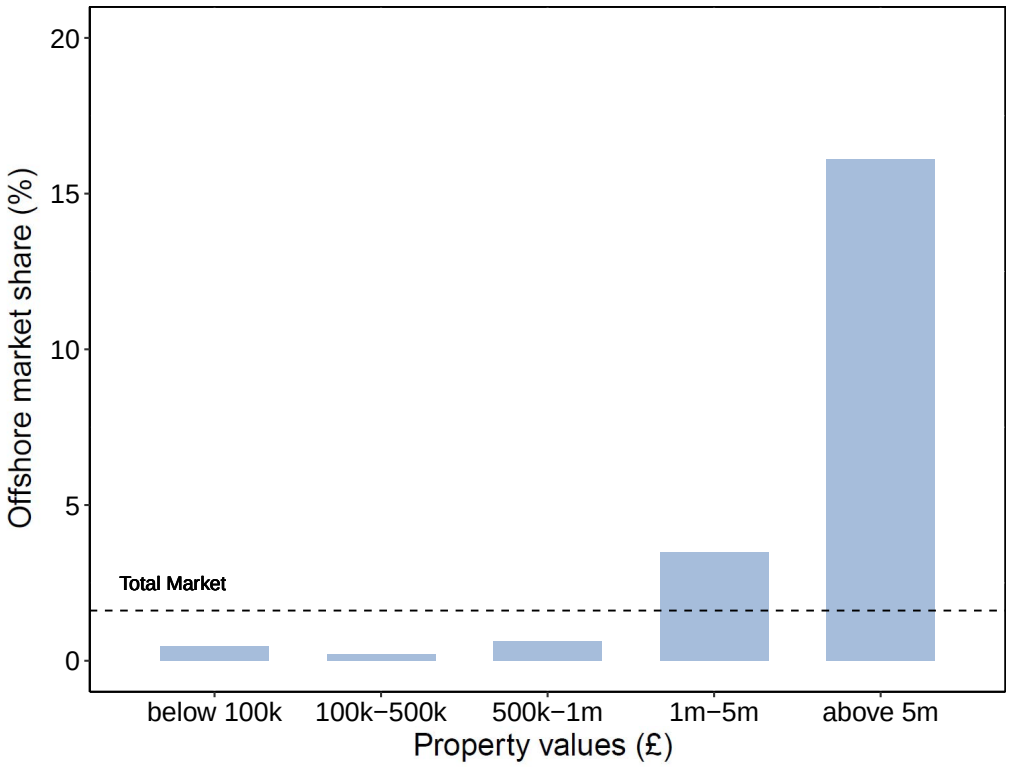


Note: The figure shows the aggregate value of residential property in England and Wales owned by corporations in each of the top-20 foreign jurisdictions in December 2019. It distinguishes offshore tax havens (blue bars) and other foreign countries (red bars) and includes direct ownership (dark color) as well as indirect ownership (light color). For details on the ownership assignment and the estimation of stock volumes, see Appendix A.2 and Appendix A.3.

properties in low-price and mid-price segments have offshore owners, the share increases monotonically with the price level and exceeds 15% in the most expensive price segment with prices above £5 million. This pattern is consistent with anecdotal accounts that associate offshore ownership with the most exclusive residential properties.

As shown in Figure 1.3, offshore ownership is highly concentrated in urban centers such as London, Manchester and Leeds and in selected medium-sized cities such as Exeter, Nottingham and Durham while it is virtually non-existent in rural areas (Panel A). This is consistent with the price gradient in offshore ownership documented above, as most top-end residential properties are located in big cities. Moreover, there is considerable variation in the offshore market share across neighborhoods within the same urban centers. Specifically, offshore ownership in London is extremely high in areas such as Shoreditch, Finsbury Park and Hatton Garden with market shares exceeding 20% and much lower in suburban areas such as Bexley, Havering and Redbridge (Panel B).

Figure 1.2. – Offshore Market Share by Price Segment



Note: The figure shows the share of aggregate residential property value in England and Wales owned by corporations in offshore tax havens, in the total market (dashed line) and in different price segments (blue bars, £) in December 2019. Offshore ownership includes direct as well as indirect ownership. For details on the ownership assignment and the estimation of market shares, see Appendix A.2 and Appendix A.3.

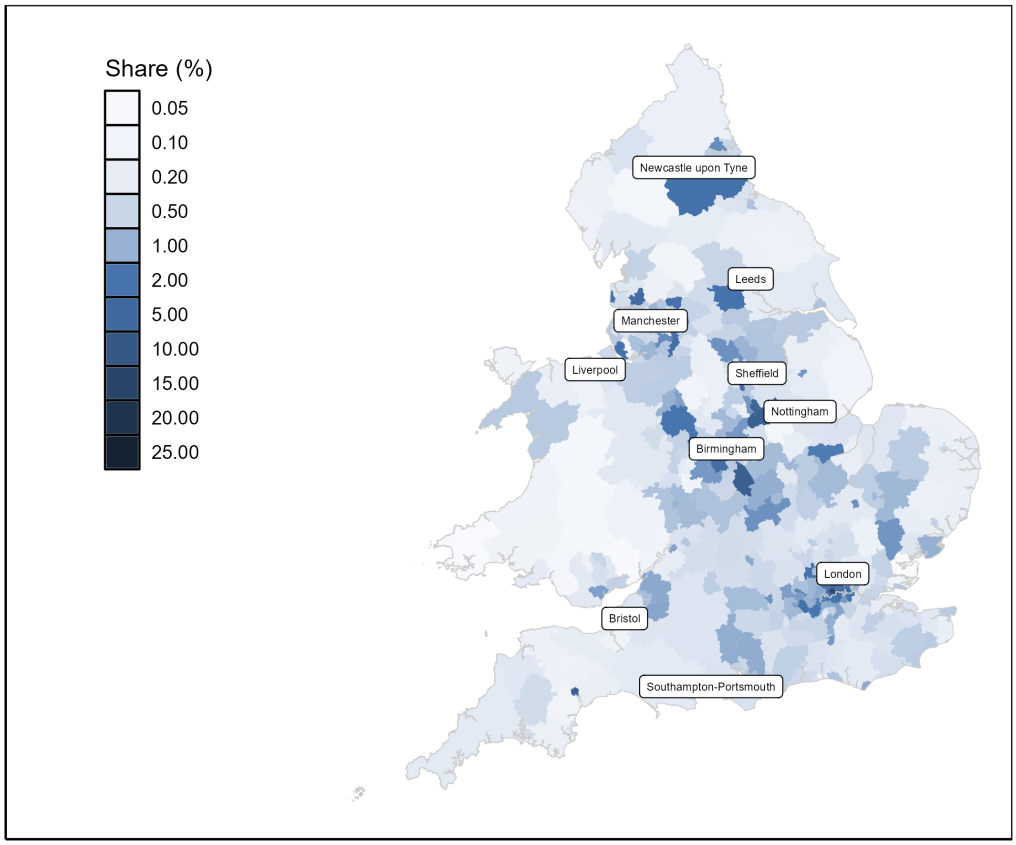
1.3.4 Long-Run Dynamics

We produce a time series for the offshore market share of residential property that goes as far back in time as the data allows.[9] For recent years where the OCOD/CCOD snapshots are available, we can employ the methodology developed in section 1.3.1 for each time period separately. For earlier years, we face one additional challenge: In the FOI/PE data, we observe directly when foreign corporations buy properties and, combining with PPD data, we can also infer when they subsequently sell. Hence, we can track whether a property is owned by a foreign corporation at any date *after* the first time it is transacted within the period with data coverage, but not *before* the first transaction.[10] We address

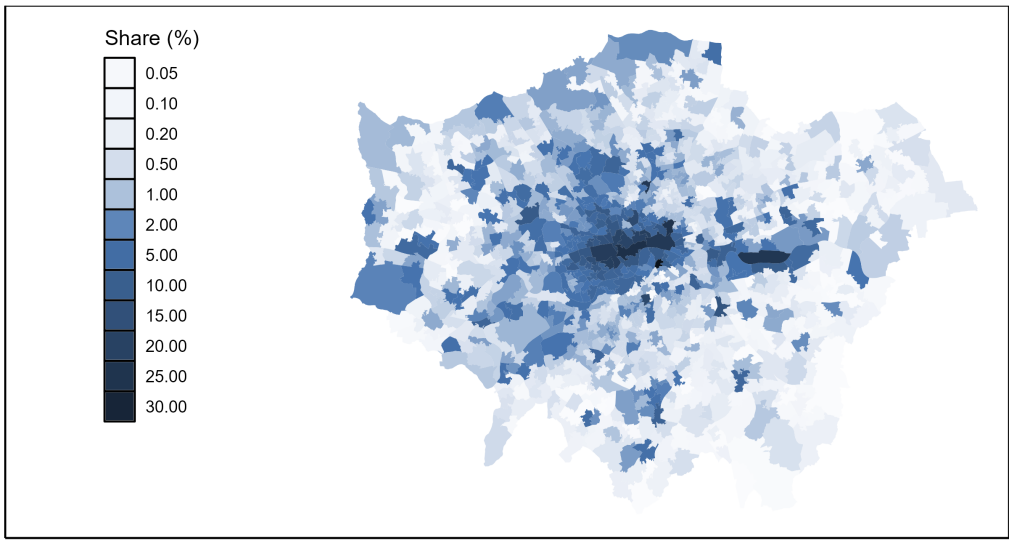
[9]We provide more details on the methodology in Appendix A.3.

[10]For example, if a property is transacted in June 1997 according to PPD, we can determine whether the buyer is a foreign corporation by checking if the transaction is included in the FOI/PE datasets. If the property is transacted again later on, say in May 2003, we can do the same. Hence, we can determine precisely when the property has a foreign corporate owner and when it does not at each point in time since 1997. However, if the transaction in 1997 is the first one within the period with data coverage, we cannot infer from the transaction data whether the seller in 1997 was a foreign corporation or not. Hence, in this case, we cannot make inference about ownership at points in time before 1997.

Figure 1.3. – Geography of Offshore Market Shares



(A) England and Wales



(B) London

Note: The figure shows the share of aggregate residential property value owned by corporations in offshore tax havens varies across local areas in December 2019. Offshore ownership includes both direct and indirect ownership. Figure 1.3a displays market shares across districts in England and Wales while Figure 1.3b displays market shares across local areas in London. For illustration, the upper panel includes labels with major cities. For details on the ownership assignment and the estimation of market shares, see Appendix A.2 and Appendix A.3.

this challenge by assigning to each of these "first transactions" a probability that the seller was a foreign corporation based on what we learn from other transactions in the dataset. Specifically, the assigned probability is the share of foreign corporate sellers across all transactions in the same year, for which we know the seller because there is an earlier transaction in the dataset.[11] Figure A.1 in the Appendix shows that this method delivers accurate imputations in the recent period where we can check them against the OCOD/CCOD dataset (right panel) whereas Figure A.2 shows that the imputed values are relative insensitive to the parametric choices.

As shown in Figure 1.4, the resulting time series reveals a striking pattern. Over the full period, the offshore market share has grown very significantly, from around 0.3% in the middle of the 1990s to the level of around 1.5% in 2019. There are notable differences across time periods: the pace of the increase was slow in the first decade (1995-2005), much faster in the second decade (2005-2015) and then again slow in the last part of the sample period (2015-2019).[12]

1.3.5 Geography of Beneficial Ownership

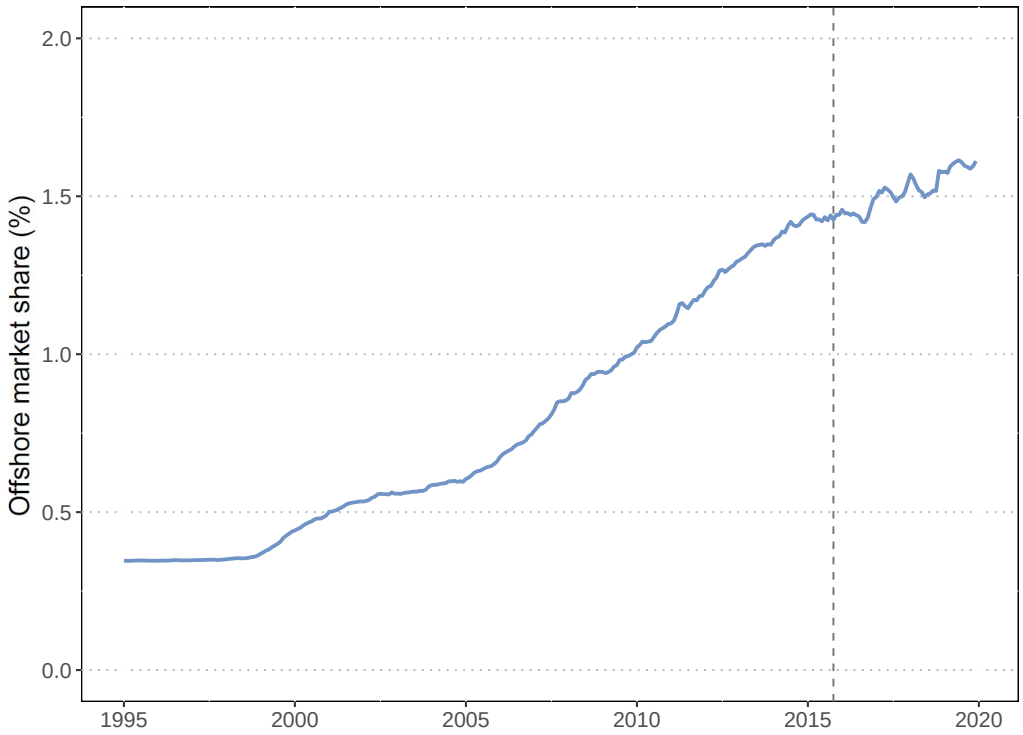
Finally, we shine light on the individuals who own properties through offshore corporations by computing the distribution of beneficial ownership across countries for the set of properties that can be linked to their beneficial owners through the ICIJ leaks. As shown in Figure 1.5, the United Kingdom is by far the largest beneficial owner country with a market share above 15%.[13] Most of the properties that are truly foreign-owned have beneficial owners in the Middle East (e.g. United Arab Emirates and Saudi Arabia), Africa (e.g. South Africa and Nigeria) and Asia (e.g. China and Malaysia) and a relatively small fraction belongs to beneficial owners in Europe and North America. The two largest beneficial owners in Europe (i.e. Monaco and Switzerland), are both popular residence countries for global economic elites and their shares of beneficial ownership therefore do not necessarily reflect beneficial ownership by European nationals. It is interesting that Russia appears to account for a relatively small share of the beneficial

[11] Consider again the example where a property's first observed transaction is in 1997. What is the probability that the seller was a foreign corporation? We identify all the transactions in 1997, for which there exists an earlier transaction in the dataset and for which we therefore know whether the seller in 1997 was a foreign corporation or not. For this set of transactions, we compute the share with a foreign corporate seller and we assume this to be the probability of a foreign corporate seller in the case at hand.

[12] Figure A.3 in the Appendix shows the evolution in the aggregate value of properties held through corporations in offshore tax havens and foreign non-havens, i.e. without scaling by the value of the total market.

[13] This is consistent with other evidence on round-tripping investment in other economic domains (Damgaard et al., 2024; Hanlon et al., 2015).

Figure 1.4. – Offshore Market Share over Time



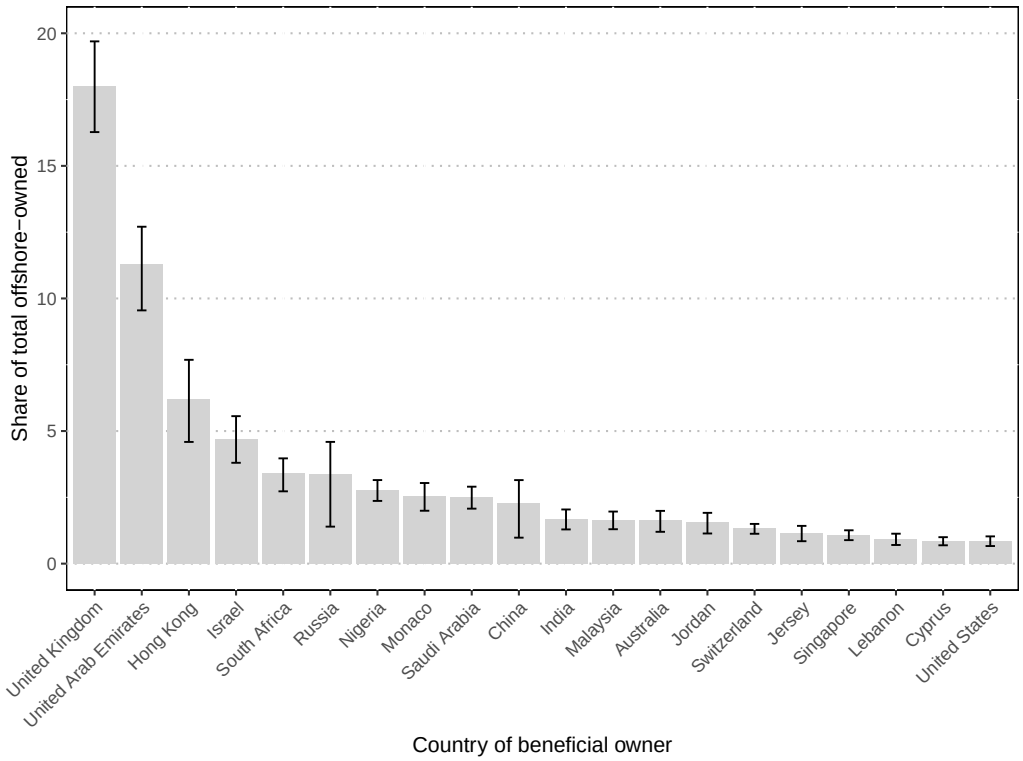
Note: The figure shows the estimated share of the aggregate residential property value in England and Wales that is owned by corporations in offshore tax havens in each month between January 1995 and December 2019. Offshore ownership includes direct and indirect ownership links. The dashed line indicates the start of the real-time snapshots for direct ownership of offshore companies (OCOD). For details on the ownership assignment and the estimation of market shares, see Appendix A.2 and Appendix A.3.

ownership given the anecdotal evidence that Russian oligarchs are heavily invested in London properties (Moore, 2022) and the massive focus on the seizure of such properties following the invasion of Ukraine.

It is important to emphasize that the results are based on a relatively small subsample. This implies that the ownership shares are estimates associated with uncertainty: individual countries may have a smaller or larger weight in the universe of beneficial owners than in the leaked subsample. The bootstrapped standard errors used to construct the confidence intervals in the figure quantify the sampling uncertainty under the assumption that beneficial owners from all countries have the same ex ante probability of being disclosed in a leak.[14]

[14]This follows the approach taken by Alstadsæter et al. (2019). In our setting, the risk of selection bias is smaller because we rely on data from five distinct data leaks and more than five distinct data sources, as some of the leaks include data from multiple sources.

Figure 1.5. – Beneficial Ownership by Home Country



Note: The figure shows the distribution of residential property in England and Wales by the country of the ultimate beneficial owners, in the sample of properties owned by corporations in offshore tax havens, for which we identify the ultimate beneficial owners in the offshore leaks (gray bars). It also illustrates the statistical uncertainty under the assumption of random selection into the leaks by plotting the interquartile range of bootstrapped shares (black vertical lines).

1.3.6 Comparison with Financial Assets

It is illustrative to compare the patterns of offshore ownership in property markets that we have identified above to offshore ownership in financial markets, which has been the focus of a number of earlier studies. In the Appendix, we document how the share of deposits in U.K. banks owned through offshore tax havens has evolved over time (Figure A.4) and illustrate how the beneficial ownership of financial assets in offshore tax havens may be distributed across countries (Figure A.5).[15]

Three interesting insights emerge from the analysis. First, the tax haven boom in property markets occurred much later than in the financial market. While we established above that offshore ownership in property markets built up relatively recently (2005-2015 as

[15]The former analysis uses comprehensive data on U.K. banks from the BIS Locational Banking Statistics. The latter draws on leaked customer data from the HSBC Switzerland, a private bank in Switzerland managing around 2% of global offshore assets at the time of the leak (Alstadsæter et al., 2019). The leaked data was published by the ICIJ under the name *Swiss Leaks*.

shown in Figure 1.4), offshore ownership in deposit markets increased most rapidly in the decade before the financial crisis (1997-2008) and stayed roughly constant, if not falling slowly, in the remainder of the sample period (2008-2019). Second, offshore ownership is generally smaller than in the property market than in financial markets except at the top of the property price distribution. Specifically, the share of offshore ownership in the overall property market (around 1.5% as shown in Table 1.2) is significantly lower than in the deposit market (around 10%); but the opposite is true when considering only the most exclusive properties (around 15% as shown in Figure 1.2). Third, in property markets, the individuals behind the offshore ownership structures are to a much larger extent residents of developing countries than in financial markets. For instance, beneficial owners from Africa and Asia account for a large share of the properties held through offshore corporations (around 20% and each) whereas their share of financial assets in offshore banks is much smaller (only around 4% and 5% respectively).

1.4 EVIDENCE ON CAUSES

This section investigates the underlying motives for using offshore corporations as vehicles for owning real estate. The literature generally highlights two advantages associated with investment through tax havens, *lower taxation* and *secrecy*, and we provide evidence on each of these motives in turn. We first study the role of taxation by estimating the behavioral response to a reform that eliminated the favorable treatment of some but not all offshore structures in terms of capital gains taxation (section 1.4.1). We then study the role of secrecy by estimating the behavioral response to a reform intended to enhance transparency for some but not all offshore structures (section 1.4.2).

1.4.1 The Tax Motive

Generally, owning real estate through an offshore shell corporation can legally reduce the effective taxation in different ways. Most prominently, it may allow owners to avoid transaction taxes and capital gains taxes in the country where the property is located, as they can trade the shares in the offshore shell corporation rather than the property itself.

We consider a reform in the United Kingdom that targets this type of tax avoidance by extending the taxation of capital gains to "indirect disposals", i.e. the trading of shares in companies whose main assets are property in the United Kingdom. The reform was

announced in November 2017 and took effect from April 2019. It allowed owners within the scope of the new rules to rebase property values so that only capital gains accruing after April 2019 would be subject to taxation.

While the reform aimed to remove one of the main advantages of holding real estate through offshore shell corporations, its effectiveness was limited by an existing double tax treaty with Luxembourg, which prevented the United Kingdom from taxing capital gains made on the disposal of shares in Luxembourg companies. The provision thus created a loophole by effectively exempting capital gains on U.K. properties, realized indirectly through the sales of Luxembourg holding companies, from taxation in the United Kingdom. The loophole was eventually eliminated with Luxembourg joining the OECD multilateral convention amending tax treaties to prevent base erosion and profit shifting and with the United Kingdom and Luxembourg signing a new treaty that explicitly covers the tax treatment of indirect disposals of properties.[16]

These events create useful variation in the tax incentive to hold property through offshore corporations. Before April 2019, offshore corporations in any tax haven would allow owners to avoid taxation of capital gains, but only corporations in Luxembourg could serve this purpose after April 2019. To the extent that tax avoidance is a motive for holding property through offshore corporations, we should therefore expect an increase in holdings through Luxembourg relative to holdings through other tax havens. Specifically, we should expect owners of property to switch to Luxembourg holding structures at some time between November 2017 and April 2019, since this would allow them to continue to avoid capital gains taxes after April 2019.[17]

To test these predictions, we use the database on corporate ownership constructed above to compare the evolution of property titles in England and Wales held by corporations in Luxembourg vs corporations in other tax havens. We make the scope of the analysis as broad as possible by including both residential and non-residential properties and estimate the following model:

$$log(y_{it}) = \sum_{t \neq Oct.2017} \beta^t d_t \times d_i^{Lux} + \gamma_t + \alpha_i + \epsilon_{it} \tag{1.1}$$

[16]We provide more details on the policy reform, the treaty amendment and their implications for taxation of capital gains in Appendix A.4.

[17]To the extent that owners anticipated that the Luxembourg loophole would eventually be closed, it may plausibly have muted the behavioral response.

where y_{it} is the number of property titles held by corporations in country i in month t, d_i^{Lux} is a dummy indicating whether country i is Luxembourg and d_t captures monthly time dummies where the omitted category corresponds to one period before the announcement of the policy in November 2017. The equation includes country fixed effects, α_i, that absorb all cross-sectional variation in holding patterns as well as time fixed effects, γ_t. The estimates of β_t have the flavor of difference-in-differences estimates capturing the percentage change in the number of properties held through Luxembourg since the reference period over and above the percentage change in properties held in other tax havens.

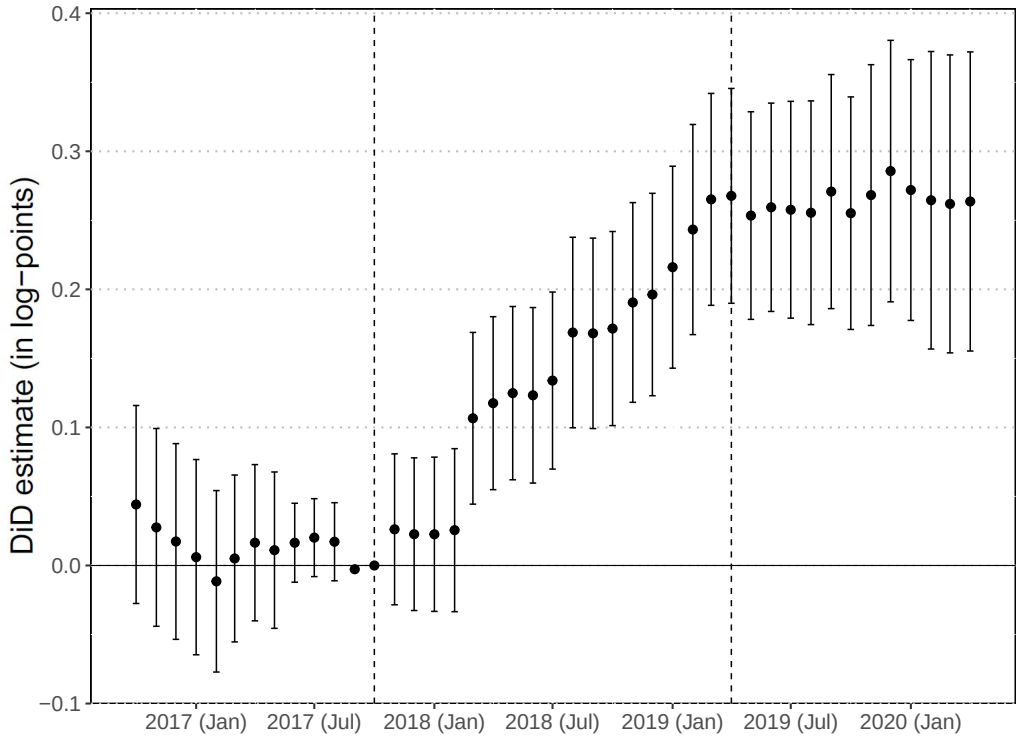
Consistent with tax avoidance being an important motive for holding real estate through offshore corporations, the number of properties held by corporations in Luxembourg increased significantly, around 0.25 log-points, relative to corporations in other tax havens through 2018 and early 2019, as illustrated in Figure 1.6. Importantly, the parallel trends prior to the policy announcement in late 2017 strengthens the case for a causal interpretation of the divergence in early 2018.

In the Appendix, we provide additional evidence on the cross-border flows of property titles underlying these results in Figure A.6. We show that the increase in properties owned through Luxembourg corporations primarily reflects shifting of property titles from corporations in the United Kingdom and nearby tax havens like Jersey and Guernsey. We also provide robustness analysis in Figure A.7. The results are similar when we apply an inverse hyperbolic sine transformation to the outcome rather than the logarithmic transformation in the baseline (Panel A) and when we use no transformation at all (Panel B). They are also robust to measuring property investment in terms of property holdings (i.e. number of postal addresses) rather than property titles (Panel C).

1.4.2 The Secrecy Motive

To investigate the importance of the secrecy motive, we consider a policy initiative targeting the secrecy provided by offshore corporations. In May 2018, the U.K. government announced that it would mandate the Overseas Territories - including well-known tax havens such as Bermuda, the Cayman Islands and the British Virgin Islands - to set up public corporate registers with information about beneficial ownership. Combined with the existing public land registers, the corporate registers would make it possible to trace the ultimate owners of real estate in England and Wales through corporate structures in these jurisdictions and thus make such structures ineffective as vehicles of asset protection and tax evasion. Importantly, the mandate did not affect the Crown Dependencies -

Figure 1.6. – The Tax Motive



Note: The figure plots the estimated coefficients from the baseline analysis of the tax motive (black dots). The estimates show the evolution of property holdings by companies in Luxembourg relative to companies in other tax havens. The omitted time category is October 2017, the month before the extension of the capital gains tax coverage was announced (first dashed line). The second dashed line refers to the implementation of the capital gains tax extension. The confidence bands are at the 95%-level and based on heteroskedasticity and autocorrelation robust standard errors (black vertical lines). For details on the capital gains tax extension, see Appendix A.4.

another group of well-known tax havens comprising Jersey, Guernsey and the Isle of Man - whose constitutional status is different and grants them a higher degree of autonomy. In June 2019, the Crown Dependencies announced that they had decided to set up corporate registers voluntarily.[18]

This policy process creates useful variation in the incentives to use corporations in otherwise similar tax havens to conceal the ultimate ownership of real estate. By May 2018, it was clear that the Overseas Territories would expose shareholders in a public register whereas no such decision was made by the Crown Dependencies before June 2019.

[18]We provide more details on the policy process leading to the creation of corporate registers in the Overseas Territories and the Crown Dependencies in Appendix A.5.

To the extent that secrecy is a motive for holding properties through offshore corporations, we should therefore expect an increase in holdings of properties through the Crown Dependencies relative to holdings through Overseas Territories over this period.[19]

As the nature of the variation in secrecy provisions is similar to the variation in tax provisions studied above, we can use the same empirical approach to study its effect on behavior. In this case, however, it is natural to restrict the sample to the Crown Dependencies and the Overseas Territories so that identification comes from a comparison of tax havens with highly similar legal institutions affected differently by the mandate of the U.K. government. We estimate the following model:

$$log(y_{it}) = \sum_{t \neq Mar.2018} \beta^t d_t \times d_i^{OT} + \gamma_t + \alpha_i + \epsilon_{it}, \tag{1.2}$$

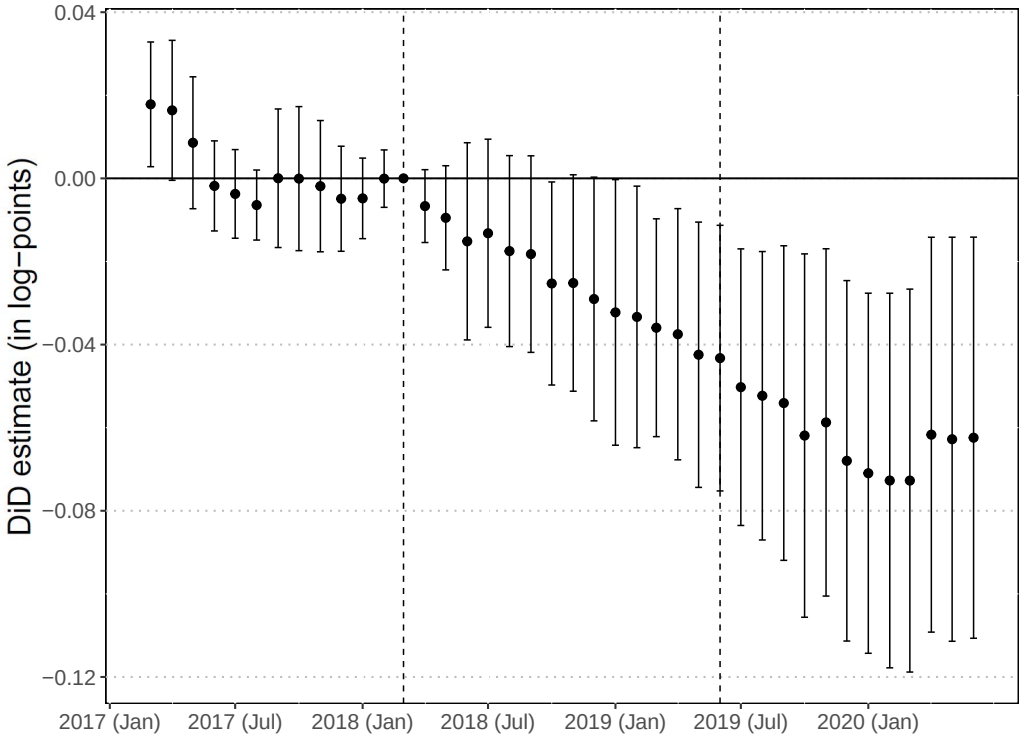
where the notation follows equation (1.1) except that d_i^{OT} is a dummy indicating whether country i is an Overseas Territory. As the government's decision to mandate the register was widely covered in the press already in April 2018, we choose March 2018 to be the omitted time period.

The estimates imply that the number of properties held by corporations in the Crown Dependencies started increasing relative to corporations in the Overseas Territories shortly after the mandate to set up corporate registers in May 2018, as illustrated in Figure 1.7. Consistent with a causal effect of the mandate, the trends in the two outcomes were parallel prior to the mandate. Moreover, there are no clear signs of differential trends after the decision by the Crown Dependencies in June 2019 to set up similar corporate registers. The finding that ownership structures respond sharply to differential changes in transparency across highly similar tax havens suggests that secrecy is an important motive for holding real estate through offshore corporations for at least some investors.

In the Appendix, we report additional results from a robustness analysis in Figure A.8. The tests are analogous to the ones we conducted for the tax shock: We show that the results are robust to alternative transformations (Panels A & B) and units (Panel C).

[19] If there was some anticipation that Crown Dependencies would eventually agree to improve transparency for corporate ownership, it may plausibly have muted the behavioral response.

Figure 1.7. – The Secrecy Motive



Note: The figure plots the coefficients from the baseline analysis of the secrecy motive (black dots). The estimates show the evolution of property holdings by companies in the Overseas Territories relative to companies in the Crown Dependencies. The omitted time category is March 2018, two months before the Sanctions and Anti-Money Laundering Act (SAMLA) amendment passed parliament (first dashed line). The second dashed line refers to the announcement of Crown Dependencies to voluntarily establish public ownership registries. The confidence bands are at the 95%-level and based on heteroskedasticity and autocorrelation robust standard errors (black vertical lines). For details on the Sanctions & Anti-Money Laundering Act (SAMLA) amendment behind this policy change, see Appendix A.5.

1.5 EVIDENCE ON CONSEQUENCES

This section studies the real effects of offshore property ownership using the surprising outcome of the referendum on continued membership of the European Union in 2016 as a natural experiment. The referendum was followed by a surge in property sales by offshore owners, presumably due to uncertainty about the conditions under which foreign nationals could live, work and invest in the United Kingdom after Brexit. Our analysis exploits that similar local property markets were differentially exposed to this shock due to variation in ex ante offshore ownership. Intuitively, when offshore owners sell their properties in response to an adverse change in the economic environment, it amounts to a larger shock in local markets where offshore ownership is initially more prevalent. We first present reduced-form evidence of the real effects of offshore demand shocks by comparing price dynamics around the Brexit referendum across similar local property markets that differ with respect to ex ante offshore ownership (section 1.5.1). We then build a framework

to retrieve a key structural parameter, i.e. the elasticity of property prices with respect to changes in offshore demand (section 1.5.2). Finally, we consider the effect of offshore ownership on vacancy rates (section 1.5.3).

1.5.1 Reduced-Form Evidence

Empirical Framework. We investigate whether the property price dynamics around the Brexit referendum are affected by ex ante offshore ownership. Indexing individual properties by i, 983 local property markets in London by m and calendar years by t, we estimate the following transaction-level model for residential properties in London in the period 2011-2019:

$$log(p_{it}) = \mu_i + \sum_{t \neq 2015} \Gamma^t d_t \times \boldsymbol{x_i} + \sum_{t \neq 2015} \beta^t d_t \times z_m^{\text{Offshore}} + \varepsilon_{it} \tag{1.3}$$

where the dependent variable is the transaction price for property i observed in year t; μ_i denotes property fixed effects; d_t denotes year dummies (the omitted category is 2015); z_m^{Offshore} denotes the ex ante share of the local property market owned by corporations in tax havens (measured end of 2015); and $\boldsymbol{x_i}$ denotes a vector of time-invariant controls at the property-level. In the baseline specification, $\boldsymbol{x_i}$ includes indicators for the percentiles of the ex ante national distribution of property prices (measured in 2015).[20] In robustness tests, we augment the model with additional controls.

The first and second terms of the model capture the underlying price dynamics within each of the 100 price categories. It is identified by comparing transaction prices for the same property in different years. The third term captures the key object of interest: the differential price dynamics in local property markets with more offshore ownership. It is identified by comparing within-property changes in transaction prices across properties that belong to the same narrow price segment but are located in local property markets within London with different ex ante shares of offshore ownership.

Main Results. Figure 1.8 illustrates the elements in the vector β. The post-2015 elements are all negative suggesting that prices decreased more following the Brexit referendum in local property markets with a higher ex ante share of offshore ownership. A causal interpretation of this result requires the identifying assumption that price trends in the

[20]Specifically, we consider each property's last transaction price prior to the Brexit referendum and determine the position of this price in the national distribution of transaction prices in the same quarter.

post-2016 period would have been parallel across local property markets with different ex ante shares of offshore ownership absent the Brexit shock (within price segments). This assumption concerns a counterfactual and is not directly testable, but the parallel trends in the pre-2015 period makes it credible.

The estimated reduced-form effects are economically and statistically significant. The point estimate of around -1 in 2018 suggests that property prices dropped by around 1% over the period 2015-2018 for each additional 1 percentage point of ex ante local offshore ownership. This implies that prices were around 1.8% lower in 2018 in local areas at the 75th percentile of offshore ownership shares, around 4.4% lower at the 90th percentile, and around 8.2% lower at the 95th percentile.[21]

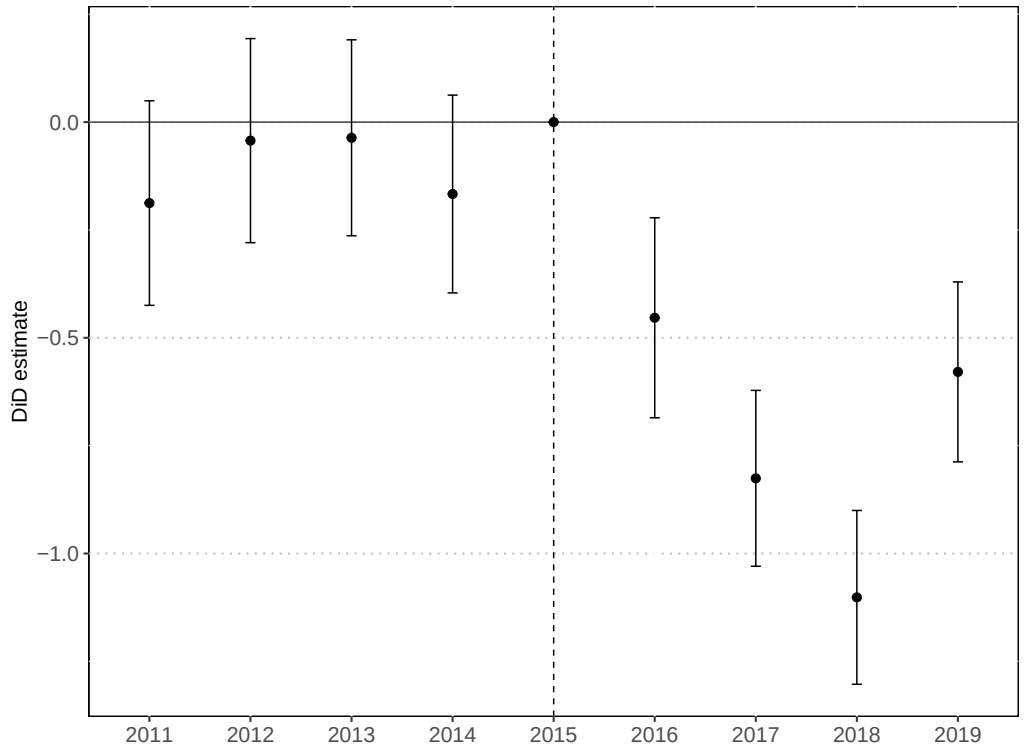
Robustness. Table 1.3 reports the results from additional robustness tests conducted in a compact framework where the vector d_t is replaced with a simple post-2015 indicator. The estimated coefficient captures the differential price change between the pre-referendum period (2011-2015) and the post-referendum period (2016-2019) associated with an additional 1 percentage point of ex ante local offshore ownership.

First, we augment the estimating equation with additional area-level controls that absorb potential confounding factors. We add the ex ante share of the local property market with any corporate ownership (interacted with time dummies). As shown in Column (2), area-level corporate ownership is not generally associated with larger price decreases in the post-referendum period. By contrast, when controlling for overall corporate ownership, the effect of offshore corporate ownership is even stronger. We also add the ex ante share of foreigners in the local population (interacted with time dummies). As shown in Column (3), the results suggests that the presence of foreigners was associated with price decreases, for instance reflecting migration responses to Brexit; however, the estimated effect of offshore ownership barely changes.

Next, we investigate whether the results are driven by offshore corporations with foreign beneficial owners, i.e. those directly affected by Brexit. For some of the offshore corporations in the sample, we observe the beneficial owners directly in the leaks (see section 1.2.3) and we rely on this information to impute beneficial ownership patterns for the re-

[21] Figure A.9 in the Appendix illustrates the distribution of ex ante offshore ownership shares across the 939 local property markets in London with non-zero offshore ownership. In most of these areas, offshore ownership accounts for less than 2% of the local market, but in a handful of areas this metric exceeds 10%.

Figure 1.8. – Offshore Ownership and Property Prices around Brexit



Note: The figure plots the coefficients from the baseline analysis of offshore ownership and property prices around the Brexit referendum (black dots). The coefficients show the estimated change in property prices relative to 2015 for each additional percentage point of offshore ownership in the local property market in 2015. The local property markets correspond to the 983 middle output areas of London. The regression is at the property-level and the sample is limited to properties in London transacted both before and after Brexit. The baseline regression includes property fixed effects, so that the underlying price dynamics is identified from repeat sales, and 100 indicators for the ex ante price level interacted with calendar time, so that the effect of offshore ownership is identified from comparisons of within-property price changes across similarly-priced properties located in property markets with different shares of offshore ownership. The confidence bands are at the 95%-level and based on robust standard errors clustered at the property level (black vertical lines).

mainder.[22] This allows us to re-estimate the baseline model while conditioning the effect of offshore ownership on foreign vs domestic beneficial ownership. As shown in Column (4), the estimates suggest that offshore ownership was associated with a differential price decrease after the Brexit referendum only when the beneficial owners were foreign.[23]

[22]Assuming random selection into the leaks within each offshore jurisdiction (Alstadsæter et al., 2019), we assign probabilities of foreign and domestic beneficial ownership to the corporations that are not in the leaks based on the ownership patterns observed in the leaks. For instance, observing that beneficial owners are foreign for 58.6% of the Jersey corporations and 94.1% of the corporations from the Seychelles in the leaks, the imputed probability of foreign beneficial ownership is 58.6% for Jersey corporations and 94.1% for Seychelles corporations not in the leaks.

[23]The reported standard errors should be interpreted with some caution, as they do not account for the fact that the two dependent variables are themselves estimates associated with uncertainty.

Finally, we examine the robustness to the treatment of outliers and serial correlation in the data. As shown in Columns (5)-(6), winsorizing at alternative levels makes the estimated effect of offshore ownership slightly more negative. As shown in Column (7), clustering standard errors more conservatively widens the standard errors, but the key estimate remains highly significant.

1.5.2 Retrieving the Structural Parameter

Empirical Framework. Ultimately, we would like to retrieve a parameter that captures the effect of offshore ownership on property prices. This is not straightforward as offshore ownership is generally endogenous to concurrent price developments. We address this challenge by harnessing the exogenous variation in offshore ownership created by the Brexit referendum in an instrumental variable framework. Specifically, the exposure of local property markets to sales by offshore owners following the referendum was highly heterogeneous and largely mechanical, reflecting the pre-determined share of offshore owners in the market just before the referendum.

It is useful to start from the following structural equation:

$$\frac{\Delta p_i}{p_i} = \alpha + \theta \frac{\Delta D_m^{\text{offshore}}}{D_m} + \varepsilon_{im} \tag{1.4}$$

which relates a percentage change in prices to a percentage change in property demand where the demand change comes from offshore corporations. Conceptually, we think of demand from offshore corporations as their property purchases net of their sales. We refer to the parameter θ as the elasticity of property prices with respect to offshore demand.

We propose to use the share of offshore ownership in the local property market before the Brexit referendum, z_m^{offshore}, as an instrument for the change in offshore demand after the referendum. Several conditions need to be satisfied for it to be valid. First, it must affect the change in offshore demand (relevance). Second, it must not correlate with other shocks to prices (random assignment). Third, it must not affect prices in other ways than through the change in offshore demand (exclusion restriction).

Intuitively, we expect the instrument to be relevant because a general surge in property sales by offshore owners, in this case triggered by the prospect of Brexit, mechanically increases sales more in local property markets where the offshore ownership share is initially high. Figure 1.9 illustrates this pattern in the data by plotting, for each quintile of the ex ante offshore ownership share, the change in sales by offshore owners. While sales

Table 1.3. – Offshore Ownership and Property Prices, Robustness

	(1) **Baseline**	(2) **Additional Area Controls**	(3)	(4) **Beneficial Ownership**	(5) **Alternative Winsorization**	(6)	(7) **Alternative Clustering**
Post x Offshore	-0.701***	-1.06***	-0.659***		-0.820***	-1.09***	-0.701***
	(0.069)	(0.087)	(0.070)		(0.068)	(0.078)	(0.153)
Post x Corporate		0.119***					
		(0.017)					
Post x Foreign Population			-0.043***				
			(0.011)				
Post x Offshore, Foreign (exp.)				-1.24***			
				(0.130)			
Post x Offshore, Not Foreign (exp.)				0.014			
				(0.100)			
Property FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
100 price bins × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustering	Property	Property	Property	Property	Property	Property	983 areas
Winsorization					2x0.5%	2x2%	
Observations	143,308	143,308	143,308	143,308	143,308	143,308	143,308
Adjusted Adjusted R^2	0.94977	0.94982	0.94978	0.94981	0.94978	0.94979	0.94977

Note: This table shows difference-in-differences estimates for our reduced-form price analysis. The sample is identical to the event study presented in Figure 1.8, i.e. we focus on property sales between 2011 and 2019 and limit the sample to properties sold both before and after the Brexit referendum. Post indicates that the year is 2016 or later. Column 1 shows the results from the baseline model. Column 2 adds a control for the total corporate market share (domestic and foreign corporations). Column 3 adds a control for the foreign population share (defined at the district level). Column 4 conditions the effect of offshore ownership on the expected residence of the beneficial owners (Foreign vs British). Columns 5-6 show results with alternative winsorization (0.5% and 2% level). Column 7 shifts the clustering from the property level to the local property market level (983 local areas). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

by offshore owners increase after the referendum in all quintiles except the bottom one, the increase is much more pronounced in local property markets with a higher ex ante share of offshore ownership.

To strengthen the case for random assignment, we add a vector of property-level controls $\boldsymbol{x_i}$ to equation 1.4. The control vector includes 100 indicators for the ex ante price level (same as in equation 1.3). Thus, the assumption is that local property markets in London *within the same narrow price segment* were not exposed to other shocks varying systematically with the offshore ownership share around the Brexit referendum. The control vector also includes controls for the timing of the property transactions. By construction, all properties in the sample are transacted pre-referendum (2011-2015) and again post-referendum (2016-2019). We include indicators for all combinations of transaction years, e.g. 2013-2018, 2014-2016 and 2015-2018.

It is useful to note that the IV system is closely related to eq. 1.3 except that it is formulated in differences rather than in levels. Specifically, eq. 1.3 is essentially the reduced form of the IV system, relating the instrument (ex ante offshore ownership) directly to the outcome (prices). While the reduced form only requires random assignment (conditional on controls), the IV system comes with an additional assumption namely the exclusion restriction.

Results. Table 1.4 reports the estimates for the full IV system. The reduced-form estimate is -0.728, which is very close to the baseline reported in Table 1.3 as expected. It implies that property prices dropped by around 0.7%, pre-referendum to post-referendum, for each additional 1 percentage point of local offshore ownership. The small discrepancy relative to the baseline reflects that we now estimate a cross-sectional model in differences rather than a multi-year panel model in levels.

The first stage has the expected negative sign and is highly statistically significant (F-stat: 37.36). The point estimate implies that offshore corporations' property demand (i.e. their purchases net of their sales) dropped by 0.54% of total demand in the market (i.e. the ex ante value of property purchases) for each additional 1 percentage point of local offshore ownership.

The second stage delivers the structural parameter of interest: the elasticity of property prices with respect to offshore demand. The point estimate suggests that a drop in property demand from offshore corporations equivalent to 1% of total demand reduces prices by around 1.3%.

Table 1.4. – Offshore Ownership and Property Prices, Structural Parameter

	(1) **Reduced Form**	(2) **First Stage**	(3) **Second Stage**
Offshore Market	-0.728***	-0.540***	
	(0.167)	(0.088)	
Demand Shock			1.35***
			(0.317)
100 price bins x pre-post	Yes	Yes	Yes
Clustering	local area	local area	local area
Observations	62,638	62,638	62,638
Adjusted R^2	0.2623	0.1570	0.2312
Kleibergen-Paap F-Statistic		37.36	

Note: This table reports the estimation results for our IV system. The dependent variable in the reduced form and the second stage is the property-level change in transaction prices (in logs) between the most recent transaction before the Brexit referendum and the first transaction after. The exogenous instrument is the share of local property market m owned by offshore corporations end of 2015. The endogenous variable is the change in average offshore demand in market m between 2011-2015 and 2016-2019 as a share of total demand in market m 2011-2015 where offshore demand is the value of property purchases net of property sales by offshore corporations and total demand is the value of all property transactions. The vector of control variables in all three regressions includes 100 indicators for the ex ante price level (computed as in the main regression) and 20 indicators for the timing of the property transactions, e.g. 2013-2018, 2014-2016 and 2015-2018. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

The large elasticity estimate suggests that property prices are highly sensitive to offshore demand, which largely constitutes new flows into the property market by foreign investors. This contrasts with domestic demand where property buyers are often also property sellers, which dampens the effect on prices. Viewed in this light, the large elasticity estimate resonates with recent research showing that new flows into the stock market have large effects on stock prices (Gabaix and Koijen, 2021).

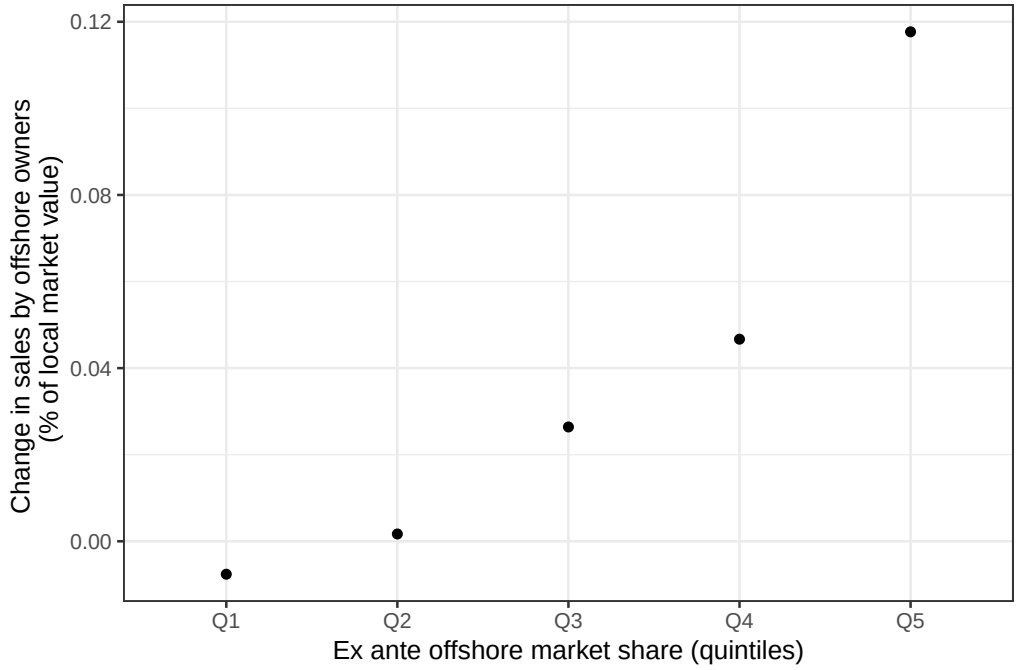
1.5.3 Vacancies

Finally, we investigate whether offshore ownership is associated with real effects in the form of vacancies. From the perspective of urban development, the emergence of ghost properties owned by absentee residents is a salient risk associated with foreign property investments motivated by tax and secrecy rather than by residence.

Our analysis uses administrative data on vacancies available at the level of local authority districts.[24] The data originates from the collection of council taxes on properties levied to fund local services. Councils require information about vacancies because the tax

[24]In particular, we relate the number of vacant dwellings by local authority district (Table 615) to the total dwelling stock (Table 125) from the UK government statistics to estimate the vacancy rate for each district.

Figure 1.9. – Offshore Capital Flight after the Brexit Referendum



Note: The figure illustrates how the exposure of local property markets to sales by offshore owners after the Brexit referendum correlates with the ex ante offshore market share. For each local property market, we compute the change in the average annual property sales by offshore corporations to other investors from before the referendum (2011-2015) to after the referendum (2016-2019), with sales scaled by the aggregate value of the local property market. The figure plots the average of this metric by quintile of the ex ante offshore market share.

treatment varies with the vacancy status. Vacant dwellings usually benefit from a tax discount, but if a dwelling has been empty for more than two years, a tax penalty may apply.[25]

Our main specification mirrors the reduced-form framework we used to study price effects, except that vacancy information is at the area-level rather than the property-level:

$$\mathrm{V}_{at} = \mu_a + \sum_t \Gamma^t d_t \times \boldsymbol{x_a} + \beta Post_t \times z_a^{\text{Offshore}} + \varepsilon_{at}$$

where V_{at} is the vacancy rate in local authority district a in year t, z_a^{Offshore} denotes the ex ante share of the property market in district a owned by corporations in tax havens and $Post_t$ is an indicator for years after 2015. The specification allows for different underlying

[25]For details, see information on the tax treatment of empty properties provided by the UK government.

trends in vacancy rates across price segments. In particular x_a includes 10 indicators of the ex ante district-level price level, each of which is interacted with time dummies.[26] We estimate the model with annual data for 33 local authority districts in London.

Table 1.5 displays the results from the main specification in Panel A. The estimates suggest that the vacancy rate decreased by 0.11 percentage points after the Brexit referendum for each additional 1 percentage point of ex ante local offshore ownership (Column 1). The estimate compares to a baseline average vacancy rate of 1.7%. The estimate is robust to controlling for the foreign population share and corporate ownership (Columns 2-3). The reduced-form estimates are consistent with the notion that the property sales by offshore owners induced by the Brexit referendum lowered the prevalence of vacant dwellings in London.

To gauge the share of dwellings with offshore owners that are vacant, we use an alternative specification that exploits the time variation in the offshore ownership share:

$$\mathrm{V}_{at} = \alpha_a + \mu_t + \omega \mathrm{Offshore}_{at} + \varepsilon_{at}$$

$\mathrm{Offshore}_{at}$ is the share of the property market in district a owned by corporations in tax havens in year t. The specification includes fixed effects for local districts (α_a) and years (μ_t). The coefficient of interest, ω, is identified by comparing the change in vacancy rates across districts with different changes in the offshore ownership share. Because both the dependent variable and the main explanatory are expressed in shares of the market, we can approximately interpret β as the fraction of properties that become vacant when acquired by offshore owners.[27]

Table 1.5 displays the results from this specification in Panel B. The baseline estimate suggests that increasing the offshore ownership share by 1 percentage point is associated with an increase in the vacancy rate by just above 0.2 percentage points (Column 1). The estimate is robust to controls for the foreign population share and corporate ownership (Columns 2-3). Taken at face value, the estimate suggests that one out of five properties owned through offshore corporations is vacant.

[26] We calculate the mean transaction price for each local authority district in England and Wales over the period 2011-2015, compute the deciles of the distribution of mean transaction prices, assign the districts in our sample to 10 income groups defined by the deciles and construct indicators for each of the groups.

[27] The interpretation is only approximate as $\mathrm{Vacancy}_{at}$ refers to the vacant share of housing units whereas $\mathrm{Offshore}_{at}$ refers to the offshore-owned share of housing values. Identification could be much improved with vacancy information at the property-level.

Table 1.5. – Offshore Ownership and Vacancies

	(1) **Baseline**	(2) **Foreign Pop.**	(3) **Dom. Corp**
Panel A: Brexit Identification			
$Post \times z^{\text{Offshore}}$	-0.11***	-0.11***	-0.12**
	(0.027)	(0.030)	(0.047)
District FE	Yes	Yes	Yes
10 price bins × year FE	Yes	Yes	Yes
Post × Foreign Pop.		Yes	
Post × Dom. Corp.			Yes
Observations	297	297	297
Adjusted R^2	0.6157	0.615	0.6156
Panel B: Variation over Time			
Offshore	0.20***	0.21***	0.21***
	(0.071)	(0.068)	(0.063)
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Foreign Pop.		Yes	
Dom. Corp.			Yes
Observations	297	297	297
Adjusted R^2	0.5725	0.5754	0.5763

Note: This table shows the regression results for the analysis of vacancy rates. Across all specifications, we use annual observations for the period 2011-2019 for each of the 33 districts in London. The dependent variable is the vacancy rate, i.e. the share of dwellings in the district that are vacant. *Post* indicates that the year is 2016 or later, z^{Offshore} is the share of the local property market owned by offshore corporations in 2015 and Offshore is the share of the local property market owned by offshore corporations in the current year. Panel A shows results from a specification that mirrors the price analysis and exploits the Brexit referendum as a shock that changes the ownership structure of properties in London. The coefficient on $Post \times z^{\text{Offshore}}$ indicates the differential change in the vacancy rate associated with a one percentage point higher ex ante offshore market share. Panel B shows results from a specification that exploits longitudinal variation in offshore ownership. The coefficient of interest on *Offshore* indicates how the vacancy rate changes when the offshore market share increases by one percentage point. Column (1) displays the baseline results. Column (2) adds controls for the foreign population. Column (3) adds controls for domestic corporate ownership. The unweighted average vacancy rate across all London districts in 2015 was around 1.7%. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

1.6 CONCLUSION

The role of offshore corporations in the U.K. residential property market has increased dramatically over the two most recent decades. It currently exceeds 1.5% in the overall residential market and reaches 15% for the most expensive properties.

Our analysis suggests that taxation and secrecy are important motives for funneling property investments through offshore tax havens. Changes in tax rules and disclosure requirements induce sharp changes in offshore ownership structures. Seemingly, offshore ownership can serve to avoid taxes on capital gains and to conceal the identity of the ultimate owners.

Our analysis also shows that the rise of offshore ownership has significant impacts on real outcomes in property markets. The Brexit referendum triggered a sharp increase in property sales by offshore owners accompanied by differential decreases in prices and vacancy rates in areas with higher ex ante offshore ownership rates. Offshore ownership appears to drive prices up and leads to more empty dwellings.

Chapter 2

Pareto-Improvements, Welfare Trade-Offs, and the Taxation of Couples

This chapter is based on co-authored work with Felix Bierbrauer, Pierre Boyer, and Andreas Peichl. A previous version has been published as a CEPR Discussion Paper. See Bierbrauer et al. (2024) for the full reference.

2.1 INTRODUCTION

This paper has a pure theory part, an applied theory part, and an empirical part. In the pure theory part, we develop a new approach for the identification of Pareto-improving and welfare-improving tax reforms in a setting with the following features: There is a tax function that potentially has many arguments. Moreover, there is multidimensional heterogeneity amongst taxpayers and every economic decision may give rise to behavioral responses both at the intensive and the extensive margin. Under the given tax system, there may be discontinuities in marginal tax rates. For all these reasons, the optimal choices of taxpayers may not satisfy first-order conditions and behavioral responses may therefore be discontinuous.

We focus on one particular application, the taxation of couples. We take as given a status quo tax system, but we do not make a priori assumptions on its functional form. We also assume that, for each spouse, "making money" comes with both fixed and variable effort costs. Depending on these costs and the properties of the tax system, couples sort themselves into the set of dual-earner couples, the set of single-earner couples, or the set of couples with no earnings. This sorting changes when the tax function is changed. In addition, earnings choices change. In the applied theory part of the paper, we combine a detailed analysis of behavioral responses in this setting with our general theory insights. This approach yields sufficient statistics that can be used to identify Pareto- or welfare-improving tax reforms in the data.

We use these sufficient statistics for an evaluation of the tax treatment of couples in the US. Our analysis is motivated by recurrent debates about reforms of the tax and transfer system with the objective to achieve more gender balance. In particular, there is the concern that current systems excessively discourage the labor supply of secondary earners, most of whom are women, see e.g. Bick and Fuchs-Schündeln (2017). Our analysis (i) identifies the parts of the income distribution with inefficiently high tax rates on secondary earnings, (ii) it shows which welfare functions approve and which ones reject reforms that lower the tax system's degree of jointness and/ or strengthen workfare elements at the bottom of the income distribution. We also present a calibration of optimal marginal tax rates for primary and secondary earners at the top of the income distribution.

Pure Theory. The paper's theory part is based on the following setting: There is a continuum of couples. A couple consists of two spouses with earnings of y_1 and y_2. Couples are confronted with a tax function $(y_1, y_2) \mapsto T(y_1, y_2)$. This tax function may be inherited from the past or it may be an object of theoretical interest such as an optimal

tax function.[1] Couples choose y_1 and y_2 optimally given the tax system. Specifically, we assume that the spouses maximize their joint surplus, i.e. the difference between their disposable income and their effort costs.[2]

Methodologically, our key innovation is the use of *conditional revenue functions*. Consider the following thought experiment. Focus on all couples who have primary earnings in a narrow bracket B_1 that starts at an income level y_1. Then focus on the subset of these couples who have secondary earnings in a narrow bracket B_2 that starts at income level y_2. For these couples, consider a small increase of the marginal tax rate on secondary earnings and denote the reform's marginal effect on tax revenue by $\mathcal{R}_2(y_2 \mid y_1)$. The conditional revenue function $y_2 \mapsto \mathcal{R}_2(y_2 \mid y_1)$ documents how this revenue effect varies as we vary the position of B_2 in the range of incomes while holding B_1 fixed. To see how such functions can help to identify inefficiencies in the tax system, suppose that the status quo is a system of progressive and joint taxation, as in France, Germany and the US. Consequently, the larger is y_1, the larger is, for every level of y_2, the marginal tax rate on secondary earnings. For very high y_1, this may discourage secondary earners so much that lower marginal taxes would be actually be self-financing. In this case $\mathcal{R}_2(y_2 \mid y_1)$ takes a negative value which is analogous to tax rates on secondary earnings being above the top of the "Laffer curve", except that the Laffer curve here allows for non-linear taxes and uses primary earnings y_1 as a conditioning variable. A reform that lowers the marginal tax rate on secondary earnings for couples with primary earnings close to y_1 and secondary earnings close to y_2 would then be Pareto-improving. If the status quo has progressive and joint taxation, then the conditioning on y_1 has the potential to make an important difference. For lower values of y_1 marginal tax rates are lower. Hence, there may not be an inefficiency.

The theory part of this paper contains three theorems. They all invoke conditional revenue functions for dual-earner and single-earner couples. Theorems 1 and 2 state, respectively, necessary and sufficient conditions for the Pareto-efficiency of a tax system. Theorem 3 characterizes the welfare implications of tax reforms. It can be used to analyze the extent to which alternative social welfare functions give rise to conflicting views about desirable directions for reform. In particular, Corollary 3 shows that a reduction of marginal tax rates for secondary earners is rejected by a Rawlsian welfare measure if the starting point

[1] Stylized alternatives are individual and joint taxation. Under the former, the marginal tax rate on the earnings of one spouse does not depend on the earnings of the other spouse. Under the latter, and with a progressive tax system, the marginal tax rate on the earnings of one spouse is increasing in the earnings of the other spouse.

[2] The maximization of the joint surplus is, under some ancillary conditions, consistent with the spouses engaging in Nash bargaining, see Appendix B.1.10.

is a tax system that satisfies the sufficient conditions for Pareto-efficiency in Theorem 2. Moreover, Theorem 3 lends itself to a characterization of welfare-optimal tax systems (Corollary 2). It can also be used to analyze the welfare implications of reforms that alter the tax system's degree of jointness (Corollary 4).

Applied Theory. The general analysis that leads to these Theorems is complemented by a detailed analysis of a couple's behavioral responses to taxation. This analysis leads to a sufficient statistics characterization of the conditional revenue functions for dual- and single-earner couples. With this characterization we can turn to a more detailed analysis of the conditions under which changes of tax rates are desirable. We find conditions for the desirability of negative marginal tax rates that are unprecedented by the previous literature on optimal taxation: When welfare weights are continuous and decreasing in income, then, at the bottom of the income distribution, negative marginal tax rates are desirable even when extensive margin elasticities are zero. We also state conditions under which an optimal tax system is piecewise linear and separable at the top of the income distribution. Thus, under these conditions, small changes in the earnings of one spouse should not affect the marginal tax rate on the earnings of the other spouse.

Empirical Analysis. The empirical part of this paper uses these insights to (i) identify Pareto-improving reforms of the current US tax system, (ii) trace out the welfare implications of changes in the US tax system's degree of jointness, (iii) discuss reforms of the tax-and transfer system that affect low-income couples, and (iv) characterize optimal marginal tax rates for primary and secondary earners at the top of the income distribution.

For this purpose, we combine data on the income distribution within and across married couples from the Current Population Survey (CPS) with information on marginal tax rates from the NBER TAXSIM microsimulation model. In addition, we use tax return data to estimate the properties of the primary and secondary earnings distribution at the top.

We present four main empirical findings. First, for the recent past, the taxation of couples is associated with inefficiently high marginal tax rates for secondary earners. Identifying these inefficiencies requires the use of conditional revenue functions. If primary earnings were averaged out and an unconditional revenue function was used, many of these inefficiencies would remain undetected. The policy implication is that, to reap those efficiency gains, the reduction of marginal tax rates for secondary earners needs to depend also on the level of primary earnings.

Second, reducing the tax system's degree of jointness is not a Pareto-improving reform. It comes with a loss of tax revenue and is therefore rejected by a Rawlsian welfare function. It is approved, however, by a welfare function with weights that are increasing in the women's income share in the couple's total income. Thus, judging the tax system's degree of jointness requires to take a stance on the what the relevant welfare objective looks like. There is a trade-off between poverty alleviation and gender balance.

Third, we identify ranges of incomes at the bottom of the income distribution where reforms towards lower marginal tax rates are "controversial" in the sense that they are approved by welfare functions with continuous weights – among them welfare functions that, say, concentrate weights on the bottom ten percent of the income distribution – while they are rejected by a Rawlsian welfare function.

Finally, there is a substantial gap in optimal top tax rates for primary and secondary earners, even if the elasticities that are capturing the behavioral responses to taxation are the same and the welfare weights for all these people at the top of the income distribution are identical. The differences are driven by the tails of the income distributions. The distribution of primary earnings has a fatter tail so that optimal top tax rates are around 12 percentage points higher for primary earners. In addition, we find that the top tax rates for primary earners in single-earner couples should be larger than for primary earners in dual-earner couples.

Related Literature. There is a rich literature that studies the optimal taxation of couples. Its starting point is the seminal paper by Boskin and Sheshinski (1983) which applied an inverse elasticities logic to the taxation of couples: When secondary earners show behavioral responses to taxation that are stronger than those of primary earners, then it is optimal to tax secondary earnings at a lower rate.[3] The subsequent literature has branched out in numerous ways.[4] Our approach differs in that we analyze reform directions in a neighborhood of a status quo tax system. While we explore some implications of our approach for optimal taxes, we are primarily interested in identifying Pareto- or welfare-improving reforms of an actual tax system.

[3]For a related discussion of gender-based taxation, see Alesina et al. (2011).

[4]Non-linear taxes have been considered, labor supply responses at the intensive and the extensive margin have been taken into account, marital status has been treated as endogenous, see e.g. Schroyen (2003), Brett (2009), Kleven et al. (2009), Immervoll et al. (2011), Cremer et al. (2012), Gayle and Shephard (2019), Malkov (2021), Alves et al. (2024), Ales and Sleet (2022), and, most recently, Golosov and Krasikov (2025).

Our formal analysis is based on a model with multi-dimensional heterogeneity amongst couples. Couples differ in the productive abilities of the primary and the secondary earner, in their fixed costs of labor market participation and, possibly, in their weights in the couple's internal bargaining procedure.[5] This framework is richer than what has previously been considered in papers that approach the optimal taxation of couples as a problem of multi-dimensional screening. Kleven et al. (2009) focus on a setting in which a primary earner only makes intensive margin choices and a secondary earner only makes an extensive margin choice. Golosov and Krasikov (2025) focus on spouses who both only make intensive margin choices.[6] Both Kleven et al. (2009) and Golosov and Krasikov (2025) present results on optimal jointness, i.e. on the optimal interdependence of marginal tax rates on primary and secondary earnings. We complement these findings with an analysis of reforms that alter a given tax system's degree of jointness. In particular, if the status quo has progressive and joint taxation, then a jointness-reducing reform can be interpreted as bringing the status quo closer to individual taxation. We identify the conditions under which such reforms can be Pareto-improving and the conditions under which such reforms come with welfare trade-offs.

We use perturbation techniques to identify reform directions that are efficiency-enhancing, or welfare-improving, starting from a given status quo. The perturbation approach is frequently used in analyses of income taxes.[7] In this paper, we extend this approach in a substantial way so that we can explicitly deal with the interdependence of primary and secondary earnings that arises through joint taxation. Problems of optimal multi-dimensional screening are known to give rise to complex patterns of bunching.[8] Our theoretical approach allows for the possibility of bunching and for the possibility of non-continuous behavioral responses to changes in marginal tax rates. We remain agnostic, however, whether these patterns arise because the status quo is an optimal tax system or because the status quo has discontinuities in marginal tax rates for other reasons. Our approach takes account, however, of the possibility that bunching regions or regions with discontinuous behavioral responses change their size in response to a tax perturbation. If the status quo is an optimal tax system, then no such perturbation can have a positive effect on welfare.

[5]See Appendix B.1.10 for a discussion of Nash bargaining in couples.

[6]In an extension, they show that fixed costs of working are without consequence for optimal taxes when there is random matching on the marriage market.

[7]References include Piketty (1997), Saez (2001), Golosov et al. (2014), Saez and Stantcheva (2016), Sachs et al. (2020), Bierbrauer et al. (2021), Jacquet and Lehmann (2021b), Jacquet and Lehmann (2021a), Bergstrom and Dodds (2021), Bierbrauer et al. (2023a), Ferey et al. (2024) and Spiritus et al. (forthcoming).

[8]For recent progress on the characterization of optimal bunching patterns, see Boerma et al. (2022), Carlier et al. (2024), see Rochet (2024) for a review of the literature.

Our empirical analysis employs the NBER TAXSIM microsimulation model and CPS micro data.[9] The microsimulation model uses rich data on individual characteristics so that we can elicit, at the level of an individual tax unit, the marginal tax rates and tax liability under the status quo and any alternative tax system. The evaluation of tax reforms rests on empirical estimates of the behavioral responses to taxation. A large literature exists estimating relevant elasticities and our assumptions on behavioral responses are informed by this literature.[10]

Our analysis yields predictions on how tax reforms would affect earnings incentives of the spouses in a couple, aggregate measures of welfare and tax revenue. A complementary literature in macroeconomics embeds the joint labor supply decisions of couples into quantitative dynamic models. It then traces out the tax and transfer system's implications for the labor market outcomes and the savings decisions of men and women or tax revenue.[11]

In a companion paper, Bierbrauer et al. (2025), we analyze the taxation of couples from a political economy perspective, with a focus on how the support for reforms towards individual taxation has evolved since the 1960s. The paper also contains an analysis of past reforms that affected the distribution of marriage bonuses. It is related to this paper in that it also uses revenue functions for an analysis of tax reforms (as have other papers). There is, however, an important methodological difference: The companion paper is tailored to a status quo with joint taxation and it does not use conditional revenue functions (which are the key innovation from this paper). The analysis in this paper does not restrict the set of tax functions. Therefore, its analysis of Pareto- and welfare-improving reforms is complete in the sense of identifying necessary and sufficient conditions for the existence of such reforms.[12]

Outline. Section 2.2 contains the theoretical analysis, and Section 2.3 the empirical results. The last section contains concluding remarks. Formal proofs and supplementary analyses are relegated to the Appendix.

---

[9] Our empirical approach builds on and extends work by Eissa et al. (2008), Bargain et al. (2015) and Bierbrauer et al. (2021).

[10] See, e.g., Blundell and Macurdy (1999), Blau and Kahn (2007), Eissa and Hoynes (2004), LaLumia (2008), Bargain et al. (2014).

[11] See, e.g., Guner et al. (2012), Bick and Fuchs-Schündeln (2017), Borella et al. (2023a) or Holter et al. (2023).

[12] It also transcends our earlier discussion paper, Bierbrauer et al. (2023b) which only covered some sufficient conditions when using unconditional revenue functions.

2.2 THEORY

The Status Quo. There is a status quo tax system for married couples $(y_1, y_2) \mapsto T^0(y_1, y_2)$ where y_1 are the earnings of spouse 1 and y_2 are the earnings of spouse 2.[13] To save on notation, we often write $y = (y_1, y_2)$. We assume that T^0 is a continuous function and write $T^0_{y_1}(y)$ and $T^0_{y_2}(y)$ for the marginal tax rates on the earnings of spouse 1 and spouse 2 at y. We write $T^0_{y_1,y_2}(y)$ and $T^0_{y_2,y_1}(y)$ for the tax function's cross derivatives.[14]

There is a status quo distribution of incomes. In the status quo y takes values in $\mathcal{Y}_0 = \mathcal{Y}_{0,1} \times \mathcal{Y}_{0,2}$, where, for $i = 1, 2$, $\mathcal{Y}_i = (0, \bar{y}_i]$ is the range of strictly positive earnings and $Y_{0,i} = [0, \bar{y}_i]$ includes earnings of zero. We denote by $F_y : \mathcal{Y}_0 \to [0, 1]$ the *cdf* the represents the joint distribution of incomes in the status quo, and by f_y its density. We use obvious notation for the conditional and marginal distributions that can be derived from F_y. We interpret the status quo distribution of incomes as resulting from the utility-maximizing behavior of couples. Specifically, we assume that the spouses in a married couple solve the following problem: Choose y to maximize $C^0(y) - K(y, \theta)$, where the couple's disposable income is given by

$$C^0(y) \quad = \quad b^0_m \; + \; y_1 + y_2 \; - \; T^0(y) \; ,$$

and its costs of productive effort are given by

$$K(y, \theta) \quad = \quad k_1(y_1, \omega_1) \; + \; \varphi_1 \, \mathbf{1}(y_1 > 0) \; + \; k_2(y_2, \omega_2) \; + \; \varphi_2 \, \mathbf{1}(y_2 > 0) \; .$$

We assume that $T^0(0, 0) = 0$. Hence, b^0_m is the intercept of the consumption schedule that couples are facing in the status quo, or, equivalently, the transfer to a couple with no earnings. For each spouse i, the generation of earnings comes with fixed and variable costs of productive effort. We assume that the variable costs are captured by the function k_i which is increasing in the first argument, $k_{i,1} > 0$, satisfies the usual Inada conditions and, moreover, is such that $k_{i,12} < 0$, so that the marginal effort costs decrease in ω_i. We refer to ω_i as a measure of spouse i's productive ability. There also is a fixed cost for each spouse i that we denote by φ_i. The indicator function $\mathbf{1}(y_i > 0)$ takes the value of 1 if $y_i > 0$ and equals zero otherwise. A couple is characterized by a pair $\omega = (\omega_1, \omega_2)$ of productive abilities and a pair of fixed costs $\varphi = (\varphi_1, \varphi_2)$. We will sometimes write for

[13] For now the indices 1 and 2 are labels with no further meaning. In empirical applications, we will assign the index 1 to the spouse with higher earnings in the status quo.

[14] If the status quo has *joint taxation* there is a function $\tilde{T}^0$ so that $T^0(y) = \tilde{T}^0(y_1 + y_2)$ for all y. Then, $T^0_{y_1}(y) = T^0_{y_2}(y)$, for all (y_1, y_2).

short $\theta = (\omega, \varphi)$ with $\theta_1 = (\omega_1, \varphi_1)$ and $\theta_2 = (\omega_2, \varphi_2)$. The joint distribution of θ_1 and θ_2, denoted by F_θ, is a primitive of the economy. We assume that it is well-behaved in that the marginal and conditional *cdf*s that can be derived from F_θ are differentiable.[15]

A Tax Reform. A tax reform replaces the status quo tax system T^0 by a new one T^1. We write $T^1(y) = T^0(y) + \tau\, h(y)$ and refer to the function h as the reform direction and to the scalar τ as the size of the reform. We will focus on marginal effects, i.e. on the effect of raising τ slightly, starting from the status quo, $\tau = 0$. We denote by $y^*(\theta, \tau, h) = (y_1^*(\theta, \tau, h), y_2^*(\theta, \tau, h))$ the utility-maximizing choices of a couple that is confronted with the post-reform budget constraint

$$C^1(y) = b_m^1 + y_1 + y_2 - T^0(y) - \tau\, h(y) \ ,$$

where b_m^1 is the post-reform intercept of the consumption function. We assume that tax revenue is rebated lump sum so that $b_m^1 = b_m^0 + R(\tau, h)$, where $R(\tau, h)$ is the reform-induced change in overall tax revenue,

$$R(\tau, h) = \mathbb{E}_\theta\left[T^1(y^*(\theta, \tau, h))\right] - \mathbb{E}_\theta\left[T^0(y^0(\theta))\right] \ ,$$

and the operator $\mathbb{E}_\theta$ indicates that expectations are take with respect to the distribution F_θ. Finally, $y^0(\theta) := y^*(\theta, 0, h)$ is a shorthand for earnings in the status quo. We write $R_\tau(0, h)$ for a reform's marginal effect on tax revenue. It is an endogenous object that depends inter alia on the behavioral responses to taxation.[16] The spouses' behavioral responses to the tax reform are captured by

$$y_{1,\tau}^0(\theta, h) := \frac{\partial}{\partial \tau}\, y_1^*(\theta, \tau, h)_{|\tau=0} \quad \text{and} \quad y_{2,\tau}^0(\theta, h) := \frac{\partial}{\partial \tau}\, y_2^*(\theta, \tau, h)_{|\tau=0} \ ,$$

whenever these derivatives exist.[17]

Individual and Social Welfare. Using the indirect utility function

$$v(\theta, \tau, h) := \max_{y_1, y_2} \quad y_1 + y_2 - T^0(y) - \tau\, h(y) - K(y, \theta) \ ,$$

[15] We discuss alternative modeling choices regarding the behavior of couples and explain the extent to which our analysis is robust to such choices in Appendix B.1.10.

[16] For various reform directions of interest, including the set of all continuous functions h, $R_\tau(0, h)$ will be fully characterized in the course of our analysis.

[17] In the presence of fixed costs, the functions y_1^* and y_2^* jump at "cutoff levels" of φ_1 and φ_2 with the implication that left-sided and right-sided derivatives do not coincide. Discontinuities in marginal tax rates may also imply that left-sided and the right-sided derivative do not coincide. We explicitly deal with these complications below.

we can express the reform induced change in indirect utility for a couple with characteristics θ as $V(\theta,\tau,h) = R(\tau,h) + v(\theta,\tau,h) - v(\theta,0,h)$. By the envelope theorem, see Milgrom and Segal (2002), the marginal effect on indirect utility, associated with direction h, is given by $V_\tau(\theta,0,h) = R_\tau(0,h) - h(y^0(\theta))$. A reform in direction h is said to be *Pareto-improving* if, for all θ, $R_\tau(0,h) - h(y^0(\theta)) \geq 0$, with a strict inequality for some θ. Given a specification of welfare weights $\tilde{g} : \theta \mapsto \tilde{g}(\theta)$ with $\mathbb{E}_\theta[\tilde{g}(\theta)] = 1$, a direction h is said to be *welfare-improving* if

$$W_\tau(0,h) \quad := \quad R_\tau(0,h) - \mathbb{E}_\theta\left[\tilde{g}(\theta)h(y^0(\theta))\right] \quad \geq \quad 0\,.$$

2.2.1 Conditional Revenue and Welfare Functions

Theorems 1 and 2 below state necessary and sufficient conditions for the existence of a Pareto-improving reform direction. Theorem 3 characterizes the revenue and welfare implications of continuous reform directions. These theorems invoke what we refer to as *conditional revenue functions*. Here, as a preliminary step, we define these objects. Conditional revenue functions are derived from a specific class of simple tax reforms. They raise the marginal tax rates on the earnings of spouse i conditional on y_i and y_{-i} belonging to particular brackets of income. For instance, a reform that raises the marginal tax rate on the earnings of spouse 2, conditional on the earnings of spouse 1 belonging to a bracket $B_1(y_{1s},\ell_1) = [y_{1s}, y_{1s}+\ell_1]$ and the earnings of spouse 2 belonging to a bracket $B_2(y_{2s},\ell_2) = [y_{2s}, y_{2s}+\ell_2]$ can be represented by a scalar τ and a function $y \mapsto h(y)$ such that

$$h(y) \;=\; \begin{cases} y_2 - y_{2s}, & \text{if} \quad y \in B_1(y_{1s},\ell_1)\times B_2(y_{2s},\ell_2)\,, \\ \ell_2, & \text{if} \quad y_1 \in B_1(y_{1s},\ell_1) \quad \text{and} \quad y_2 \geq y_{2s}+\ell_2\,, \\ 0, & \text{otherwise}\,. \end{cases}$$

Consequently, the marginal tax rates on the earnings of spouse 2 change when the earnings of the spouses belong to the relevant brackets and do not change otherwise,

$$T^1_{y_2}(y) - T^0_{y_2}(y) = \begin{cases} \tau, & \text{if} \quad y \in B_1(y_{1s},\ell_1)\times B_2(y_{2s},\ell_2)\,, \\ 0, & \text{otherwise}\,. \end{cases}$$

Simple reforms so that B_1 and B_2 are brackets with positive length affect the earnings incentives of dual-earner couples. We use different notation for simple reforms that affect the earnings incentives of single-earner couples. For instance, a pair (τ, h^{s1}), with

$$h^{s1}(y) \;=\; \begin{cases} y_1 - y_{1s}, & \text{if} \quad y_1 \in B_1(y_{1s},\ell_1) \quad \text{and} \quad y_2 = 0\,, \\ \ell_1, & \text{if} \quad y_1 \geq y_{1s}+\ell_1 \quad \text{and} \quad y_2 = 0\,, \\ 0, & \text{otherwise}\,, \end{cases}$$

raises marginal tax rates on the earnings of spouse 1 conditional on $y \in B_1(y_{1s}, \ell_1)$ and spouse 2 having no earnings, $y_2 = 0$.

For $\tau >> 0$, a simple tax reform creates dominated regions in the space of primary and secondary earnings, i.e. regions where no couple chooses its earnings. Consider Figure 2.1, a simple tax reform generates a cliff between the purple rectangle with $h(y) = \ell_2$ and the neighboring rectangles with $h(y) = 0$. There also is a discontinuity in marginal tax rates when transiting from the blue rectangle with $h(y) = 0$ to the red rectangle with $h(y) = y_s - y_{2s}$. Typically, there will be bunching at such kink points. By Proposition 6 in part B.1.5 of the Appendix, this creates no complications for the analysis of such a reform's implications for tax revenue when τ is small. In particular, the impact of behavioral responses on tax revenue is then entirely due to those couples who, in the status quo, choose their earnings y in the region with $h(y) \neq 0$.

Conditional Revenue Functions for Dual-Earner Couples. When we seek to emphasize the parameters that define a simple tax reform we write $R_i(y_{1s}, \ell_1, y_{2s}, \ell_2)$ for its marginal impact on tax revenue, rather than using the more concise notation $R_\tau(0, h)$. The subscript $i \in \{1, 2\}$ indicates whether the marginal tax rates for spouse 1 or spouse 2 are affected by the reform. Of interest is also the derivative of $R_i(y_{1s}, \ell_1, y_{2s}, \ell_2)$ with respect to ℓ_i, evaluated at $\ell_i = 0$, see the graph on the left of Figure 2.4 for an illustration. We define the shorthand

$$\mathcal{R}_i(y_{is} \mid B_{-i}(y_{-is}, \ell_{-i})) \quad := \quad \frac{\partial}{\partial \ell_i} R_i(y_{1s}, \ell_1, y_{2s}, \ell_2)_{|\ell_i = 0} \, .$$

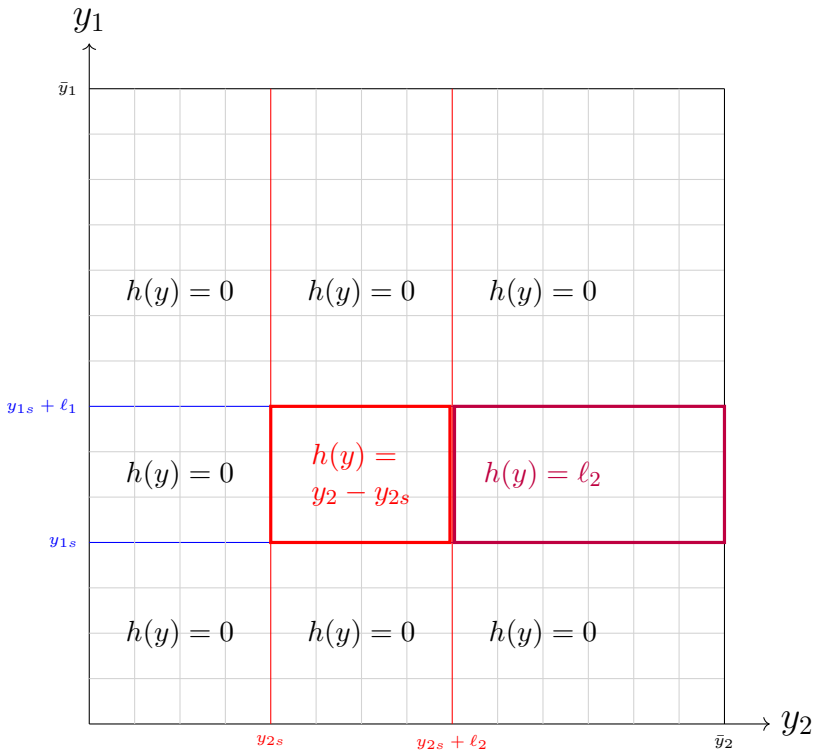
Finally, we define

$$\mathcal{R}_i(y_{is} \mid y_{-is}) \quad := \quad \frac{\partial}{\partial \ell_{-i}} \mathcal{R}_i(y_{is} \mid B_{-i}(y_{-is}, \ell_{-i}))_{|\ell_{-i} = 0} \, .$$

Henceforth, we refer to $y_i \mapsto \mathcal{R}_i(y_i \mid B_{-i})$ and to $y_i \mapsto \mathcal{R}_i(y_{is} \mid y_{-is})$ as *conditional revenue functions*; $\mathcal{R}_i(y_{is} \mid B_{-i})$ gives the revenue effect from a slight increase of the marginal tax rates that applies to a spouse i with earnings in a narrow bracket starting at the income level y_{is}, conditional on spouse $-i$ having earnings in a (possibly wide) bracket B_{-i}; $\mathcal{R}_i(y_{is} \mid y_{-is})$, by contrast, conditions on spouse $-i$ having earnings in a narrow bracket that starts at income level y_{-is}.[18]

[18] More formally, note that for τ close to zero, a Taylor approximation implies $R(\tau, h) \simeq \tau \, R_i(y_{1s}, \ell_1, y_{2s}, \ell_2)$. A repeated application of this argument implies that for τ, ℓ_i and ℓ_{-i} close to zero $R(\tau, h) \simeq \tau \, \ell_i \, \ell_{-i} \, \mathcal{R}_i(y_{is} \mid y_{-is})$.

Figure 2.1. – A Simple Tax Reform



Note: For couples with y in the red rectangle, the marginal tax rate that applies to y_2 is increased by τ. For couples with y in the purple rectangle, the tax burden increases by ℓ_2, while there is no change in marginal tax rates. For τ and ℓ_2 "small", $\tau\,\ell_2\mathcal{R}_2(y_{2s} \mid B_1[y_{1s}, \ell_1])$ is the marginal effect on tax revenue that results when, for couples with y in the red area, the marginal tax rate that applies to y_2 is increased by τ, with the consequence that, for couples with y in the purple area, the tax burden is increased by ℓ_2.

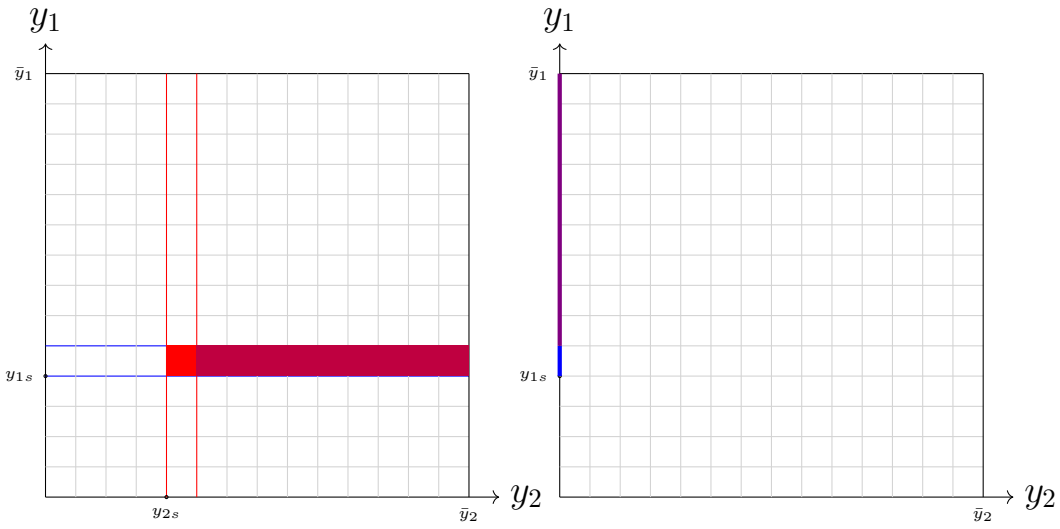
Conditional Revenue Functions for Single Earner Couples. For a simple tax reform that affects single-earner couples with, say, spouse 1 as the single earner, we write $R_1^s(y_{1s}, \ell_1)$ rather than $R_\tau(0, h)$. Of interest is also the derivative of $R_1^s(y_{1s}, \ell_1)$ with respect to ℓ_1, evaluated at $\ell_1 = 0$, , see the graph on the right of Figure 2.2 for an illustration. We define the shorthand

$$\mathcal{R}_1^s(y_{1s}) \quad := \quad \frac{\partial}{\partial \ell_1} R_1^s(y_{1s}, \ell_1)_{|\ell_1=0} \,.$$

Conditional Welfare Functions. Conditional welfare functions give the welfare rather than the revenue implications of simple reforms that affect marginal tax rates over narrow ranges of income. The conditional welfare functions for dual earner couples are formally defined by

$$\mathcal{W}_i(y_{is} \mid y_{-is}) \;:=\; \mathcal{R}_i(y_{is} \mid y_{-is}) - \frac{\partial}{\partial \ell_{-i}} \left(\frac{\partial}{\partial \ell_i} \mathbb{E}_\theta[\tilde{g}(\theta) h(\theta)]_{|\ell_i=0} \right)_{|\ell_{-i}=0} .$$

Figure 2.2. – Conditional Revenue Functions



(A) Dual-Earner Couples (B) Single-Earner Couples

Note: Panel (a): For τ, ℓ_2 and ℓ_1 "small", $\tau\,\ell_2\,\ell_1\,\mathcal{R}_2(y_{2s} \mid y_{1s})$ is the marginal effect on tax revenue that results when, for couples with y in the red area, the marginal tax rate that applies to y_2 is increased by τ, with the consequence that, for couples with y in the purple area, the tax burden is increased by ℓ_2. Panel (b): For τ, and ℓ_1 "small", $\tau\,\ell_1\,\mathcal{R}_1^s(y_{1s})$ is the marginal effect on tax revenue that results when, for couples with y on the blue line segment, the marginal tax rate that applies to y_1 is increased by τ, with the consequence that, for couples with y on the violet line segment, the tax burden is increased by ℓ_1.

One can show that

$$\mathcal{W}_i(y_{is} \mid y_{-is}) \;:=\; \mathcal{R}_i(y_{is} \mid y_{-is}) - (1 - F_{y_i}(y_{is} \mid y_{-is}))\mathcal{G}(y_{is} \mid y_{-is}) f_{y_{-i}}(y_{-is})\;.$$

where $\mathcal{G}(y_{is} \mid y_{-is}) \;:=\; \mathbb{E}_y(g(s_1, s_2) \mid s_i \geq y_{is} \text{ and } s_{-i} = y_{-is})$ is the average welfare weight of couples where spouse $-i$ has earnings of y_{-is} and spouse i has earnings exceeding y_{is}, and $g(y_1, y_2) := \mathbb{E}_\theta[\tilde{g}(\theta) \mid (y_1^0(\theta), y_2^0(\theta)) = (y_1, y_2)]$ is the average welfare weight of couples with incomes of y_1 and y_2 in the status quo. The obvious adjustments for the case of single earner couples yield conditional welfare functions of the form $\mathcal{W}_i^s(y_{is}) \;:=\; \mathcal{R}_1^s(y_{1s}) - (1 - F_{y_i}(y_{is} \mid 0))\mathcal{G}(y_{is} \mid 0) F_{y_j}(0)\;.$

2.2.2 Theorems

Pareto-Efficiency. Let $\mathcal{H}$ by the set of functions $h : y \mapsto h(y)$ that take at least two different values on $y^0(\Theta)$. A tax system is Pareto-efficient only if, for any reform direction $h \in \mathcal{H}$,

$$R_\tau(0, h) \;-\; \max_\theta\; h(y^0(\theta)) \quad \leq \quad 0\,. \tag{2.1}$$

If this condition was violated for some $h \in \mathcal{H}$, there would by a reform (τ, h) that makes some couples strictly better off without making others worse off. Theorem 1 states necessary conditions for the Pareto-efficiency of a tax system. Its proof exploits that, under a Pareto-efficient tax system, condition (2.1) has to hold for all h in the previously introduced class of simple tax reforms.

Proposition 1. *If a tax system is Pareto-efficient then: For all i and j and for all B_j, the conditional revenue functions $y_i \mapsto \mathcal{R}_i(y_i \mid B_j)$ and $y_i \mapsto \mathcal{R}_i^s(y_i)$ are (i) non-increasing, (ii) bounded from below by 0, and (iii) bounded from above by 1.*

Theorem 1 extends a result from Bierbrauer et al. (2023a) to a tax function with two arguments. Pareto-efficiency implies that conditional revenue functions have to be non-increasing. If this condition is violated one can Pareto-improve the tax system by means of a self-financing tax cut that applies to all incomes that lie in the income range over which the conditional revenue function is increasing. Pareto-efficiency also implies that conditional revenue functions have to be bounded from below by zero. This is the analogue to the famous Laffer-condition. If there is y so that $\mathcal{R}_i(y_i \mid y_j) < 0$, then lowering the marginal tax rate for spouse i by means of a simple tax reform is self-financing. Finally, there is the "reverse Laffer condition": If $\mathcal{R}_i(y_i \mid y_j)$ gets too large, then increasing the marginal tax rate for spouse i by means of a simple tax reform generates so much additional revenue that even those who now face higher taxes are better off. This is incompatible with Pareto-efficiency in the status quo. Upon letting the length of the conditioning income range vanish, we obtain a Corollary to Theorem 1 that is without precedence in earlier work. As will become clear shortly, the Corollary enables us to transition from the necessary conditions for Pareto-efficiency in Theorem 1 to the sufficient conditions in Theorem 2.

Corollary 1. *If a tax system is Pareto-efficient, then, for all i and j and for all $y \in \mathcal{Y}$, the functions $y_i \mapsto \mathcal{R}_i(y_i \mid y_j)$ are (i) non-increasing, (ii) bounded form below by 0, and (iii) bounded from above in the following sense: There is $\bar{\ell}$, so that $\ell_j < \bar{\ell}$ implies*

$$\ell_j \, \mathcal{R}_i(y_i \mid y_j) \quad \leq \quad 1 \, . \tag{2.2}$$

Proposition 2. *Under the conditions listed in Corollary 1, if h is continuous, then*

$$R_\tau(0, h) \; - \; \max_{y \in \mathcal{Y}} \; h(y) \quad \leq \quad 0 \, . \tag{2.3}$$

Theorem 2 shows that the conditions in Corollary 1 are sufficient for the local Pareto-efficiency of a tax system: If these conditions are met there is no Pareto-improving direction in the set of continuous functions. Together Corollary 1 and Theorem 2 imply that all conceivable violations of Pareto-efficiency can be found by checking whether the conditional revenue functions have the properties listed in Corollary 1.

The proof of the Theorem proceeds as follows: We take an arbitrary continuous function h as given. We show that any such function can be approximated arbitrarily well by a sum of simple tax reforms. Specifically, we consider functions h_m that are piecewise linear over a partition of $\mathcal{Y}_0$ into m^2 squares and which converge to h as $m \to \infty$. We show that, under the conditions in Corollary 1, there is no function h_m that yields a Pareto-improvement and then extend the argument to the function h. This requires to verify that, for any m, (i) there is no Pareto-improving direction in the set of reforms affecting marginal tax rates only for primary earners, (ii) that there neither is a Pareto-improving direction in the set of reforms affecting marginal tax rates only for secondary earners, and (iii) finally, that there is no Pareto-improving direction that combines changes of marginal tax rates for primary earners and changes of marginal tax rates for secondary earners. The proofs of Theorem 2 and also of Theorem 3 below make intense use of the linearity of Gateaux differentials which implies that the revenue implications of h_m can be written as a linear combination of conditional revenue functions.[19]

Social Welfare. The following theorem provides a characterization of an arbitrary reform direction h's marginal impact on tax revenue and social welfare. Its significance lies in two observations. First, this impact is linear in the changes of marginal tax rates – denoted by h_1 for the earnings of spouse 1 and by h_2 for the earnings of spouse 2 – that are implied by h. Second, the impact associated with a change of marginal tax rates is weighted by the relevant conditional revenue and welfare functions. Thus, the revenue and welfare implications of continuous reforms directions can be written as linear combinations of the revenue and welfare implications of simple reforms.

Proposition 3. *For any continuous reform direction $y \mapsto h(y)$, the marginal effects on tax revenue and social welfare are respectively given by*

$$\begin{aligned} R_\tau(0,h) \;\; = \;\; & \int_{\mathcal{Y}} h_1(y)\, \mathcal{R}_1(y_1 \mid y_2)\, dy \;+\; \int_{\mathcal{Y}_1} h_1(y_1,0)\, \mathcal{R}_1^s(y_1)\, dy_1 \\ & + \int_{\mathcal{Y}} h_2(y)\, \mathcal{R}_2(y_2 \mid y_1)\, dy \;+\; \int_{\mathcal{Y}_2} h_2(0,y_2)\, \mathcal{R}_2^s(y_2)\, dy_2 \;, \end{aligned} \tag{2.4}$$

[19] Assumptions which guarantee the linearity of the Gateaux differential in the direction h are discussed in Bierbrauer et al. (2023a).

and

$$\begin{aligned} W_\tau(0,h) &= \int_{\mathcal{Y}} h_1(y)\,\mathcal{W}_1(y_1 \mid y_2)\,dy + \int_{\mathcal{Y}_1} h_1(y_1,0)\,\mathcal{W}_1^s(y_1)\,dy_1 \\ &\quad + \int_{\mathcal{Y}} h_2(y)\,\mathcal{W}_2(y_2 \mid y_1)\,dy + \int_{\mathcal{Y}_2} h_2(0,y_2)\,\mathcal{W}_2^s(y_2)\,dy_2\,. \end{aligned} \tag{2.5}$$

The observation that welfare implications are linear in the changes of marginal tax rates has an immediate consequence for welfare-maxima. A welfare maximum has the property that $W_\tau(0,h) = 0$, holds for all h. Given the characterization of $W_\tau(0,h)$ in Theorem 3, this is possible if and only if all conditional welfare functions are equal to zero, everywhere.

Corollary 2. *The following statements are equivalent:*

1. *For all continuous functions $h : \mathcal{Y}_0 \to \mathbb{R}$, $W_\tau(0,h) = 0$.*

2. *For all $y \in \mathcal{Y}$, $\mathcal{W}_1(y_1 \mid y_2) = \mathcal{W}_2(y_2 \mid y_1) = \mathcal{W}_1^s(y_1) = \mathcal{W}_2^s(y_2) = 0$.*

Below we will draw on Corollary 2 for a characterization of optimal top tax rates both on primary and secondary earnings. The following Corollary illustrates the possibility to use Theorem 3 also for an assessment of reforms starting from a possibly suboptimal status quo. Specifically, it shows that the lowering of marginal tax rates on secondary earnings reduces Rawlsian measures of social welfare when the starting point is an efficient tax system.

Corollary 3. *Consider a status quo so that y is in the support of F_y only if $y_2 \leq y_1$. Consider a subset of incomes $\mathcal{Y}_0'$ and tax reform so that $h_2(y) < 0$ if $y \in \mathcal{Y}_0' \subset \mathcal{Y}_0$ and $h_2(y) = 0$, otherwise. Let $h_1(y) = 0$, for all $y \in \mathcal{Y}_0$. Suppose that $g(y) \leq \bar{g}^d$, whenever $y_1 > 0$ or $y_2 > 0$. Suppose that there is no Pareto-improving direction for reform. If $\bar{g}^d$ is sufficiently small, then $W_\tau(0,h) < 0$.*

The Corollary is based on the premise that $y_2 \leq y_1$. Thus, the spouse with index 2 is the secondary earner. It then considers a reform so that the marginal tax rates in some set of secondary earnings are reduced while all other marginal tax rates remain unchanged. It shows that this reform is welfare-damaging when two conditions are fulfilled: First, the status quo is an efficient tax system in the sense that the conditions in Theorem 2 are satisfied. Second, the welfare weights of couples with positive earnings are sufficiently low. The Corollary follows from Theorem 3 upon noting that $\mathcal{R}_2(y_2 \mid y_1) > 0$, for all y,

when the status quo is an efficient tax system. Thus, for $\bar{g}^d$ close to zero, $\mathcal{W}_2(y_2 \mid y_1) > 0$. Hence, all else equal, a lowering of marginal tax rates for (a subset of) secondary earners, $h_2(y) < 0$, yields a loss of social welfare.

2.2.3 Using the Theorems to Analyze Changes of Jointness

The term jointness refers to the interdependence of the spouses marginal tax rates. Under a system of progressive and joint taxation, there is positive jointness: the marginal tax rate on the earnings of spouse i is a non-decreasing function of the earnings of spouse $-i$. Under individual taxation, by contrast, there is zero jointness. In the following, we discuss reforms that reduce the tax system's degree of jointness. If the status quo has progressive and joint taxation, then a jointness-reducing reform can be interpreted as bringing the status quo closer to individual taxation.

A reform in direction h alters the tax system's jointness when the cross-derivative h_{12} differs from zero over parts of the income distribution,

$$T^1_{y_1,y_2}(y) \quad = \quad T^0_{y_1,y_2}(y) \; + \; \tau \, h_{12}(y) \; .$$

The following Corollary to Theorem 3 gives the welfare implications associated with a reform in direction h via the implied changes of *jointness*.

Corollary 4. *For any continuous reform direction $y \mapsto h(y)$,*

$$\begin{aligned} W_\tau(0,h) \;\; = \;\; & \textstyle\int_{\mathcal{Y}} h_{12}(y)\, \mathbf{W}(y) dy \\ & + \textstyle\int_{\mathcal{Y}_1} h_1(y_1,0) \left(\mathcal{W}^s_1(y_1) + \int_{\mathcal{Y}_2} \mathcal{W}_1(y_1 \mid y_2) dy_2 \right) dy_1 \\ & + \textstyle\int_{\mathcal{Y}_2} h_2(0,y_2) \left(\mathcal{W}^s_2(y_2) + \int_{\mathcal{Y}_1} \mathcal{W}_2(y_2 \mid y_1) dy_1 \right) dy_2 \; , \end{aligned}$$

where

$$\mathbf{W}(y) \quad := \quad \int_{y_1}^{\bar{y}_1} \mathcal{W}_2(y_2 \mid y_1) dy_1 \; + \; \int_{y_2}^{\bar{y}_2} \mathcal{W}_1(y_1 \mid y_2) dy_2 \; .$$

An Example. Consider two brackets $B_1(y_{1s}, \ell_1)$ and $B_2(y_{2s}, \ell_2)$ with $y_{1s}, y_{2s} > 0$ and a reform in direction h so that $h(y_1, y_2) \; = \; h^1(y_1)\, h^2(y_2)$ where, for $i = 1, 2$,

$$h^i(y_i) \;\; = \;\; \begin{cases} 0 & \text{if} \;\; y_i \leq y_{is} \; , \\ y_i - y_{is}, & \text{if} \;\; y_i \in B_i(y_{is}, \ell_i) \; , \\ \ell_i, & \text{if} \;\; y_i \geq y_{is} + \ell_i \; . \end{cases}$$

Note that h is the product of two continuous functions and hence also continuous. Also, note that for $y_1 \in B_1(y_{1s}, \ell_1)$ and $y_2 \in B_2(y_{2s}, \ell_2)$, $h_{12}(y) = 1$. Finally, note that the reform does not involve changes of marginal tax rates for single-earner couples. Therefore, by Corollary 4, the welfare implications of this reform are given by

$$W_\tau(0, h) = \int\limits_{B_1 \times B_2} \mathbf{W}(y, \lambda)\, dy\ . \tag{2.6}$$

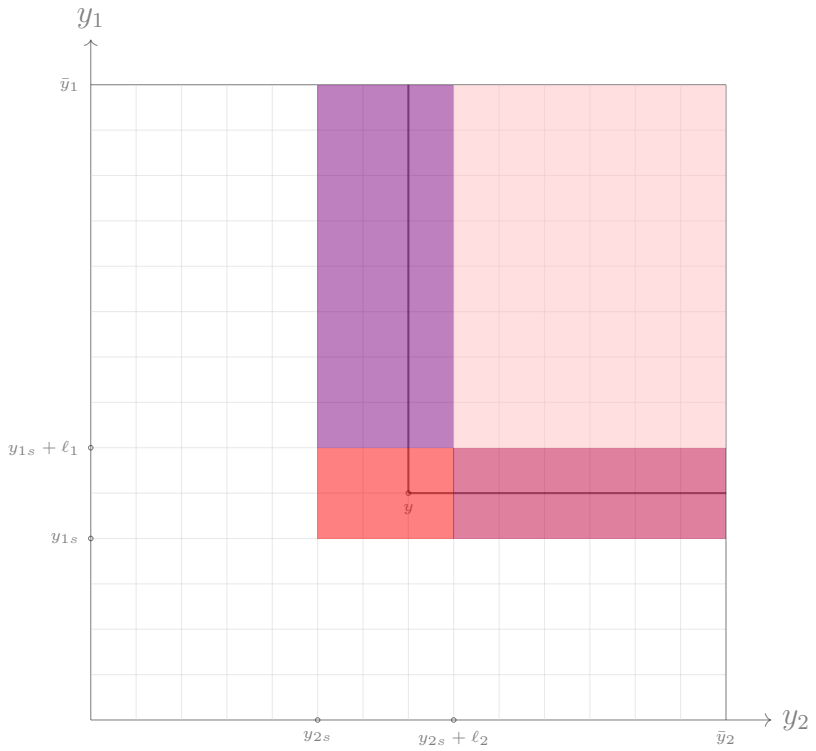
Equation (2.6) has implications that are noteworthy: First, suppose the status quo is efficient so that all revenue functions are strictly positive. Also suppose that the welfare measure is Rawlsian so that, for all i, and y

$$\mathcal{W}_i(y_i \mid y_j) \quad = \quad \mathcal{R}(y_i \mid y_j) \quad > \quad 0\ .$$

Then, $\mathbf{W}(y, \lambda) > 0$ for all y and hence $W_\tau(0, h) > 0$. A reduction of jointness ($\tau < 0$) then yields a loss of tax revenue and of Rawlsian welfare. This observation echoes Corollary 2 above. If the status quo is efficient, a move towards individual taxation via a reduction of jointness is rejected by a Rawlsian welfare measure. Second, suppose that a reduction of jointness is instead Pareto-improving. Then, it has to be the case that $W_\tau(0, h) < 0$ for all possible welfare measures, including the Rawlsian measure. Consequently, given a Rawlsian welfare measure, there must be y, so that $\mathbf{W}(y, \lambda) < 0$. But this is possible only if there are i and j so that $\mathcal{R}(y_i \mid y_j) < 0$; i.e. if one of the sufficient conditions in Theorem 2 is violated. More specifically, if a jointness-reducing reform is Pareto-improving, then there is also a simple reform – lowering of the marginal tax rates on the earnings of one spouse only – that is Pareto-improving. Conversely, if the scope for Pareto-improvements via simple reforms has been exhausted, then there is also no possibility to Pareto-improve the tax system via a modification of jointness. To see where all this is coming from, consider Figure 2.3. The figure illustrates that a reduction of jointness implies a reduction of marginal tax rates in different parts of the income space. The question whether a reduction of jointness is desirable therefore amounts to asking whether these reductions of marginal tax rates are desirable.

Figure 2.3 is also helpful to clarify the relation between Theorem 3 and Corollary 4. Theorem 3 traces welfare implications back to changes in marginal tax rates. Corollary 4 traces them back to changes in jointness. To describe the connection, we take a closer look at $\mathbf{W}(y)$. Referring to Figure 2.3, what this expression does is the following: It takes an arbitrary point in the red rectangle and then sums the welfare implications of changes of marginal tax rates in vertical direction (for the spouses with index 1) and the welfare

Figure 2.3. – A Change in Jointness



Note: The figure illustrates the implications that the reform in the example above has for jointness and marginal tax rates: Jointness changes only in the red rectangle. There, $h_{12}(y) = 1$, and $h_{12}(y) = 0$ elsewhere. Moreover, in the red area the marginal tax rate on the earnings of spouse 1 and also the marginal tax rate on the earnings of spouse 2 change since $h_1(y) = y_2 - y_{2s}$ and $h_2(y) = y_1 - y_{1s}$. In the violet area, the marginal tax rate on the earnings of spouse 1 does not change, $h_1(y) = 0$, but the marginal tax rates on the earnings of spouse 2 changes since $h_2(y) = \ell_1$. Analogously, in the purple area, there is no change in the marginal tax rate for spouse 2, but a change in the marginal rate of spouse 1, $h_1(y) = \ell_2$. In the pink area, there is no change in marginal tax rates, but the total tax burden changes since $h(y) = \ell_1\ell_2$.

implications of changes of marginal tax rates in horizontal direction (for the spouses with index 2.) The consequence for the overall welfare measure is obtained by "summing over" all the points in the red rectangle.

Corollary 4 gives an intuitive understanding what a change in jointness means: Reducing the jointness for couples where spouse 1 has close to 80,000 USD per year and spouse 2 has close to $60,000$ USD, amounts to a transition into (i) lower marginal tax rates for the earnings of spouse 1 when spouse 1 has more than 80,000 USD, holding spouse 2 fixed at 60,000 USD, and (ii) lower marginal tax rates for the earnings of spouse 2 when spouse 2 has more than 60,000 USD, conditional on spouse 1 making 80,000 USD.

Beyond the Example. Corollary 4 implies that Equation (2.6) gives the welfare implications for any reform so that (i) $h_{12}(y) = 1$, for $y \in B_1 \times B_2$, and $h_{12}(y) = 0$ otherwise,

and (ii) there are no changes of marginal tax rates for single-earner couples. Thus, there is a whole class of different h functions which all have the same welfare implication and the one in the example above is just one function in this class.

In the following, we develop a sufficient statistics characterization of the conditional revenue and welfare functions. This will ultimately enable us to identify the scope for Pareto- or welfare-improving empirically. An intermediate step on this path is relegated to Appendix B.1.5 where we clarify what behavioral responses look like in our framework and, moreover, what they imply for tax revenue.

2.2.4 Conditional Revenue Functions as Sufficient Statistics

Conditional revenue functions give the revenue implications of simple tax reforms. Consider an increase of the marginal tax rates on the earnings of spouse 2 conditional on $y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2)$ and focus on the limit case $\ell_1 \to 0$ and $\ell_2 \to 0$. The revenue effect of this reform is denoted by $\mathcal{R}_2(y_{2s} \mid y_{1s})$. Proposition 1 below provides a characterization.[20] An analogous characterization of the revenue functions for single-earner couples can be found in Appendix B.1.7. We first introduce notation that will enable us to state the Proposition in a concise way.

The Extensive Margin. The mass of couples who experience an increase of their tax burden due to an infra-marginal increase of tax rates is given by $1 - F_{y_2}(y_{2s} \mid y_{1s})$. To get the reform's impact on the tax revenue that is collected form this set of individuals, this mass is multiplied by $1 - \mathcal{E}_x^d(y_{2s} \mid y_{1s})$, where $\mathcal{E}_x^d(y_{2s} \mid y_{1s})$ is an extensive margin elasticity, reflecting that the reform turns some of these dual-earner couples into single-earner couples.[21]

The Intensive Margin. Conditional on being a dual-earner couple, a couple's optimal earning choices depend only on ω. Intensive margin responses arise when these choices are affected by a tax reform. Henceforth, we denote by $\sigma^{d0}(\omega)$ the mass of dual earner couples among those with productive abilities ω in the status quo.

[20] A characterization of $\mathcal{R}_1(y_{1s} \mid y_{2s})$ can be obtained along the same lines.

[21] Formally, $\mathcal{E}_x^d(y_{2s} \mid y_{1s})$ is the percentage change in the mass of dual earner couples that is due to a percentage change of the tax advantage of dual-earner over single-earner couples.

In Proposition 1 below, the implications of behavioral responses at the intensive margin for tax revenue are captured by

$$\bar{\beta}^{d}_{I,2}(y_{1s}, y_{2s}) \;=\; \mathbb{E}_{\omega_1}\Big[\beta^{d0}_{I,2}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1^{d0*}(\omega) \;=\; y_{1s}\Big] \,,$$

where the expectation is over dual-earner couples with productive abilities ω who, in the status quo, have an income in the region where marginal tax rates on y_2 go up. For such a couple,

$$\beta^{d0}_{I,2}(\omega) \;=\; \sigma^{d0}(\omega)\Big(T^0_{y_1}(y^{d0}(\omega)) y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega)) y^{d0}_{2,\tau}(\cdot)\Big)$$

captures the revenue implications of behavioral responses at the intensive margin.[22] The expectation operator conditions on y_1 being equal to y_{1s} and then goes over the "relevant" combinations of ω_1 and ω_2. The way in which expectations are computed captures, in particular, the possibility that the status quo schedule has a kink at y_{2s}. To see this, for any y_2', and any ω_1, let

$$F_{y_2}(y_2' \mid \omega_1) \quad := \quad F_{\omega_2}(\bar{\omega}_2(y_2' \mid \omega_1) \mid \omega_1) \,,$$

where, given ω_1, $\bar{\omega}_2(y_2' \mid \omega_1)$ is the largest value of ω_2 so that $y_2^{d0*}(\omega_1, \omega_2) \leq y_2'$. Thus, $F_{y_2}(y_2' \mid \omega_1)$ is a mass of couples who choose $y_2 \leq y_{2s}$. Upon differentiating with respect to y_2', we obtain

$$f_{y_2}(y_2' \mid \omega_1) \quad := \quad f_{\omega_2}(\bar{\omega}_2(y_2' \mid \omega_1) \mid \omega_1) \frac{\partial}{\partial y_2'} \bar{\omega}_2(y_2' \mid \omega_1) \,.$$

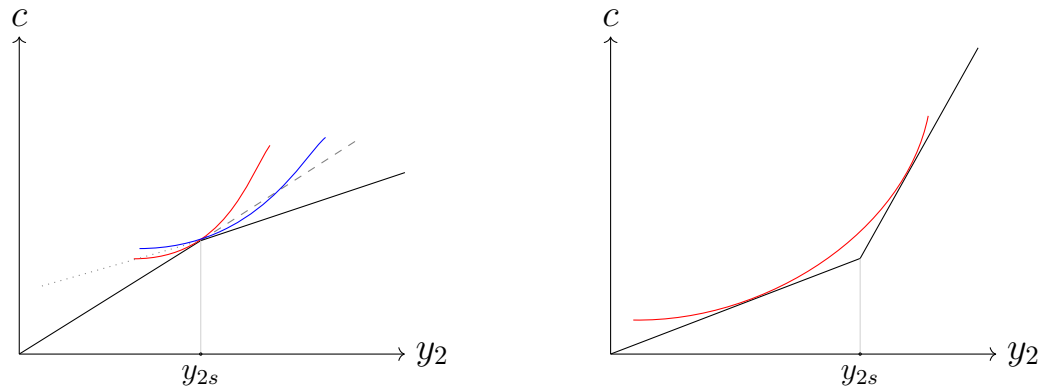
If couples with $y_2 = y_{2s}$ satisfy a first order condition, then $f_{y_2}(y_{2s} \mid \omega_1)$ is a measure of the mass of couples who choose y_2 very close to y_{s2}. If, by contrast, there is a kink because marginal tax rates jump upwards, see the left part of Figure 2.2, then there is a mass point at y_{2s} and $f_{y_2}(y_{2s} \mid \omega_1)$ indicates how this mass changes when the marginal tax rate changes at y_{2s}. Finally, if there is a kink because marginal tax rates jump downwards, see the right part of Figure 2.2, then $f_{y_2}(y_{2s} \mid \omega_1)$ indicates how the mass of people who choose y_2 strictly smaller than y_{2s} changes.

Proposition 1. *Consider an increase of the marginal tax rates on the earnings of spouse 2 conditional on $y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2)$. For the limit case $\ell_1 \to 0$ and $\ell_2 \to 0$, the revenue effect of this reform is given by*

$$\mathcal{R}(y_{2s} \mid y_{1s}) \quad = \quad f_{y_1}(y_{1s}) \left(\bar{\beta}^{d}_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \left(1 - \mathcal{E}^{d}_{x}(y_{2s} \mid y_{1s})\right)\right) \,,$$

[22] See Appendix B.1.8 for a detailed characterization of $\beta^{d0}_{I,2}(\omega)$.

Figure 2.4. – Intensive Margin Responses with Discontinuities in Marginal Tax Rates



(A) Bunching (B) Jumping

Note: The black line is the frontier of the couple's budget set conditional on $y_1 = y_{1s}$. Panel (a): There is a kink at y_{2s} because marginal tax rates jump upwards. Holding fixed the productive ability of spouse 1, ω_1, and the fixed costs φ_1 and φ_2, there is a mass of couples who choose $y_2 = y_{2s}$. These couples differ in the value of ω_2. The red line is the indifference curve of the couple with the minimal value of ω_2. The blue line is the indifference curve of the couple with the maximal value of ω_2. When the marginal tax rate changes at y_{2s} then this maximal value changes and hence the mass of couples who bunch at the kink. Panel (b): There is a kink at y_{2s} because marginal tax rates jump downwards. Holding fixed the productive ability of spouse 1, ω_1, and the fixed costs φ_1 and φ_2, there is a mass of couples who choose $y_2 < y_{2s}$ and a mass of couples who choose $y_2 > y_{2s}$. These couples differ in the value of ω_2. The red line is the indifference curve of the couple with the critical value of ω_2 that separates these two sets. Note that the earnings function $y_2^*(\cdot)$ jumps upwards at this critical level When the marginal tax rate changes at y_{2s} then this critical value changes and hence the mass of couples who choose $y_2 < y_{2s}$.

where $y_1 \mapsto f^{y_1}(y_1)$ *is the density associated with the marginal distribution of* y_1 *and* $y_2 \mapsto F^{y_2}(y_{2s} \mid y_{1s})$ *is the cdf of* y_2*, conditional on* y_1 *being equal to* y_{1s}*.*

The revenue function simplifies further under an additional assumption: When the tax system is piecewise linear, effort-cost functions are iso-elastic

$$k_2(y_2, \omega_2) \quad = \quad \frac{1}{1+\frac{1}{\varepsilon_2}} \left(\frac{y_2}{\omega_2}\right)^{1+\frac{1}{\varepsilon_2}} \ ,$$

and earnings choices are characterized by first order conditions, then[23]

$$\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) \quad = \quad -f_{y_2}(y_{2s} \mid y_{1s}) \, \frac{T^0_{y_2}(y_{1s}, y_{2s})}{1 - T^0_{y_2}(y_{1s}, y_{2s})} \, y_{2s} \, \varepsilon_2 \ . \tag{2.7}$$

[23]The derivation can be found in part B.1.8 of the Appendix.

For later reference, we note that Proposition 1 can be extended so as to obtain a sufficient statistics characterization also of the corresponding conditional welfare functions. This yields

$$\begin{aligned}\mathcal{W}_2(y_{2s} \mid y_{1s}) &= f_{y_1}(y_{1s})\Big(\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \times \\ &\qquad \Big(1 - \mathcal{G}(y_{2s} \mid y_{1s}) - \mathcal{E}^d_x(y_{2s} \mid y_{1s})\Big)\Big) \,.\end{aligned} \tag{2.8}$$

2.2.5 Tax Rates at the Bottom and Optimal Tax Rates at the Top

In the following, we will use the sufficient statistics characterization of conditional revenue and welfare functions to discuss two specific policy design questions: (a) The desirability of workfare elements in the tax and transfer system, and (b) optimal tax rates at the top of the income distribution.

Marginal Tax Rates at the Bottom. The following assumption is empirically plausible at the bottom of the income distribution.

Assumption 1. *Suppose that there is a range of incomes $Y' = Y'_1 \times Y'_2$ over which the following properties all hold true: (i) $\mathcal{G}(y_2 \mid y_1) \geq 1$ and $0 \leq \mathcal{E}^d_x(y_{2s} \mid y_{1s}) < 1$, with at least one of the weak inequalities being strict. (ii) The choices of dual earner couples with $y^{d0}(\theta) \in Y'$ satisfy first order conditions. (iii) Effort cost functions are iso-elastic. (iv) The tax system in is piecewise linear. (v) Marginal tax rates are (weakly) positive.*

To see that part (i) is plausible for low incomes, suppose that $(y_1, y_2) \mapsto \mathcal{G}(y_2 \mid y_1)$ is continuous and non-increasing both in y_1 and y_2. Since $\mathbb{E}_y[g(y)] = 1$ this implies that $1 - \mathcal{G}(y_{2s} \mid y_{1s}) \leq 0$ for y_{1s} and y_{2s} both sufficiently close to zero. This inequality is strict if the function $(y_1, y_2) \mapsto \mathcal{G}(y_2 \mid y_1)$ is strictly decreasing. Part (i) allows both for extensive margin elasticities of zero and for positive ones. Empirically, it has been shown that extensive marginal elasticities are largest at low levels of income, while the assumption that extensive margin elasticities are bounded from above by 1 is also in line with the empirical literature. Part (ii) is a regularity condition that simplifies the analysis. Part (iii) is another simplification which, moreover, relates the welfare implications of tax

changes to an elasticity that is frequently estimated, the Frisch elasticity. Since any continuous tax function can be well approximated by piecewise linear tax system, part (iv) does not seem to be strong restriction.[24]

Proposition 2. *Under Assumption 1,*

$$\mathcal{W}(y_{2s} \mid y_{1s}) \quad < \quad 0 \quad < \quad \mathcal{R}(y_{2s} \mid y_{1s})\,.$$

We omit a formal proof of Proposition 2. It follows immediately from squaring Proposition 1 with Equations (2.7) and (2.8). The Proposition gives conditions under which it is desirable to lower marginal tax rates on the earnings of spouse 2.[25] Moreover, when $\mathcal{W}(y_{2s} \mid y_{1s}) < 0$, and $0 < \mathcal{R}(y_{2s} \mid y_{1s})$, then the lowering of marginal tax rates has two effects: First, it reduces the intercept of the consumption schedule, and therefore is harmful to those with zero earnings both before and after the reform. A Rawlsian welfare function therefore rejects the reform. Second, it makes the secondary earner's budget line in a (c, y)-space steeper, so that the couple's disposable income rises more strongly in secondary earnings. This resembles introducing or strengthening elements of workfare in the tax and transfer system: Less unconditional transfers and more high-powered earnings incentives. It is remarkable that such reforms are desirable for a large class of social welfare functions, including ones with significant inequality aversion. Proposition 2 also implies that optimal marginal tax rates are negative at the bottom of the income distribution: If parts (i) - (v) of Assumption 1 hold, then the condition $\mathcal{W}(y_{2s} \mid y_{1s}) = 0$, yields negative marginal tax rates. Note in particular, that with $\mathcal{G}(y_{2s} \mid y_{1s}) > 1$, negative marginal tax rates are desirable even with an extensive margin elasticity of zero. This contrasts with results from Mirrleesian models with a single-decision maker. In those models, extensive margin responses are necessary for the possibility to rationalize negative marginal tax rates when welfare weights are monotonic in income.[26]

Top Tax Rates. We impose a set of assumptions that are analogous to the ones used by Diamond (1998) and Saez (2001) in their characterizations of top tax rates in the basic Mirrless model that does not have couples.

[24]The following paragraph, moreover, presents conditions under which the conditions for welfare optimality can be satisfied fully – as opposed to approximately – with a piecewise linear tax system.

[25]A symmetric argument applies to the earnings of spouse 1.

[26]The seminal reference is Saez (2002), see Bierbrauer et al. (2023a) for a more extensive discussion of this literature.

Assumption 2. *Suppose that there is a range of incomes $Y' = Y_1' \times Y_2'$ over which the following properties all hold true: (i) There are numbers $\bar{g}_1$ and $\bar{g}_2$ so that $\mathcal{G}(y_1 \mid y_2) = \bar{g}_1$ and $\mathcal{G}(y_2 \mid y_1) = \bar{g}_2$, for all $(y_1, y_2) \in Y'$. (ii) There are numbers $\bar{\alpha}_1$ and $\bar{\alpha}_2$ so that*

$$\frac{f_{y_1}(y_1 \mid y_2)\, y_1}{1 - F_{y_1}(y_1 \mid y_2)} \;=\; \bar{\alpha}_1 \quad \textit{and} \quad \frac{f_{y_2}(y_2 \mid y_1)\, y_2}{1 - F_{y_2}(y_2 \mid y_1)} \;=\; \bar{\alpha}_2 \;,$$

for all $(y_1, y_2) \in Y'$. (iii) The choices of dual earner couples with $y^{d0}(\theta) \in Y'$ satisfy first order conditions. Moreover, their extensive margin elasticities are zero. (iv) Effort cost functions are iso-elastic.

Again, these assumptions do not formally refer to the top of the income distribution. Empirically, however, they are most plausible near the top. For Y' near the top, part (i) says that welfare weights do not discriminate between "the rich" and "the super rich". They may discriminate, however, between rich primary and rich secondary earners which is the case when $\bar{g}_1 \neq \bar{g}_2$. Part (ii) says that, in the relevant income range, conditional income distributions are (well approximated by) Pareto-distributions, and moreover that the Pareto-parameter for the conditional income distribution of spouse i is, locally, independent of the earnings of spouse $-i$. Parts (iii) and (iv) are as in the previous section on marginal tax rates at the bottom of the income distribution.

Proposition 3. *Suppose that $g(y) > 0$ for all $y \in \mathcal{Y}$. Define t_i^* implicitly as the solution to*

$$\frac{t_i^*}{1 - t_i^*} \;=\; \Big(1 - \bar{g}_i\Big)\, \frac{1}{\bar{\alpha}_i}\, \frac{1}{\varepsilon_i}.$$

Under Assumption 2, $T_{y_1}^0(y) = t_1^*$ and $T_{y_2}^0(y) = t_2^*$, solves $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$, for all $(y_1, y_2) \in Y'$.

By Corollary 2, a welfare-maximization implies $\mathcal{W}_1(y_1 \mid y_2) = \mathcal{W}_2(y_2 \mid y_1) = 0$. Assumption 2 makes it possible to find an analytical solution to this system of equations. Near the top the solution takes the form of two optimal tax formulae, one applying to y_1, one applying to y_2, that are both akin to the famous ABC formula developed in Diamond (1998). In particular, under Assumption 2 the optimal tax function is separable and piecewise linear with zero jointness for $y \in Y'$.

2.3 EMPIRICAL ANALYSIS

We now build on the theoretical framework to empirically estimate conditional revenue functions. We use these revenue functions to illustrate inefficiencies in the tax treatment

of married couples (Section 2.3.2), and to describe the welfare consequences of reducing jointness (Section 2.3.3). Section 2.3.4 analyzes the welfare implications of lowering marginal tax rates at the bottom of the income distribution while we estimate optimal marginal top tax rates for primary and secondary earners in Section 2.3.5.

2.3.1 Conditional Revenue Functions

To estimate conditional revenue functions, we draw on detailed household micro data from the Annual Social and Economic Supplement of the CPS (CPS-ASEC).[27] The detailed demographic characteristics in the CPS allow us to obtain information on the (conditional) income distributions of primary and secondary earners in married couples. Information on tax rates is obtained through the NBER TAXSIM (v32) microsimulation model.[28] By combining CPS data with TAXSIM, we estimate tax rates and tax liabilities under the status quo at the household level. This allows us to capture complexities in the tax code that go beyond the information embedded in statutory marginal tax rates. We focus on the tax system as of 2019.

From Theory to Data. In our empirical application we use conditional revenue functions, where we condition on the other spouse's earnings y_{-i} belonging to a bracket B_{-i}. The relation between these conditional revenue functions characterized in Proposition 1 and the conditional revenue functions used in the empirical application arises via[29]

$$\mathcal{R}_i(y_i \mid B_{-i}) := \int_{y_{-i} \in B_{-i}} \mathcal{R}_i(y_i \mid y_{-i})\, dy_{-i} \,. \tag{2.10}$$

In our empirical application we deal with the US tax system which is piecewise linear. We, moreover, assume that variable effort costs are captured by an iso-elastic function.

[27] See Flood et al. (2021) and https://cps.ipums.org for a detailed description of CPS data. Appendix B.2.1 provides details on the data preparation. We use CPS data because it provides separate demographic and earnings information for both spouses. In contrast, the tax return micro data (SOI-PUF) from the Internal Revenue Service (IRS) used in Bierbrauer et al. (2021) does not contain this information (except for the year 1974; see Figure B.4 for a comparison).

[28] See Feenberg and Coutts (1993) and https://users.nber.org/~taxsim/ for detailed information on the TAXSIM microsimulation model, and Appendix B.2.1 for details on its combination with the CPS data.

[29] For later reference, we also define the unconditional revenue function

$$\mathcal{R}_i(y_i) := \int_{\mathcal{Y}_{-i}} \mathcal{R}_i(y_i \mid y_{-i})\, dy_{-i} \,, \tag{2.9}$$

which gives the revenue implication of raising the marginal tax rate at y_i, irrespective of the partner's earnings.

Finally, to simplify the analysis, we assume that choices in the status quo satisfy first-order conditions. From Proposition 1 and Equations (2.10) and (2.7), it then follows that

$$\mathcal{R}_2(y_{2s} \mid B_1) := s_1(B_1)\Big(I_2^d(y_{2s}, B_1) \; + \; M_2^{xd}(y_{2s}, B_1)\Big) \,, \tag{2.11}$$

where

$s_1(B_1) = F_{y_1}(\bar{b}_1) - F_{y_1}(\underline{b}_1)$ is the mass of couples with $y_1 \in B_1 = [\underline{b}_1, \bar{b}_1]$.

$$I_2^d(y_{2s}, B_1) \quad := \quad -\mathbb{E}_{y_1}\left[f_{y_2}(y_{2s} \mid y_1)\frac{T_{y_2}^0(y_1, y_{2s})}{1 - T_{y_2}^0(y_1, y_{2s})}\; y_{2s}\; \varepsilon_2 \mid y_1 \in B_1\right]$$

gives the revenue effect of behavioral responses at the intensive margin, where

$$M_2^{xd}(y_{2s}, B_1) \quad := \quad \mathbb{E}_{y_1}\Big[(1 - F_{y_2}(y_{2s} \mid y_1))(1 - \mathcal{E}_x^d(y_{2s} \mid y_1)) \mid y_1 \in B_1\Big]$$

is a term that compounds a simple tax reform's mechanical effect on tax revenue and the effect due to behavioral responses at the extensive margin.[30]

Conditioning Brackets. Our analysis in the main text is based on a specification of brackets based on quantiles of the income distribution: We condition on primary earnings belonging to specific deciles of the distribution of primary earnings. This gives us the possible values of B_1. Similarly, when constructing revenue functions conditional on secondary earnings, we consider the deciles of the secondary earner income distribution that will give us B_2.[31]

Behavioral Responses. We consider a baseline scenario with an intensive margin elasticity for primary earners (resp. secondary earners) of 0.25 (resp. 0.75). We also consider

[30] In Appendix B.2.2, we describe in detail, under which assumptions we estimate the ingredients of the conditional revenue function empirically. There, we also discuss the estimation of unconditional revenue functions. In particular, we assume that conditional revenue functions are constant over the brackets of interest. This constant value then coincides with the bracket average.

[31] See Table B.3 for the thresholds of the respective income deciles.

alternatives that are documented in Table B.2. We consider an extensive margin elasticity of 0.2. We show that our main findings also hold without extensive margin responses, i.e., an extensive margin elasticity of 0.

2.3.2 Inefficiencies

Figure 2.5 displays conditional revenue functions for primary and secondary earners.[32] Figure 2.5a shows where marginal tax rates on secondary earnings are inefficiently high: The inefficiencies are concentrated in the upper deciles of the distribution of secondary earnings. This is most pronounced when primary earnings are in a middle range. For instance, consider primary earnings above the median (Q6). The associated conditional revenue function is positive and monotonically decreasing until secondary earnings exceed \$70,000. Then, the function turns negative indicating inefficiently high marginal tax rates on secondary earnings. Hence, a reduction of secondary earners marginal tax rates for all couples with secondary earnings above \$70,000 and primary earnings in the sixth income decile would be self-financing.

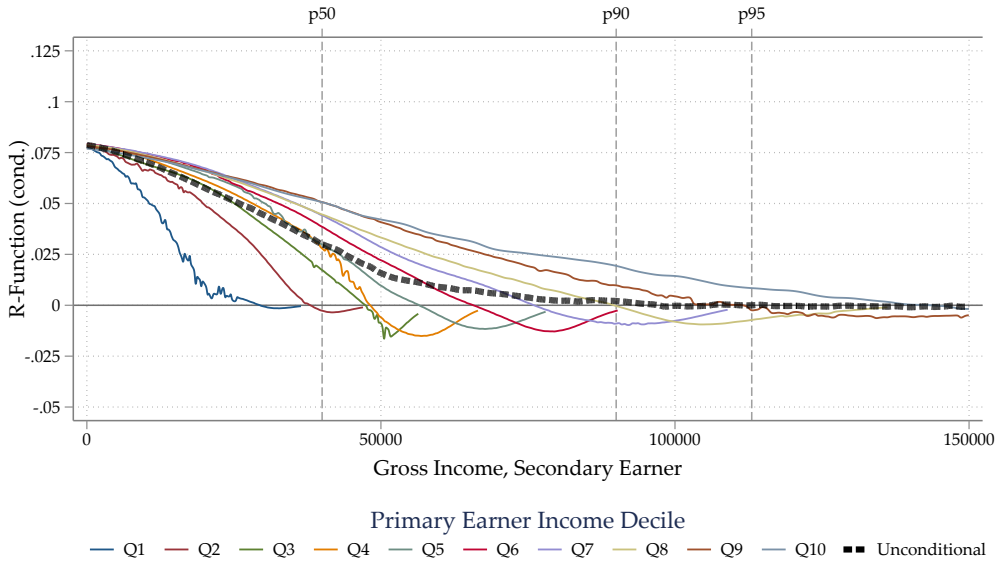
An inspection of the unconditional revenue function shows that these inefficiencies can only be eliminated with tax cuts that are tailored to the relevant decile of the primary earnings distribution. To see this, note that the unconditional revenue function at secondary earnings of \$70,000 is positive. Thus, cutting marginal tax rates on secondary earnings for all couples with y_2 close to \$70,000 would not be self-financing.

When applied to primary earnings, this test does not identify inefficiencies (see Figure 2.5b). Thus, it is not the case that marginal tax rates are inefficiently high across the board and by implication also for secondary earnings. They are too high only for secondary earnings which are close to primary earnings.

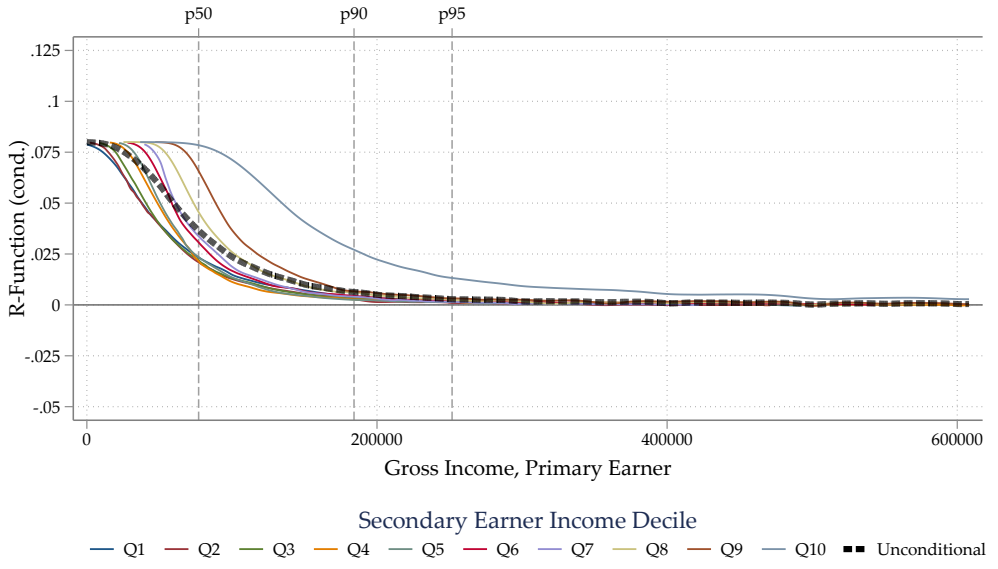
In the Appendix, we replicate the analysis by conditioning on earnings quintiles (Figure B.30), vingtiles (Figure B.31), and on statutory income tax brackets (Figure B.32). In line with the main figure, the magnitude of inefficiencies for secondary earners are largest in the middle of the primary earnings income distribution or, alternatively, for couples with primary earners in "middle" tax brackets.

[32]Empirical ingredients to these revenue functions, such as marginal tax rates, conditional income distributions, and intensive margin responses are reported in the Appendix, see Figures B.7 – B.16 (B.17–B.26).

Figure 2.5. – Conditional Revenue Functions, Deciles (2019)



(A) Secondary Earners



(B) Primary Earners

Note: This figure shows conditional revenue functions for secondary (resp. primary) earners in married dual earner couples conditional on primary (resp. secondary) earnings income deciles in Panel a (resp. b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (resp. 0.75) for primary (resp. secondary) earners and an extensive margin elasticity of 0.2. The figure also displays modified unconditional revenue functions for secondary and primary earners where unconditional revenue functions have been scaled by 0.1 to facilitate comparability with the conditional revenue functions. Results without extensive margin responses are shown in Figure B.27. Results for alternative elasticities (Table B.2) are shown in Figures B.28 and B.29. *Source:* Own calculations based on CPS-ASEC (2019).

2.3.3 Reducing Jointness

By Corollary 4, if the status quo tax system is efficient, a Rawlsian welfare measure objects to a reform that reduces jointness. However, as shown in the previous section, the status quo tax system is not efficient. It involves inefficiently high marginal tax rates in parts of the income space. Whether Rawlsian welfare is improved by a jointness-reducing reform is thus an open question that we discuss in this section. We also provide answers for alternative welfare measures.

We consider reductions of jointness tailored to specific combinations of income brackets for primary and secondary earnings. By Corollary 4, for τ small, the welfare implications of a jointness reducing reform are well approximated by

$$\begin{aligned} W &= \tau \int\limits_{B_1 \times B_2} \left[\int_{y_1}^{\bar{y}_1} \mathcal{W}_2(y_2 \mid y_1) dy_1 \; + \; \int_{y_2}^{\bar{y}_2} \mathcal{W}_1(y_1 \mid y_2) dy_2 \right] dy \\ &= \tau \int\limits_{B_1 \times B_2} \left[\mathcal{R}_2 \left(y_2 \mid [y_1, \bar{y}_1] \right) + \mathcal{R}_1 \left(y_1 \mid [y_2, \bar{y}_2] \right) - 2 \, s \left(y_1, y_2 \right) \mathcal{G} \left(y_1, y_2 \right) \right] dy \, , \end{aligned} \tag{2.12}$$

where $s\left(y_1, y_2\right)$ is the share of couples with primary earnings above y_1 and secondary earnings above y_2, and $\mathcal{G}(y_1, y_2) = \mathbb{E}[g(s_1, s_2) | s_1 \geq y_1, s_2 \geq y_2]$ is the average welfare weight for couples in this group. The second line in (2.12) expresses the welfare implications of a jointness-reducing reform using conditional revenue functions, which capture the loss in overall revenue and cumulative welfare weights which capture the gains of those who benefit from a lower tax burden.

If the relevant brackets are not too wide, we can approximate the above expression for W by

$$W = \tau \; \ell_1 \; \ell_2 \left[\mathcal{R}_2 \left(y_{2s} \mid [y_{1s}, \bar{y}_1] \right) + \mathcal{R}_1 \left(y_{1s} \mid [y_{2s}, \bar{y}_2] \right) - 2 \, s \left(y_{1s}, y_{2s} \right) \mathcal{G} \left(y_{1s}, y_{2s} \right) \right], \tag{2.13}$$

where y_{1s} and y_{s2} are, respectively, the income levels at which the relevant brackets start, whereas ℓ_1 and ℓ_2 are their respective lengths.

In our application of this formula, we consider brackets of lengths $l_1 = l_2 = 500$ and we set $\tau = -0.01$, corresponding to a one percentage point reduction of marginal tax rates. We select the thresholds y_{1s} for primary earnings based on the deciles of the primary earner income distribution, and the thresholds for secondary earnings y_{2s} based on the deciles of the distribution of secondary earnings conditional on the respective primary earner income decile. The relevant bracket combinations are shown in Table B.6. We, moreover, focus on specific welfare functions that enable us to highlight the trade-offs associated with a

reduction of jointness. Specifically, we consider a welfare function with equal weights for all couples. With preferences that are quasi-linear in consumption, welfare is then simply a measure of the total economic surplus. Changes in this welfare measure indicate how the distortions due to the behavioral responses to taxation change. We also consider welfare functions that, for lack of a better term, we call quasi-Rawlsian. Those welfare functions assign positive weights only to couples with joint income $y_1 + y_2$ in the bottom decile of the joint income distribution of dual earner couples. This welfare function indicates whether a tax reform is in the interest of "the poor". Finally, we consider "feminist" welfare functions with weights that increase in the women's share in a couples joint income. This welfare function indicates whether a reform is in the interest of "working women".[33]

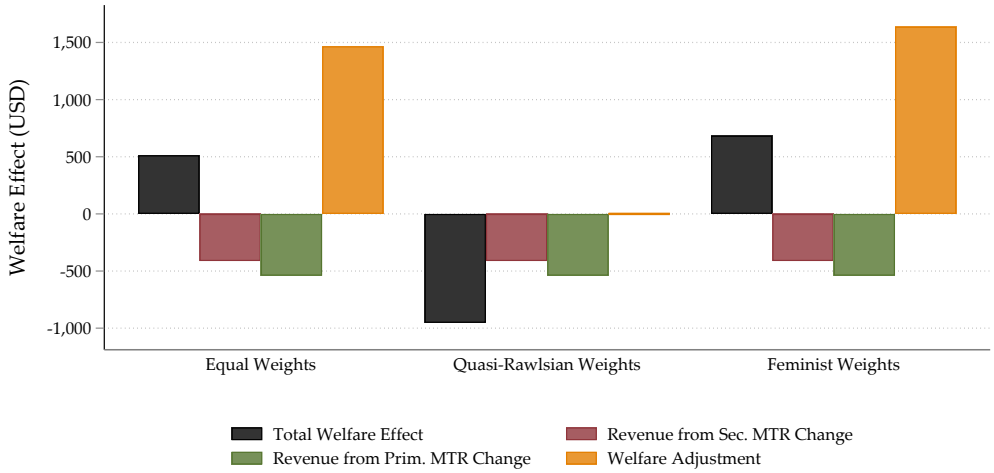
Reducing Jointness Around the Median. Figure 2.6 illustrates the welfare effect for a reform that reduces jointness in $B_1 \times B_2$, where B_1 starts at the median level of primary earnings and y_2 starts at the median of secondary earnings conditional on $y_1 \in B_1$. It provides a decomposition into the various components that appear on the right hand side of Equation (2.13). The left hand side corresponds to the black bar. The Figure shows that the reform is rejected by a quasi-Rawlsian welfare function and approved both by a Feminist one and by one with equal weights.

Note that the revenue effects coming from the respective changes in secondary and primary earner marginal tax rates do not necessarily point into the same direction. For instance, when marginal tax rates are inefficiently high, e.g. under strong behavioral responses and for relatively rich secondary earners (see Figure B.34), the (conditional) cut of secondary earner marginal tax rates might lead to a revenue gain. The aggregate revenue and welfare effects then depend on the relative magnitude of revenue effects coming from the tax rate changes on primary earnings and on secondary earnings.

Different Combinations of Brackets. We can repeat the exercise across all combinations of primary and secondary earnings shown in Table B.6 and evaluate the welfare consequences of jointness-reducing reforms in the respective brackets. Figure 2.7 summarizes the results for all three welfare measures and all bracket combinations. Under equal weights, a reduction in jointness is always considered beneficial, no matter at which parts of the income distribution it is carried out. The positive welfare impact is larger if behavioral responses are stronger (see Figure B.35). The feminist welfare function gives a similar picture, but with more pronounced welfare gains for reform in the range of high

[33]For the detailed specification of welfare weights, see Table B.5. Appendix B.2.3 further provides detailed information on the required ingredients and construction of empirical welfare effects.

Figure 2.6. – Welfare, Reducing Jointness at the Median (2019)



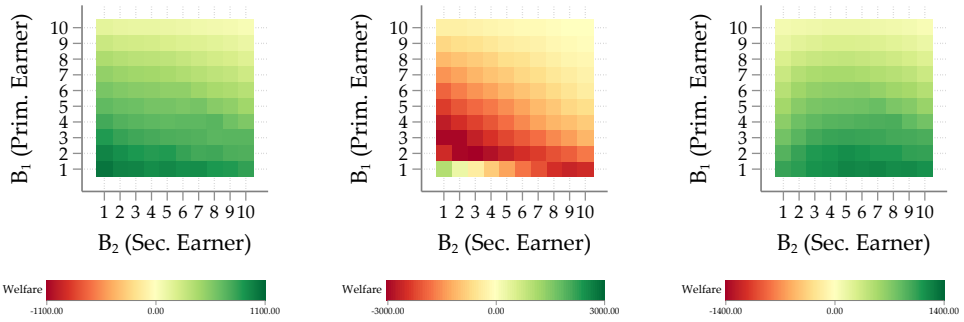
Note: This figure visualizes the welfare effects implied by a reduction in jointness as of 2019 for $(y_1, y_2) \in B_1 \times B_2$, where B_1 is the 6th decile of the distribution of primary earnings and B_2 is the 6th decile in the distribution of secondary earnings conditional on $y_1 \in B_1$. The aggregate welfare effect and its components is shown for three different welfare weights (see Table B.5) and is based on the bracket lengths $\ell_1 = \ell_2 = 500$ and $\tau = -0.01$. We take intensive margin responses according to the baseline elasticity scenario (see Table B.2) into account and also include extensive margin responses. The black bar indicates the aggregate welfare effect. The red (resp. green) bar visualizes the revenue change coming from the decrease of secondary (resp. primary) earnings. The orange bar illustrates the welfare adjustment corresponding to the last term in equation (2.13). For details on the decomposition of the welfare effect, see also Appendix B.2.3. Figure B.33 illustrates the conditional revenue functions, the two revenue effects are based on.
Source: Own calculations based on CPS-ASEC (2019).

secondary earnings. The quasi-Rawlsian welfare measure rejects reductions in jointness, except at the very bottom. The opposition is weak however for reductions of jointness in brackets with high secondary earnings as those reforms are close to self-financing (see Section 2.3.1).

2.3.4 Lowering Marginal Tax Rates at the Bottom

Proposition 2 points to the possibility that welfare functions with continuous weights and a Rawlsian welfare functions may give rise to different conclusions regarding the desirability of reforms that lower marginal tax rates at the bottom of the income distribution. In this section, we demonstrate that this situation is empirically relevant. Specifically, we will show that a strictly Rawlsian and a quasi-Rawlsian welfare function reach opposite conclusions.

Figure 2.7. – Welfare Implications of Decreasing Jointness (2019)



(A) Equal (B) Rawlsian (C) Feminist

Note: This figure displays the welfare implications of decreasing jointness as of 2019 by decreasing secondary and primary marginal tax rates conditional on particular brackets of primary earnings and secondary earnings. Bracket thresholds are shown in Table B.6. The reform applies to a bracket of length $l(B_1) = l(B_2) = 500$ and has the magnitude $\tau = -0.01$. The figure distinguishes between three different forms of welfare weights (see Table B.5) and is based on intensive margin responses according to the baseline elasticity scenario (see Table B.2). All results are shown including extensive margin responses. Results for different elasticity scenarios and without extensive margin responses are shown in Figures B.35 and B.36.
Source: Own calculations based on CPS-ASEC (2019).

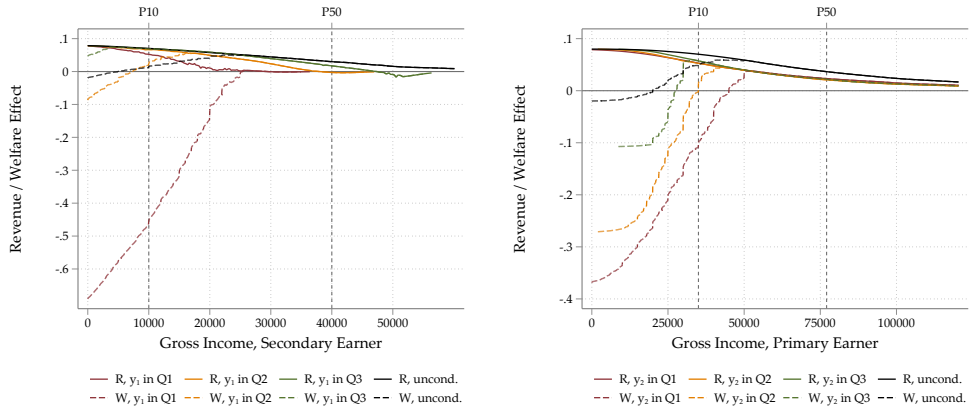
Our analysis makes use of the empirical conditional revenue functions that were already used in our analysis of inefficiencies in Section 2.3.1. To get to the welfare implications, we extend the approach that was used to construct these functions to the conditional welfare functions that were introduced in Section 2.2.[34]

Figure 2.8 shows results for a reduction of marginal tax rates at a specific level of secondary earnings conditional on primary earnings belonging to the three bottom deciles of the primary earnings distribution. It also presents analogous results for a lowering of marginal tax rates on primary earnings. The solid lines are the conditional revenue functions from Section 2.3.1 which describe the revenue effects of the tax change. The dashed lines indicate the welfare effect of the tax change. The figure also displays the revenue and welfare effects of reforms that lower marginal tax rates on secondary or primary earnings unconditionally (black lines).

We can extract three insights about tax reforms at low incomes. First, in line with the theoretical reasoning in Proposition 2, the quasi-Rawlsian welfare function takes negative values at low secondary earnings conditional on low primary earnings (Figure 2.8a). Thus, reducing marginal tax rates is desirable according to this welfare measure. This is also true at low primary earnings conditional on low secondary earnings (Figure 2.8b). Second, the conditioning on the other spouse's earnings makes an important difference.

[34]Our empirical implementation of the conditional welfare functions uses equation (2.8) and then follows our empirical implementation of the conditional revenue functions.

Figure 2.8. – Welfare Effects of Reforms at the Bottom (2019)



(A) Secondary Earnings (B) Primary Earnings

Note: This figure shows the revenue and welfare effects for secondary (resp. primary) earners in married dual earner couples conditional on primary (resp. secondary) earnings income deciles at the bottom of the income distribution in Panel a (resp. b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (resp. 0.75) for primary (resp. secondary) earners and an extensive margin elasticity of 0.2. The figure displays conditional revenue functions (solid lines, also displayed in detail in Figure 2.5) and the welfare effect based on quasi-Rawlsian welfare weights according to Table B.5 (dashed lines). Black lines indicate unconditional revenue and welfare functions. Results without extensive margin responses are shown in Figure B.38. Results for different elasticity scenarios in line with Table B.2 are shown in Figure B.39.
Source: Own calculations based on CPS-ASEC (2019).

An unconditional quasi-Rawlsian welfare function takes negative value over a much more narrow income range at the very bottom.[35] Third, lowering marginal tax rates is costly in terms of tax revenue and therefore rejected by a strictly Rawlsian welfare function. This holds true both for conditional and for unconditional revenue functions. Lowering marginal tax rates would thus be akin to an introduction of workfare elements into the tax-and-transfer system: The intercept of the consumption schedule goes down but its slope goes up as the net-of-tax-rate increases.

2.3.5 Optimal Top Tax Rates

Recall that under the premises of Proposition 3 top tax rates for primary and secondary earners are given by

$$\frac{t_i^*}{1-t_i^*} = (1-\bar{g}_i)\,\frac{1}{\bar{\alpha}_i}\frac{1}{\varepsilon_i} \qquad i=1,2\,. \tag{2.14}$$

[35]When excluding extensive margin responses, the unconditional quasi-Rawlsian welfare function is positive throughout, while the unconditional ones take negative values (Figure B.38)— a fact driven by the behavior of (un)conditional welfare weights (Figure B.37).

Table 2.1. – Pareto Coefficients and Optimal Tax Rates (2018)

	P95		P99	
	Primary	Secondary	Primary	Secondary
Panel A: Pareto coefficients				
Wages on W2 form	1.83	3.06	1.75	3.12
Panel B: Optimal tax rates				
Elasticity = .25	69%	57%	70%	56%
Elasticity = .5	52%	40%	53%	39%
Elasticity = .75	42%	30%	43%	30%

Note: This table shows Pareto coefficients and optimal top tax rates. Panel A displays Pareto coefficients for primary and secondary earners based on a generalized Pareto interpolation using tabulated data on wages on W2 forms for joint return taxpayers with wage income. We distinguish between interpolations at the 95th and the 99th percentile. Panel B displays optimal top tax rates associated with these Pareto coefficients. We distinguish between different elasticities.
Source: Own calculations based on SOI Tax Stats - Individual Information Return Form W-2 Statistics (2018).

We seek to trace out the implications of differences in the Pareto-coefficients for primary and secondary earnings for top tax rates. We therefore assume, for ease of exposition, that "rich people" have welfare weights of zero, $g_1 = g_2 = 0$. As in the previous sections, in our baseline, we assume that $\varepsilon_1 = 0.25$ and $\varepsilon_2 = 0.75$.[36]

Pareto Coefficients. We estimate Pareto coefficients for primary and secondary earners among jointly filing tax units using tabulated income tax data from the Statistics of Income (SOI) program of the Internal Revenue Service (IRS).[37] We use the tabulated tax return data and apply generalized Pareto interpolation techniques by Blanchet et al. (2022) to estimate properties of the wage income distribution. In the main text, we focus on 2018, the most recent available year for this data. Panel A of Table 2.1 shows the estimated Pareto coefficients at the 95th and 99th percentile of the wage income distribution. At both percentiles, the Pareto coefficient of the primary earner wage income distribution is around 40 percent lower than the respective coefficient for secondary earners.

Results. Based on the estimated Pareto coefficients, Panel B of Table 2.1 displays optimal top tax rates for different elasticities. Even under the assumption that $\varepsilon_1 = \varepsilon_2$ at the top of the income distribution, the optimal tax rates for primary earners are around 12 percentage points higher than the one for secondary earners.[38]

[36] We proceed under Assumption 2. In particular, we take elasticities to be constant over the relevant part of the income distribution. An implication is that the Pareto coefficient does not vary with the tax rates at the top.

[37] For details on the data source and preparation, see Appendix B.2.1.

[38] In Appendix B.2.5, we show that this gap is also present when using CPS data to estimate Pareto coefficients and optimal top tax rates.

As shown by Figure B.40, for the time period where tabulated tax data for primary and secondary earners is available (2008-2018), the optimal top tax rate gap stayed relatively constant. In Appendix B.2.5, we also provide heterogeneity analyses that estimate top tax rates for primary earners separately for single earner and dual earner couples. The results show that marginal tax rates for primary earners in single earner couples should be higher than for primary earners in dual earner couples, since the Pareto coefficients of the former group are much lower. In particular, optimal top tax rates for primary earners in dual earner couples should be closer to the secondary earners optimal top tax rate.

2.4 CONCLUDING REMARKS

In this paper, we have introduced a theory of tax reforms for a setting with (i) multi-dimensional heterogeneity amongst taxpayers, (ii) multiple economic decisions that are all subject to both fixed and variable costs, and therefore involve multiple discontinuities in the behavioral responses to taxation, and (iii) a tax function that treats all these decisions as separate arguments. For such a setting, the theorems in this paper provide a complete characterization for the existence of Pareto- or welfare-improving tax reforms.

We then explored the implications of these Theorems for a particular policy design problem, the taxation of couples. For this application, we developed sufficient statistics that can be used to check (i) whether and where marginal tax rates on secondary earnings are inefficiently high, (ii) whether and where changes to the tax system's degree of jointness would be welfare-improving, (iii) which welfare functions approve and which welfare functions reject reforms that introduce or strengthen workfare elements at the bottom of the income distribution, and (iv) to characterize optimal marginal tax rates at the top of the income distribution.

Finally, we provided answers to these questions against the background of the tax treatment of couples in the US. We showed that (i) marginal tax rates on secondary earnings are inefficiently high in parts of the income distribution, (ii) that a reduction of jointness is approved by Feminist welfare measures, but rejected by Rawlsian ones, (iii) that "more" workfare is desirable for a large class of social welfare functions, including ones that concentrate welfare weights at the bottom, but rejected by a Rawlsian welfare function, and (iv) at the top, marginal tax rates on secondary earners should be lower than marginal tax rates on primary earnings (even if "rich men" and "rich women" receive the same welfare weights and show the same behavioral responses to taxation) because the distribution of primary earnings has fatter tails than the distribution of secondary earnings.

Chapter 3

The Political Economy of Joint Taxation

This chapter is based on co-authored work with Felix Bierbrauer, Pierre Boyer, and Andreas Peichl. A previous version has been published as a CEPR Discussion Paper. See Bierbrauer et al. (2023b) and Bierbrauer et al. (2025) for the full reference.

3.1 INTRODUCTION

A basic lesson from the theory of optimal taxation is that the intensity of behavioral responses matters for the design of taxes. The larger the elasticities that capture those behavioral responses, the greater the distortionary effects of taxation. Consequently, whatever the welfare function, optimal taxes are lower when elasticities are larger. As has been noted by Boskin and Sheshinski (1983), when applied to the taxation of married couples, this implies that primary earnings should be taxed at a higher rate than secondary earnings. Secondary earners are more likely to work part-time, if at all, and these choices are more responsive to tax incentives than those of primary earners who mostly work full time. Recent literature has moreover shown that tax incentives play an important role in explaining differences in female labor force participation across countries.[1] In some countries – including France, Germany and the US – the incomes of the spouses in a couple are taxed jointly; that is, the tax base is the sum of primary and secondary earnings. A direct implication is that the marginal tax rates on primary and secondary earnings are equal. This suggests that it is possible to improve welfare and raise more tax revenue by a reform that leads away from *joint taxation.* Such a reform would also bring more women to the labor market and improve the earnings incentives of those women who already are on the labor market. If all that is so clear, why are these reforms not taking place?

In this paper, we want to see whether political economy forces can provide an explanation. There is a rich body of literature on the political economy of taxation, which we discuss in more detail below. A focal question of this literature has been how changes in the inequality of market incomes affect the demand for redistributive policies.[2] To the best of our knowledge, there is no prior literature on the political economy of joint taxation. A contribution of this paper therefore is to add gender and marital status to the political economy of taxation. In recent decades, the population shares of unmarried individuals and of individuals living in dual-earner couples have gone up, whereas the population share of single earner couples has gone down. Consequently, the population share of those who benefit the most from the marriage bonuses that are implied by joint taxation has gone down. This leads to the question whether the support for alternatives to joint taxation has increased over time and whether such alternatives might eventually become politically feasible.

[1] See, e.g., Gustafsson (1992), Blundell and Macurdy (1999), Blau and Kahn (2007), Eissa and Hoynes (2004), LaLumia (2008), Kaygusuz (2010), Bargain et al. (2014), Selin (2014), and Neisser (2021).

[2] The Meltzer-Richard hypothesis is a well-known result from this literature: If the gap between mean and median income goes up, the size of government is predicted to go up as a result.

In the theory part, we develop tools for such an analysis. These tools can be used for an analysis of hypothetical reforms, in particular for reforms that, in line with the prescriptions of optimal tax theory, lower marginal tax rates for secondary earners and increase them for primary earners. Getting there requires a detailed analysis of how primary and secondary earners respond to changes of the tax system, and an aggregation that allows us to keep track of implications for tax revenue. We then exploit the mathematical properties of functional derivatives to determine the extent to which marginal tax rates on secondary earnings can be reduced if the marginal tax rates on primary earnings are slightly increased, ensuring the reform remains revenue-neutral. Once we have all of this available, we can draw the line between couples who are made better off and couples who are made worse off if such a reform is undertaken. This constitutes the paper's methodological contribution.[3] While it focuses on the marginal effects of reforms that stay in a vicinity of the status quo, we show that our approach can be extended to large reforms and this enables us to evaluate not just hypothetical reforms, but also reforms that actually took place. We use these insights in the empirical part of the paper. Specifically, we analyze what the major reforms of the federal income tax in the US since the 1960s implied for the relative tax treatment of singles, single-earner and dual-earner couples. This empirical analysis of joint taxation in the US is the paper's substantive contribution.

Theory. We consider a status quo tax system with joint taxation. Couples are assumed to maximize their joint surplus, defined as the difference between the couples' disposable income and the spouses' effort costs. There are both fixed and variable costs of productive effort so that behavioral responses to taxation are at the intensive and at the extensive margins. As we formally show, with joint and progressive taxation as the status quo, there is an interdependence of primary and secondary earnings. If, say, the marginal tax rate on secondary earnings goes down, and secondary earnings increase as a response, then joint earnings go up and hence also the tax rate faced by the primary earner, who now has an incentive to adjust his or her earnings in response. We provide a detailed analysis of such behavioral responses to tax reforms and their implications for overall tax revenue.

We assume that revenue changes are rebated lump-sum. Consequently, a reform is politically feasible – in the sense of being preferred by a majority of individuals over the status

[3]This approach of first constructing revenue-neutral reforms and then checking their political feasibility or their implications for social welfare can also be applied to other policy design problems. Thus, the paper's methodological contribution is not tied to an analysis of joint taxation. It could, for instance, be used, to analyze changes of the mix between direct and indirect taxes. Bierbrauer (2024) uses a related approach for an evaluation of emission-neutral changes of climate policy.

quo – if the extra tax revenue outweighs increases of the tax burden for at least half of the population. A reform is desirable according to some measure of social welfare if the gains of the reform winners are weighted higher than the losses of the reform losers.

Revenue-neutral reforms towards individual taxation are of particular interest to us. These are reforms that reduce marriage bonuses: Marginal tax rates on secondary earnings are decreased and marginal tax rates on primary earnings are increased, with the tax changes calibrated such that the reform is revenue-neutral overall. We derive simple, intuitive, and easily applicable sufficient-statistics formulas that can be used to evaluate such reforms from both from a political economy and a social welfare perspective.

The Political Economy of Joint Taxation in the US. Our empirical analysis of past reforms in the US makes use of the Current Population Survey (CPS) and of NBER's TAXSIM microsimulation model. We go through the eleven major reforms of the federal income tax since the 1960s and present a detailed analysis of how marriage penalties and bonuses were affected.[4]

We do not identify systematic changes in the tax treatment of married couples relative to singles from 1960 to today. The distribution of marriage bonuses in the recent past looks, by and large, very much like the distribution of marriage bonuses in the early 1960s. Moreover, none of those reforms led away from joint taxation. As of today and as of 1960, the tax base for married couples is the joint income.

Our analysis of hypothetical reforms that lead away from joint taxation provides a potential explanation. We show that the share of married individuals who would benefit from a revenue-neutral reform towards individual taxation has been increasing since the 1960. The share was 20 percent in 1960 and surpassed the 50 percent threshold only recently. Thus, even today, moving away from joint taxation may provoke substantial political backlash.

It is also not clear that such a reform would be desirable from a welfare perspective. Revenue-neutral reforms towards individual taxation create winners and losers. Losers are couples with the lion's share of the joint income being due to the primary earner. For

[4]We present a more detailed analysis of theses reforms in part C.3 of the Appendix. There, we analyze separately for singles and for married couples where the reform winners and the reform losers were located in the income distribution. We also look into whether the reform winners formed a majority and we check whether social welfare measures went up or down.

such couples, the increase of the tax rates on primary earnings is the dominant effect. The lower rates on secondary earnings can mitigate, but not offset, this effect. Winners, by contrast, are couples with secondary earnings close to primary earnings.

Single-earner couples are more concentrated in the bottom deciles of the couples' income distribution. Consequently, a Rawlsian social welfare function would not approve a revenue-neutral reform towards individual taxation. By contrast, an "Affirmative Feminist" social welfare function – one that assigns greater weight to couples where the woman has higher earnings – would approve it. This shows that, given a status quo with joint taxation, a reduction of marriage bonuses may be controversial also from a welfare perspective.

The reforms discussed so far reduce marriage bonuses in a particular way: They drive a wedge between the marginal tax rate of the primary earner and the marginal tax rate of the secondary earner. This wedge is the same everywhere in the income distribution. We also explore reform directions that break with individual taxation in a different way. Specifically, we show that a revenue-neutral reform that increases marginal tax rates on primary earnings only in the upper half of the income distribution and lowers the tax rates on secondary earnings across the board, would make roughly 70 percent of all married individuals better off. The reform would moreover be approved both by a Rawlsian and by a Feminist welfare measure.

Finally, we investigate reform directions that are inspired by reforms that took place in Sweden in 1971 and the UK in 1990. Both reforms introduced individual taxation. In Sweden, the reform was accompanied by tax cuts at the bottom of the income distribution. In the UK, there were tax cuts for all incomes. We find that, as of today, a "Swedish reform" would achieve majority support in the US, whereas a "British reform" would not.

The general insight that emerges from these analyses is that breaking with joint taxation can be a hardship for single-earner couples at the bottom of the income distribution. A reform towards individual taxation that seeks to achieve both majority support and an unambiguous welfare improvement therefore needs to make sure that these couples are not made worse off.

Outline. The next section discusses related literature. Section 3.3 introduces a conceptual framework to analyze tax reforms, given a status quo with joint taxation. We present an empirical analysis of historical tax reforms in the US in Section 3.4. We develop tools for an analysis of revenue-neutral reforms towards individual taxation in Section 3.5

and apply them to data in Section 3.6. Concluding remarks can be found in Section 3.7. Formal proofs, additional empirical findings and robustness checks are relegated to Appendices.

3.2 RELATED LITERATURE

Seminal contributions to the political economy of taxation have focused on linear taxes, see Roberts (1977) and Meltzer and Richard (1981), and established median-voter results.[5] More recent contributions have focused on the complications associated with non-linear income taxation while maintaining an emphasis on redistribution from "the rich" to "the poor".[6]

We bring a political economy perspective to the rich literature that studies the optimal taxation of couples. The literature following Boskin and Sheshinski (1983) has covered non-linear taxes and a wide range of behavioral responses.[7] We analyze potential reform directions in the vicinity of a status quo tax system that has been inherited from the past.[8] Kleven et al. (2009) and Golosov and Krasikov (2025) approach the optimal taxation of couples as a problem of multi-dimensional screening. Bierbrauer et al. (2024) study a related setup using a dual rather than a primal approach and provide conditions under which Pareto- or welfare improving reforms can be found.[9]

This paper is related to Bierbrauer et al. (2021) in that it does not provide a game-theoretic analysis of political competition over non-linear taxes. Instead, it focuses on the existing status quo tax system and examines whether reforms to this system could garner majority support. There are also important differences, however. Here, we bring in heterogeneity in

[5]See, for a review, Acemoglu et al. (2015).

[6]See, e.g., Farhi et al. (2012), Scheuer and Wolitzky (2016), Acemoglu et al. (2008), Acemoglu et al. (2010), Bierbrauer and Boyer (2015), Brett and Weymark (2017), and Bierbrauer et al. (2022). See, for a review, Berliant and Boyer (2024).

[7]See, e.g., Kleven et al. (2009), Immervoll et al. (2011), Cremer et al. (2012), Gayle and Shephard (2019), Malkov (2021), Alves et al. (2024), Ales and Sleet (2022), and Golosov and Krasikov (2025).

[8]This perturbation approach is frequently used in optimal tax theory. References include Piketty (1997), Saez (2001), Golosov et al. (2014), Saez and Stantcheva (2016), Lorenz and Sachs (2016), Sachs et al. (2020), Jacquet and Lehmann (2021b), Jacquet and Lehmann (2021a), Bergstrom and Dodds (2021), Bierbrauer et al. (2023a), Ferey et al. (2024), or Spiritus et al. (forthcoming). Gender-based taxation, see Alesina et al. (2011), is a related topic.

[9]A complementary strand of literature examines the implications of joint taxation in quantitative dynamic models see, e.g. Guner et al. (2012), Guner et al. (2014), Bick and Fuchs-Schündeln (2017), Borella et al. (2023b), Borella et al. (2023a), and Olsson (2025). Holter et al. (2023) use such a framework to show that the transition from joint to individual taxation comes with an increase in the government's ability to generate tax revenue.

marital status and in the composition of joint earnings, whereas Bierbrauer et al. (2021) only distinguish "tax units" that differ in income. Moreover, they focus on a class of monotonic tax reforms, which have the property that majority support coincides with support from tax units with incomes close to the median. This characterization does not apply to the reforms which are at the center in this paper.

Theorem 1 and Proposition 1 below constitute this paper's theoretical contribution. Together these results clarify the conditions under which revenue-neutral reforms towards individual taxation are politically feasible and /or desirable from a social welfare perspective. The proofs use some insights from Bierbrauer et al. (2024). Otherwise, the papers do not overlap.

We complement our theoretical analysis with an empirical approach using the TAXSIM microsimulation model and CPS microdata.[10] The microsimulation model leverages detailed data on individual characteristics, allowing us to determine the implications of tax reforms on individual welfare at the tax-unit level.

3.3 THE MODEL

The Status Quo. We consider a status quo tax system in which married couples are taxed according to their joint income. Formally, the status quo tax system consists of two tax functions. The tax function that applies to singles is denoted by $T_{s0} : y_s \mapsto T_{s0}(y_s)$, where y_s is a single's before-tax income. Married couples are taxed according to the function $T_{m0} : y_m \mapsto T_{m0}(y_m)$, where $y_m = y_1 + y_2$ is the couple's joint income, y_1 is the income of the primary earner and y_2 is the income of the secondary earner.[11] We assume that T_{0s} and T_{0m} are increasing, continuous and convex.

A tax reform replaces the status quo tax functions (T_{s0}, T_{m0}) by new tax functions (T_{s1}, T_{m1}). We introduce a formalism for the analysis of such tax reforms below. Before that, we clarify how we measure marriage penalties and bonuses.

[10] Our empirical approach builds on and extends work by Eissa et al. (2008), Bargain et al. (2015) and Bierbrauer et al. (2021). Similar approaches have also been used for the purpose of ex-ante policy evaluation, see Immervoll et al. (2007) for a prominent example.

[11] For the purposes of the theory, "primary" and "secondary" are labels that carry no further meaning. In our empirical application, we assign the label "primary earner" to the spouse whose earnings are higher.

Marriage Penalties and Bonuses. Given a tax system (T_s, T_m), we define the splitting function $y_m \mapsto \sigma(y_m)$ so that $\sigma(y_m)$ is the solution to

$$T_m(y_m) \quad = \quad \sigma(y_m)\, T_s\left(\frac{y_m}{\sigma(y_m)}\right) . \tag{3.1}$$

For instance, the tax system in Germany has $\sigma(y_m) = 2$, for all y_m. That is, all married couples are taxed as if they consisted of two singles who each contribute fifty percent to the joint income. The splitting function $\sigma : y_m \mapsto \sigma(y_m)$ allows for more general forms of income splitting. The interpretation is that married couples are taxed is as if each partner was assigned a fraction $\frac{1}{\sigma(y_m)}$ of the couple's joint income, and then the couple is treated as if it had a number of $\sigma(y_m)$ individuals who are all taxed according to the schedule for singles. We can then say that a couple with joint income y_m benefits from a marriage bonus if

$$\sigma(y_m)\, T_s\left(\frac{y_m}{\sigma(y_m)}\right) \quad < \quad T_s(y_1) + T_s(y_2) ,$$

and suffers from a marriage penalty with the reverse inequality. With progressive taxation, $\sigma(y_m) \geq 2$ implies a marriage bonus and $\sigma(y_m) \leq 1$ implies a marriage penalty, for all possible triplets (y_m, y_1, y_2) with $y_1 > y_2$. When $\sigma(y_m) \in (1, 2)$, there is an intermediate value $\bar{\sigma}(y_m)$ which solves

$$\bar{\sigma}(y_m)\, T_s\left(\frac{y_m}{\bar{\sigma}(y_m)}\right) = T_s(y_1) + T_s(y_2) ,$$

so that there is neither a bonus nor a penalty. The intermediate value not only depends on the married couple's total income but also on the income share of the primary earner. Specifically, for given y_m, $\bar{\sigma}(y_m)$ decreases in the income share of the primary earner.[12] Thus, given y_m and $\sigma(y_m) \in (1, 2)$, spouses with rather unequal incomes benefit from a marriage bonus and spouses with more equal incomes suffer from a marriage penalty.

[12] To see this, let $\pi = \frac{y_1}{y_m}$ be the income share of the primary earner and write

$$\bar{\sigma}(y_m)\, T_s\left(\frac{y_m}{\bar{\sigma}(y_m)}\right) = T_s(\pi\ y_m) + T_s((1-\pi) y_m) .$$

Employing the implicit function theorem and using the convexity of T_s makes it possible to verify that $\bar{\sigma}(y_m)$ is decreasing in π.

In our empirical analysis of joint taxation in the US, we first estimate the splitting function and then document how it has shifted over time, with upward shifts implying that more couples benefited from marriage bonuses and downward shifts implying that more couples suffered from marriage penalties (see Figure 3.2 below).[13]

Tax Reforms. We introduce a framework for an analysis of tax reforms in this setting. We then distinguish reforms *in* this system and reforms *of* this system. Reforms *in* the system yield changes of tax rates while the tax base for married couples does not change. It's the joint income. Consequently, the primary and the secondary earner face the same marginal tax rate both before and after the reform. Reforms *of* the system, by contrast, drive a wedge between the marginal tax rates on primary earnings and those on secondary earnings.

A tax reform replaces the status quo tax functions (T_{s0}, T_{m0}) by new tax functions (T_{s1}, T_{m1}) so that

$$T_{s1}(y_s) = T_{s0}(y_s) + \tau_s \; h_s(y_s) \; , \quad \text{and}$$

$$T_{m1}(y_1, y_2) = T_{m0}(y_m) + \tau_m \; h_m(y_1, y_2) \; .$$

We refer to the functions $h_s : y_s \mapsto h_s(y_s)$ and $h_m : (y_1, y_2) \mapsto h_m(y_1, y_2)$ as reform directions, whereas the scalars $\tau_s \geq 0$ and $\tau_m \geq 0$ are measures of reform intensity. For some of our analysis, we focus on reforms that stay in the vicinity of the status quo. Then, τ_s and τ_m are close to zero.

A reform *in* the system is such that h_m is a function of $y_m = y_1 + y_2$. All changes in marriage penalties and bonuses that occurred in the US since the 1960s were implied by reforms *in* the system, in particular by changes in the relative tax treatment of couples and singles. A reform *of* the system, by contrast, breaks with joint taxation. To give an example, let

$$h_m(y_1, y_2) = \tau_1 \; y_1 + \tau_2 \; y_2 \; ,$$

with $\tau_1 > 0$ and $\tau_2 < 0$. Then, after the reform, the marginal tax rate on primary earnings is higher than in the status quo, and the marginal tax rate on secondary earnings is lower:

$$\frac{\partial T_{m1}(y_1, y_2)}{\partial y_1} \quad = \quad T'_{m0}(y_1 + y_2) + \tau_m \; \tau_1 \quad > \quad T'_{m0}(y_1 + y_2) \; , \quad \text{and}$$

$$\frac{\partial T_{m1}(y_1, y_2)}{\partial y_2} \quad = \quad T'_{m0}(y_1 + y_2) + \tau_m \; \tau_2 \quad < \quad T'_{m0}(y_1 + y_2) \; .$$

[13] For details on the estimation of the empirical splitting function and its ingredients, see Appendix C.2.3.

Earnings Choices. We assume that there is a bounded set of feasible earnings levels $\mathcal{Y} = [0, \bar{y}]$, where $\bar{y}$ can be arbitrarily large. Singles choose $y_s \in \mathcal{Y}$ so as to maximize $C_0(y_s) - K_s(y_s, \theta_s)$, where a single's disposable income is given by

$$C_s(y_s) \quad = \quad b_s + y_s - \Big(T_{s0}(y_s) + \tau_s \, h_s(y)\Big)$$

and a single's costs of productive effort is given by

$$K_s(y_s, \theta_s) \quad = \quad k_s(y_s, \omega_s) + \phi_s \, \mathbf{1}(y_s > 0) \; .$$

We let $T_{0s}(0) = h_s(0) = 0$. Hence, b_s is the intercept of the consumption schedule that singles are facing, or, equivalently, the transfer to a single with no earnings. To allow for behavioral responses to taxation both at the intensive and the extensive margin, we assume that the generation of earnings comes with fixed and variable costs. The latter are captured by the function k_s which is increasing in the first argument, $k_{s,1} > 0$, satisfies the usual Inada conditions and, moreover, is such that $k_{s,12} < 0$, i.e., marginal effort costs decrease in ω_s. We refer to ω_s also as a single's productive ability. The fixed cost of generating positive earnings is denoted by ϕ_s. We write $\theta_s = (\omega_s, \phi_s)$ for a single's type and denote the utility-maximizing earnings choice by $y_s^*(\tau_s, h_s, \theta_s)$. We denote the pre-reform choice by $y_s^0(y_s) = y_s^*(0, h_s, \theta_s)$. The distribution of θ_s in the subpopulation of singles is represented by a *cdf* F_s.

The spouses in a married couple solve the following problem: Choose $y = (y_1, y_2)$ to maximize $C_m(y) \; - K(y, \theta)$, where the couple's disposable income is given by

$$C_m(y) \quad = \quad b_m \; + \; y_1 + y_2 \; - \; \Big(T_{m0}(y_1 + y_2) + \tau_m \, h_m(y_1, y_2)\Big) \; .$$

and its costs of productive effort are given by

$$K(y, \theta) \quad = \quad k_1(y_1, \omega_1) \; + \; \varphi_1 \, \mathbf{1}(y_1 > 0) \; + \; k_2(y_2, \omega_2) \; + \; \varphi_2 \, \mathbf{1}(y_2 > 0) \; .$$

Again, we assume that $T_{0m}(0, 0) = h_m(0) = 0$. The spouses costs functions are defined analogously to the ones for singles. Note, however, that the cost functions k_s, k_1 and k_2 are not assumed to be the same. A couple is characterized by a pair $\omega = (\omega_1, \omega_2)$ of productive abilities and a pair of fixed costs $\varphi = (\varphi_1, \varphi_2)$. We will sometimes write for short $\theta = (\omega, \varphi)$ with $\theta_1 = (\omega_1, \varphi_1)$ and $\theta_2 = (\omega_2, \varphi_2)$. We denote the utility-maximizing earnings choices by $y^*(\tau_m, h_m, \theta)$ and write $y^0(\theta) = (y_1^0(\theta), y_2^0(\theta))$ for earnings in the status quo. The joint distribution of θ_1 and θ_2 is denoted by F_θ.

Under the additional assumption that disposable income is a public good for the spouses, the assumption that couples maximize their joint surplus is consistent with them engaging in Nash bargaining to determine their respective contributions to the couples' joint income.[14]

The mass of tax units is normalized to 1, with the shares of singles and married couples denoted respectively by ν_s and $\nu_m = 1 - \nu_s$. The mass of individuals is therefore $\nu_s + 2\,\nu_m$.

Preferences over Tax Reforms. A generic tax reform is henceforth represented by (τ, h) with $h = (h_1, h_2)$ and $\tau = (\tau_s, \tau_m)$. We derive preferences over tax reforms from indirect utility functions. Let $V_s(\tau, h, \theta_s)$ be the indirect utility realized by a single of type θ_s after a tax reform (τ, h) has taken place. We denote the pre-reform level by $V_{0s}(\theta_s)$. Then $V_s(\tau, h, \theta_s) - V_{0s}(\theta_s)$ is the reform-induced change in indirect utility for a single with characteristics θ_s. Analogously, $V(\tau, h, \theta) - V_0(\theta)$ is the reform-induced change in indirect utility experienced by the spouses in a married couple with characteristics θ.

We denote by $R(\tau, h)$ the change in overall tax revenue due to the reform. This is an endogenous object that will be more fully characterized below. We assume that changes in tax revenue, if any, are rebated lump-sum at the tax unit level. Together with the assumption that the joint after-tax income is a public good for the spouses in a married couple, this implies that all individuals draw the same benefits from changes in overall tax revenue. Differences in preferences over tax reforms are therefore entirely due to differences in how a reform affects individual tax burdens.[15]

Political Feasibility. Given a tax reform (τ, h), we measure the support amongst singles by

$$S_s(\tau, h) \quad := \quad \mathbf{E}_{\theta_s}\left[\mathbf{1}\Big(V_s(\tau, h, \theta_s) - V_{0s}(\theta_s) > 0\Big)\right] ,$$

i.e., by the fraction of singles who are made better off. Analogously, the opposition amongst singles is given by

$$O_s(\tau, h) \quad := \quad \mathbf{E}_{\theta_s}\left[\mathbf{1}\Big(V_s(\tau, h, \theta_s) - V_{0s}(\theta_s) < 0\Big)\right] .$$

We define support and opposition amongst married individuals analogously and write $S(\tau, h)$ for the fraction of married couples that benefit from a reform and $O(\tau, h)$ for the fraction of married couples that are made worse off. We say that a reform achieves,

[14] See Appendix A.10 of Bierbrauer et al. (2024).

[15] A more general formalism that entails the possibility of different tax rebates for singles and for married couples can be found in Bierbrauer et al. (2023b).

respectively, majority support amongst singles or married individuals if $S_s(\tau, h) > O_s(\tau, h)$ or $S(\tau, h) > O(\tau, h)$. We say that a reform is politically feasible if the reform winners outnumber the reform losers, i.e. if

$$\nu_s\Big(S_s(\tau,h) - O_s(\tau,h)\Big) \; + \; 2\nu_m\Big(S(\tau,h) - O(\tau,h)\Big) \quad > \quad 0\,.$$

This notion of political feasibility requires that a reform be preferred by a sufficiently large fraction of individuals over the status quo. For concreteness, we identify "sufficiently large" with "majority support". Political feasibility then holds provided that the reform would win a majority against the status quo if there was a referendum with universal participation by all taxpayers. In part C.3.2 of the Appendix we provide detailed information on which of the major tax reforms in the US achieved (i) majority support amongst singles, (ii) majority support amongst married individuals, and (iii) majority support amongst all taxpayers. The next section contains a summary of this analysis.

Social Welfare. The change in welfare due to a tax reform is given by

$$\begin{aligned} \mathcal{W}(\tau,h) \;\; &= \;\; \nu_s \; \mathbf{E}_{\theta_s}\Big[g_s(\theta_s)\Big(V_s(\tau,h,\theta_s) - V_{0s}(\theta_s)\Big)\Big] \\ &\quad + 2\nu_m \mathbf{E}_{\theta}\Big[g(\theta)\Big(V(\tau,h,\theta_m) - V_0(\theta)\Big)\Big]\;, \end{aligned} \tag{3.2}$$

where the functions $\theta_s \mapsto g_s(\theta_s)$ and $\theta \mapsto g(\theta)$ specify, reflectively, welfare weights for singles and married individuals as functions of their characteristics θ_s and $\theta = (\theta_1, \theta_2)$.

In part C.3.3 of the Appendix, we present a detailed evaluation of the major tax reforms in the US using various social welfare functions. We summarize the results below.

Marginal Effects. The derivatives of V_s and V with respect to τ, evaluated at $\tau = 0$ give respectively, the marginal effect of reform in direction h on the utility realized by singles and couples. By the envelope theorem (see Milgrom and Segal (2002)):

$$\frac{\partial}{\partial\tau}V_s(\tau,h,\theta_s)_{|\tau=0} \;\; := \;\; V_{s,\tau}(h,\theta_s) \quad = \quad R_\tau(h) - h_s(y_s^0(\theta_s))\;, \tag{3.3}$$

and

$$\frac{\partial}{\partial\tau}V(\tau,h,\theta)_{|\tau=0} \;\; := V_\tau(h,\theta) \quad = \quad R_\tau(h) - h_m(y^0(\theta))\;, \tag{3.4}$$

where $R_\tau(h) := \frac{\partial}{\partial\tau}R(\tau,h)_{\tau=0}$ is the reform's marginal impact on tax revenue, evaluated at the status quo, i.e. for $\tau = 0$. These equations show that whether a tax unit benefits from a reform simply depends on how the change of transfers compares to the change

in the tax obligation due to the reform. A single benefits when $R_\tau(h) - h_s(y_s^0(\theta_s)) > 0$ and is made worse off otherwise. Likewise, the spouses in a married couple benefit if $R_\tau(h) - h_m(y^0(\theta)) > 0$ and lose otherwise.

Normalizing welfare weights so that $\nu_s \, \mathbf{E}_{\theta_s} [g_s(\theta_s)] + 2\nu_m \mathbf{E}_\theta [g(\theta)] = 1$, the marginal effect on welfare associated with reform direction h is given by

$$\begin{aligned} \mathcal{W}_\tau(h) \quad &= \quad R_\tau(h) - \nu_s \, \mathbf{E}_{\theta_s} [g_s(\theta_s) h_s(y_s^0(\theta_s)] \\ &\qquad - 2\nu_m \mathbf{E}_\theta [g(\theta) h_m(y^0(\theta))] \quad . \end{aligned} \tag{3.5}$$

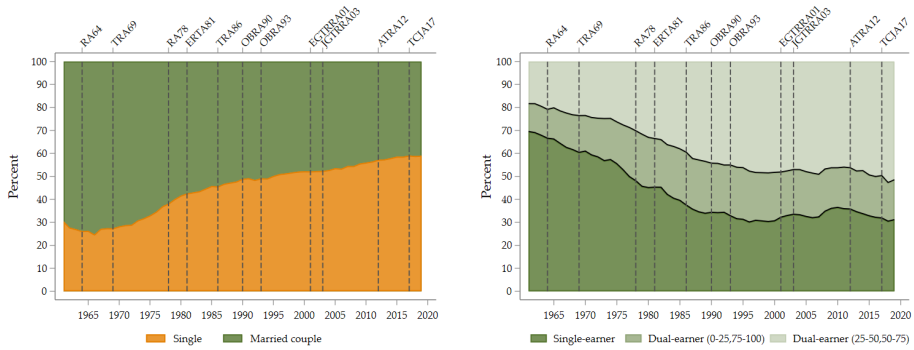
An extension of our framework that enables an analysis of large reforms can be found in part C.1.3 of the Appendix. From our analysis of small reforms, we know the marginal effects of a given reform *direction* on the indirect utilities of singles and couples. We then integrate over these marginal effects, with the bounds of integration reflecting the *scale* of the reform. We thereby obtain upper and lower bounds for a large reform's impact on individual welfare. From there, an analysis of the political economy and the welfare implications of large reforms is straightforward.

3.4 MARRIAGE PENALTIES AND BONUSES IN THE US SINCE THE 1960S

In the following, we use data from the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC) combined with the TAXSIM (v32) microsimulation model.[16] Thereby, we obtain detailed information on the characteristics of tax units, such as their sources of income or their number of children. This detailed information enables us to compute what taxes a single or a married couple has to pay under a given tax system T_0. It also makes it possible to compute what taxes they would have to pay under an alternative tax system T_1 and hence, the change in the tax burden associated with a reform that replaces T_0 by T_1.

[16] See Flood et al. (2021) and https://cps.ipums.org for a detailed description of CPS data. Appendix C.2.1 provides details on the data preparation. We use CPS data because it provides separate demographic and earnings information for both spouses. In contrast, the tax return microdata (SOI-PUF) from the Internal Revenue Service (IRS) used in Bierbrauer et al. (2021) does not contain this information (except for the year 1974; see Figure C.1 for a comparison). See Feenberg and Coutts (1993) and https://users.nber.org/~taxsim/ for detailed information on the TAXSIM microsimulation model.

Figure 3.1. – Demographic Change over Time



(A) Tax Unit Types (B) Couple Types

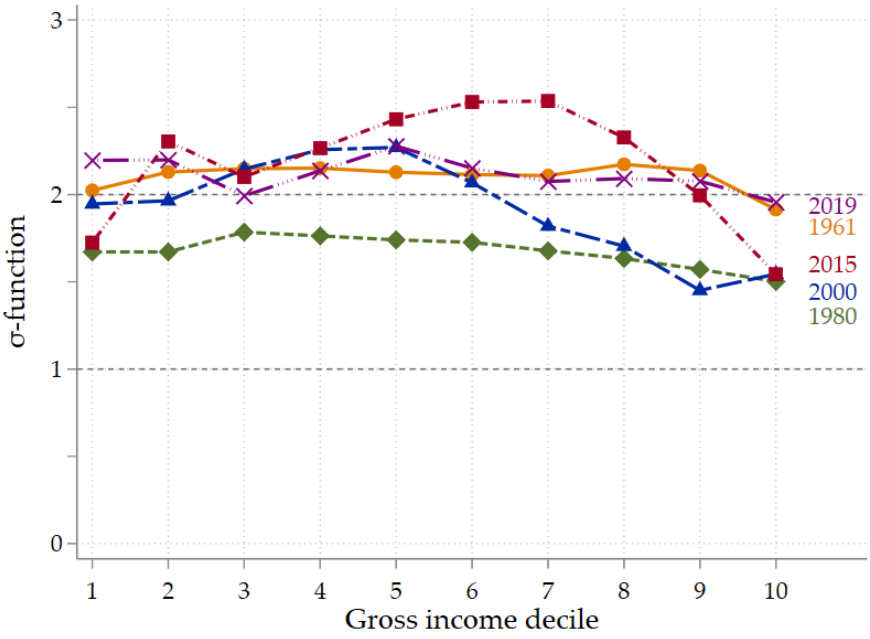
Note: This figure shows the distribution of tax unit types over time. Figure 3.1a displays the share of single tax units (orange area) and the share of couple tax units (green area). Figure 3.1b displays the share of single-earner and dual-earner couples. A single-earner couple refers to a married couple, in which one spouse is not employed (dark green area). The figure further displays the share of dual-earner couples in which both spouses are employed and (i) one spouse earns between 0 and 25 percent (mid green area) and (ii) between 25 and 50 percent of total earnings (light green area). Earnings shares are computed on the basis of wage, business and farm income. Reforms of the federal income tax code are displayed as vertical lines. All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old. Figure C.38 replicates this figure for the full adult population. *Source:* Own calculations based on CPS-ASEC.

Demographic Trends. Since the 1960s, the share of singles relative to married couples has increased in the US. Also, the share of dual earner couples has increased relative to single-earner couples. These changes have taken place continuously (see Figure 3.1) and have implications for the taxation of families, which we discuss below.

The Splitting Function over Time. Figure 3.2 shows the splitting-function σ in selected years – years close to reforms that had a significant impact on marriage penalties and bonuses – to give an indication of how marriage bonuses and penalties have evolved over time due to changes in the tax system.[17] The fact that σ is nowhere below one indicates that, throughout, couples with very unequal incomes benefited from a marriage bonus, irrespective of whether they belonged to the upper or the lower part of the income distribution. Interestingly, the splitting function in the recent past looks similar to the one from the early 1960s: it is, by and large, close to 2 for all levels of income, indicating that the occurrence of penalties and bonuses does not systematically vary with income. In between, there have been pronounced departures from this pattern. For instance, in 2015 there were larger marriage bonuses in the upper middle class. In 2000, there were marriage

[17]The estimated σ-function relates the mean average tax rates across *all* singles (baseline) to the mean average tax rates across all couples (comparison). However, average tax rates can vary within the group of couples and singles, most notably due to the presence of children. Figure C.10 explicitly differentiates between different baseline and comparison groups by estimating σ while varying the number of children in the baseline and comparison group. For details, see Appendix C.2.3.

Figure 3.2. – The Splitting Function σ over Time



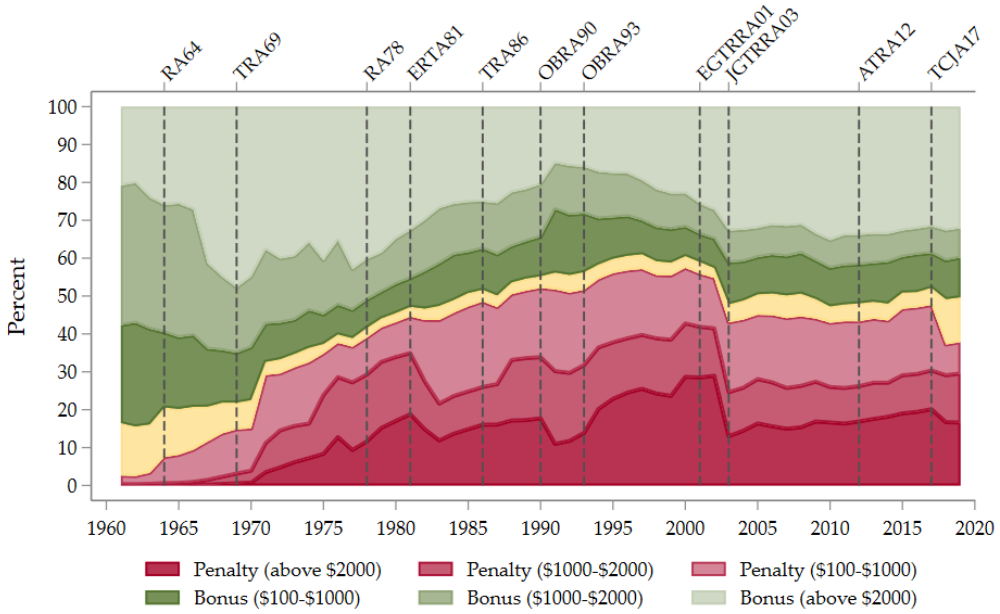
Note: This figure shows estimates of the splitting function σ for selected years. The σ-function is calculated for tax units by estimating mean average tax rates of couples and singles (see Figure C.9). Mean average tax rates are used to solve numerically for σ (see Appendix C.2.3). Deciles refer to the gross income distribution of couples in the respective year. All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old. Figure C.39 replicates this figure for the full adult population.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

penalties for "rich" dual earner couples with high secondary earnings. In Appendix C.3.1, we show how specific tax reforms shifted the splitting function, and how marriage bonuses and penalties changed in particular parts of the income distribution.

The fact that σ in the recent past looks, by and large, very much like the one in the early 1960s does not mean that the distribution of marriage bonuses have been constant. Since the 1960s, Figure 3.3 illustrates that the share of couples experiencing a marriage bonus has decreased, while the share of couples facing a marriage penalty has increased. The picture mirrors the change in couple types displayed in Figure 3.1b.

Political Economy and Welfare Implications. Figure 3.4 reports the change in tax liability net of tax revenue changes, accounting for behavioral responses and lump-sum adjustment of any loss or gain in tax revenue (top panel) for four selected reforms. It also documents, respectively, the shares of singles and couples who benefited from various tax reforms (bottom panel). In addition, it shows the positions of the reform winners and

Figure 3.3. – Distribution of Marriage Bonuses and Penalties in USD over Time



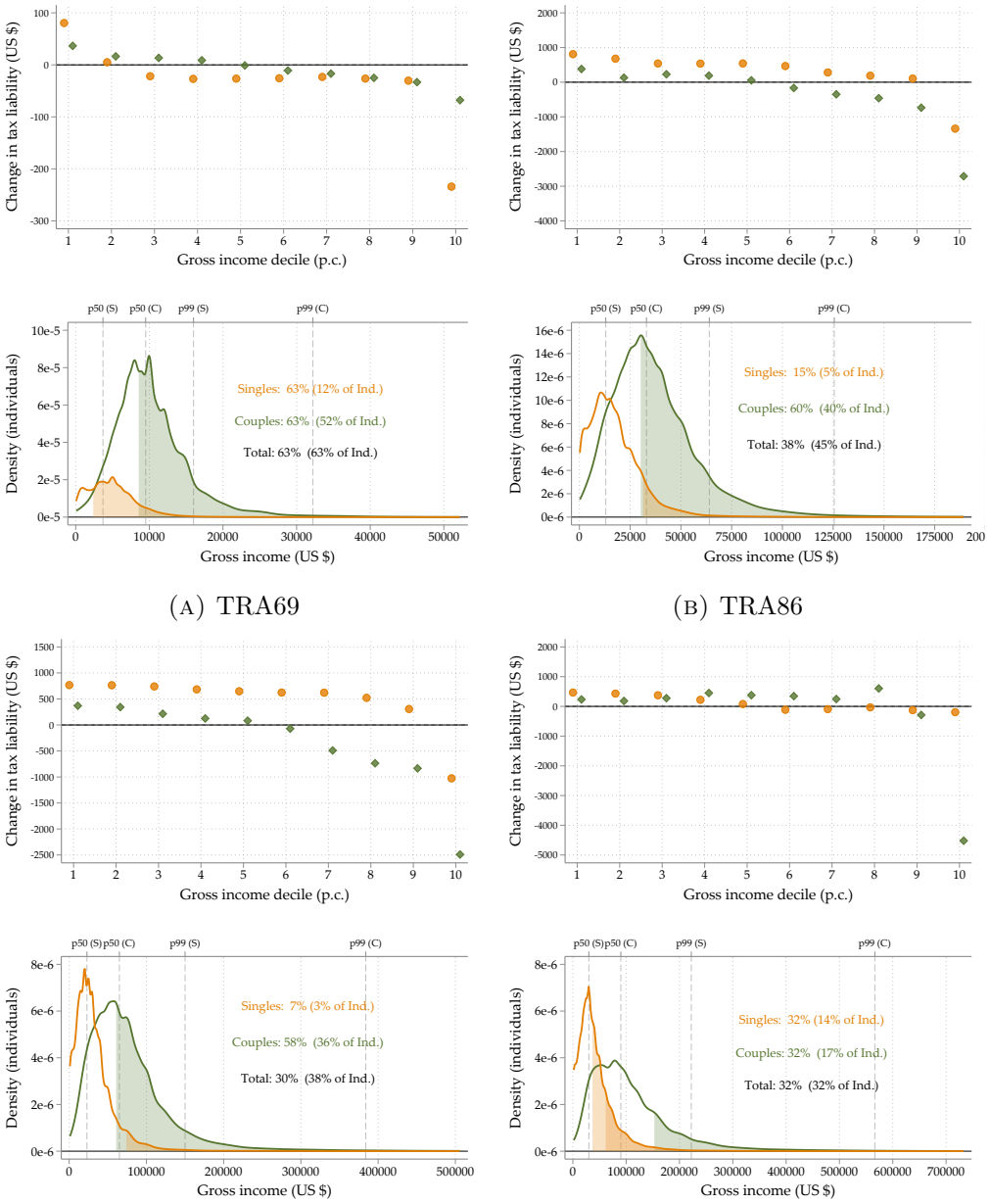
Note: This figure shows how the distribution of absolute marriage bonuses and penalties changed over time. Marriage bonuses and penalties have been constructed by estimating for every married couple a counterfactual tax burden of two singles with their respective individual incomes. The counterfactual tax burden is an average over two tax burdens that allocate dependent to either spouse. Absolute marriage bonuses are CPI-adjusted and measured in 2019 US dollars. Reforms of the federal income tax code as described in Table C.4 are displayed as vertical lines. All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

losers in the income distribution.[18] In part C.3.2 of the Appendix we report in detail – i.e. separately for each of the 11 major reforms – how much support they had amongst singles and amongst married individuals. Here we summarize the key results.

Most tax reforms – nine out of eleven – reduced marginal tax rates. They, moreover, reduced marginal tax rates in a monotonic way, i.e., richer taxpayers benefited from larger tax savings than poorer taxpayers. Under our baseline assumptions about the elasticities that govern the behavioral responses to taxation, those reforms made both singles and married individuals in the upper deciles of the income distribution better off, whereas individuals in the bottom deciles of the income distribution were made worse off. Exceptions from this pattern are two tax reforms by the Clinton and Obama administrations.

[18]Figure C.29 in the Appendix contains a robustness check for the analysis in this section. It assumes that tax is rebated lump sum at the individual level, rather than at the tax unit level. While this modification leads to changes in political support among different tax unit types, aggregate support is qualitatively similar.

Figure 3.4. – Political Feasibility



(A) TRA69 (B) TRA86

(C) JGTRRA03 (D) TCJA17

Note: This figure shows the change in tax liability net of tax revenue changes (upper panel) and winners of the reform (lower panel) for singles (orange shaded area) and couples (green shaded area). The upper panel shows the average change in tax liability per capita (PC) for each of the 10 per capita gross income deciles. The location of winners and losers across the income distribution are visualized by evaluating the average tax liability change for 25 gross income quantiles. The share of winners among tax units and the corresponding share of individuals is shown next to the distribution. We account for behavioral responses at the intensive margin (baseline elasticity scenario from Table C.5) and the extensive margin. Tax revenues are rebated lump sum at the tax unit level. Figures for all reforms are shown in Figure C.27. Figure C.28 shows an alternative analysis based on lump-sum adjustments at the individual level. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Figure C.40 replicates this figure for the full adult population.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

These reforms led to higher marginal tax rates at the top of the income distribution. They made individuals in the bottom deciles of the income distribution better off and individuals in the top deciles worse off.

Out of the eleven tax reforms between 1964 and 2017, six received majority support under our baseline assumptions. While tax reforms that increased tax rates received high levels of political support, the picture is mixed for reforms that cut tax rates. For instance, the 1969 tax cut by the Nixon administration made 63 percent of all individuals better off, with political support coming from the upper parts of the income distribution (see Figure 3.4a). Other reforms like the Reagan tax cut in 1986 or the 2003 tax reform by the Bush administration did achieve majority support among couples, but failed to obtain support in the electorate at large, in particular because only few single households benefited from the reform (see Figures 3.4b and 3.4c). In contrast, as shown by Figure 3.4d, the Trump tax cut did not achieve majority support in either of the two groups and only benefited rich households.

We also evaluate these reforms with various measures of social welfare (see Appendix C.3.3). The reforms involving tax cuts are rejected by Rawlsian social welfare functions, i.e. by welfare functions that concentrate welfare weights at the bottom of the income distribution. They are approved, however, by a welfare function with equal weights. The reforms by the Clinton and Obama administrations show the reverse pattern.

We evaluate reforms also using feminist welfare measures, i.e., welfare measures that concentrate weights on women and, for married women, have weights that are increasing in the woman's share in the couple's joint income. These measures can, in particular, be used to trade-off the welfare of single mothers at the bottom of the income distribution and the welfare of married women. We find that some of the tax cutting reforms are approved by feminist welfare measures when we consider pronounced behavioral responses to taxation. The stronger the behavioral responses, the lower is the loss of tax revenue associated with a tax cut. Thus, with strong behavioral responses, the poor single mothers do not suffer too much when married women in the upper deciles of the income distribution benefit from a tax cut.

Summary. Over time there has not been a fundamental change in marriage penalties and bonuses in the United States. While there have been significant adjustments in between, the distribution of marriage penalties and bonuses in the recent past resembles their counterpart from the early 1960s. As we document in part C.6 of the Appendix, marriage penalties and bonuses have been an important topic in election campaigns and,

more broadly, the public discourse about tax policy. In the following section we explore a potential explanation for why despite of all these debates the tax base continued to be the sum of primary and secondary earnings: Reforms towards individual taxation lack political feasibility.

3.5 REFORMS TOWARDS INDIVIDUAL TAXATION

We focus on a particular class of reforms that we refer to as *revenue neutral reforms towards individual taxation*. We study how the political support for such reforms has evolved since the 1960s and we also study their appeal from a welfare perspective. Subsequently, we discuss other ways of reforming the status quo tax system towards individual taxation.

One-Bracket Reforms. Two classes of reforms play a significant role in our analysis. We refer to them, respectively, as one-bracket reforms affecting primary or secondary earnings. More general tax reforms can be interpreted as combinations of several such reforms. So, as a preliminary step, we introduce these particular classes of reforms.

A one-bracket reform of primary earnings can represented by a pair (τ_1, h_1) with

$$\tau_1 \; h_1(y_1') = \begin{cases} 0, & \text{for} \quad y_1' \leq y_1 \; , \\ \tau_1(y_1' - y_1), & \text{for} \quad y_1' \in (y_1, y_1 + \ell) \; , \\ \tau_1 \; \ell, & \text{for} \quad y_1' \geq y_1 + \ell \; . \end{cases}$$

Thus, when primary earnings y_1' are in a bracket of length ℓ that starts at y_1, then the marginal tax rate on primary earnings is increased by τ_1. When primary earnings are smaller than y_{1s} the couples' tax burden is unaffected. When they exceed $y_1 + \ell$, there is no change of the marginal tax rate, but the couples tax burden is larger. A "small" one-bracket reform has τ_1 and ℓ close to zero. We denote by $\mathcal{R}_1 : y_1 \mapsto \mathcal{R}_1(y_1)$ the revenue implications of such a small reform.[19]

One bracket reforms that alter marginal tax rates only for secondary earnings are denoted by (τ_2, h_2) and defined in the analogous way. Their revenue implications are captured

[19]More formally, let $R_1(\tau_1, \ell, y_1)$ be the revenue from a one-bracket reform, as a function of y_1 where the relevant bracket starts, the length ℓ of the bracket, and the change of marginal tax rates within the bracket, τ_1. Then,

$$\mathcal{R}_1(y_1) \quad := \quad \lim_{\ell \to 0} \frac{\partial}{\partial \ell} \lim_{\tau_1 \to 0} \frac{\partial}{\partial \tau_1} R_1(\tau_1, \ell, y_1) \; .$$

by the function $\mathcal{R}_2 : y_2 \mapsto \mathcal{R}_2(y_2)$. Characterization of the functions $\mathcal{R}_1$ and $\mathcal{R}_2$ via sufficient statistics formulas can be found below (Proposition 4); i.e. these functions can be computed computed using data on the status quo tax system, income distribution, and intensive and extensive margin elasticities that capture behavioral responses to taxation. In part C.2.5 of the Appendix we explain in detail how we obtain the calibrations of these functions that we use in our empirical analysis.

From Simple Reforms to More General Ones. When the functions $\mathcal{R}_1$ and $\mathcal{R}_2$ are known, the revenue implication of any continuous reform direction $y_1 \mapsto h_1(y_1)$ or $y_2 \mapsto h_2(y_2)$ can be computed via[20]

$$R_\tau(h_1) = \int_{\mathcal{Y}} h_1'(y_1)\mathcal{R}_1(y_1)\, dy_1 \quad \text{and} \quad R_\tau(h_2) = \int_{\mathcal{Y}} h_2'(y_2)\mathcal{R}_2(y_2)\, dy_2 \,. \tag{3.6}$$

Revenue-Neutral Reforms Towards Individual Taxation. A revenue-neutral reform towards individual taxation raises the marginal tax rates on primary earnings and lowers the marginal tax rates on secondary earnings. Moreover, the increased revenue from the higher taxes on primary earnings is used to finance the tax cuts for secondary earners. Revenue neutrality implies, in particular, that such a reform is without consequence for singles. It has distributive effects only among married couples. It tends to make couples with a rather equal within-couple distribution better off at the expense of couples with a dominant primary earner. Formally, we consider reform directions so that

$$h_m(y_1, y_2) = \tau_1\, h_1(y_1) + \tau_2\, h_2(y_2) \,,$$

with τ_1 and τ_2 chosen such that

$$\begin{aligned} R_\tau(h_m) &= \tau_1 \int_{\mathcal{Y}} h_1'(y_1)\mathcal{R}_1(y_1)\, dy_1 \;+\; \tau_2 \int_{\mathcal{Y}} h_2'(y_2)\mathcal{R}_2(y_2)\, dy_2 \\ &= 0 \,. \end{aligned}$$

The Uniform Case. A special case of interest is that marginal tax rates are increased for all primary earners and decreased for all secondary earners: $h_1(y_1) = y_1$, for all y_1 and $h_2(y_2) = y_2$, for all y_2. If the status quo has joint taxation, such a reform drives a wedge between the marginal tax rate on primary earnings and the marginal tax rate on secondary earnings which is the same for all possible combinations of primary and sec-

[20] A formal proof can be found in Appendix A.3 of Bierbrauer et al. (2024).

ondary earnings. For now, we focus on this special, but discuss alternative specifications below. Such a reform is revenue neutral if

$$\frac{\tau^2}{\tau^1} \quad = \quad -\frac{\int_{\mathcal{Y}} \mathcal{R}_1(y_1)dy_1}{\int_{\mathcal{Y}} \mathcal{R}_2(y_2)dy_2} \quad := \quad -r \ . \tag{3.7}$$

We will repeatedly refer to the ratio on the right hand side of Equation (3.7) in the following. For ease of reference, we use r as a shorthand.

Political Feasibility. A married couple that has earnings of y_1^0 and y_2^0 in the status quo is made better off if $\tau_1\ y_1^0 + \tau_2\ y_2^0 < 0$ or, equivalently, if $y_1^0 < r\, y_2^0$. This inequality will prove useful for our analysis of whether reforms towards individual taxation would have had majority support in the US since the 1960s. Specifically, we will plot the line $y_1^0 = r\, y_2^0$ in a y_2^0-y_1^0-diagram. All couples with (y_2^0, y_1^0) below the line are reform winners, all couples with (y_2^0, y_1^0) above are reform losers. To determine political feasibility, we simply need to check whether the households above the line outnumber those below the line. To check how political feasibility has evolved, we look into how this line and the distribution of primary and secondary earnings has shifted over time.[21]

Welfare Implications. We also examine the implications for social welfare, employing various welfare functions. We will focus on welfare functions with high weights on "the poor" and on welfare functions with weights that increase in secondary earnings.

We use Equation (3.5) for the welfare evaluation. Since the reform is revenue-neutral and affects only married individuals, such a reform raises social welfare if and only if

$$\mathbf{E}_\theta[g(\theta)h_m(y^0(\theta))] = \tau_1 Y_1^g \ + \ \tau_2 Y_2^g > 0,$$

where $Y_1^g := \mathbf{E}_\theta[g(\theta)y_1^0(\theta)]$ and $Y_2^g := \mathbf{E}_\theta[g(\theta)y_2^0(\theta)]$. Thus, from the perspective of a generic welfare function, a revenue neutral reform with $h_m(y_1, y_2) = \tau_1 y_1 + \tau_2 y_2$ is desirable if and only if $Y_1^g < rY_2^g$. The following Theorem summarizes these results (the proof is in Appendix C.1.1).

[21] Note that r shifts with the behavioral responses that shape the functions $\mathcal{R}_1 : y_1 \mapsto \mathcal{R}_1(y_1)$ and $\mathcal{R}_2 : y_2 \mapsto \mathcal{R}_2(y_2)$. The less elastic primary earnings are relative to secondary earnings, the larger is r.

Proposition 4. *Consider a revenue neutral reform towards individual taxation with* $h_1(y_1) = y_1$*, for all* y_1*, and* $h_2(y_2) = y_2$*, for all* y_2*.*

(i) Such a reform direction is politically feasible if the share of married couples with $y_1^0 < r\, y_2^0$ *exceeds* $\frac{1}{2}$*.*

(ii) Such a reform direction is welfare-improving if $Y_1^g < r\, Y_2^g$*.*

While the reform considered in Theorem 4 is, by construction, revenue neutral, if $Y_1^g < rY_2^g$, then it is also possible to design a reform that extracts positive revenue, without violating the condition $Y_1^g < rY_2^g$. Since the reform, moreover, lowers the marginal tax rates for secondary earners, it also increases female labor supply. If (ii) holds and (i) fails then such a reform is desirable according to the given social welfare function, but not politically feasible. If (i) and (ii) both hold, then reform is both desirable and politically feasible.

Revenue Functions: From Theory to Empirics. In order to bring our theory to data, we state the revenue functions under two further assumptions in the following Proposition: First, the piecewise linear, as is the case in the US. Second, the primary and the secondary earners' effort costs are, respectively, represented by isoleastic functions and the Frisch elasticities governing the primary and secondary earners' intensive margin responses are denoted by ε_1 and ε_2. Extensive margin responses are captured by $\bar{\pi}_{dec}(y_2 \mid y_1)$ and $\bar{\pi}_{sec}(y_2)$, defined respectively as average extensive margin elasticities for dual-earner and single-earner couples.

Proposition 4. *When the tax system is piecewise linear and effort costs functions are iso-elastic, then the revenue functions are*

$$\mathcal{R}_2(y_2) \;=\; \mathcal{I}_2(y_2) + \mathcal{X}_{2,sec}(y_2) + \mathcal{X}_{2,dec}(y_2)\;,$$

where

$$\mathcal{I}_2(y_2) \;=\; -y_2\varepsilon_2 f_2(y_2)\mathbf{E}_{y_1}\left[\tfrac{T_m'(y_1+y_2)}{1-T_m'(y_1+y_2)} \;\middle|\; y_2\right]\;,$$

$$\mathcal{X}_{2,sec}(y) \;=\; F^{y_1}(0)\;\textstyle\int_{y_2}^{\bar{y}} \frac{T^0(y_2)}{y_2-T^0(y_2)}\;\bar{\pi}_{sec}(y_2)\;f(y_2 \mid 0)\;dy_2\;,$$

and

$$\mathcal{X}_{2,dec}(y) = (1-F^{y_1}(0))\,\mathbf{E}_{y_1}\left[\int_{y_2}^{\bar{y}} \frac{T^0(y_1+y_2)}{y_1+y_2-T^0(y_1+y_2)}\bar{\pi}_{dec}(y_2 \mid y_1)f(y_2 \mid y_1)dy_2 \mid y_1 > 0\right].$$

Analogously for $\mathcal{R}_1(y_1)$*.*

3.6 EMPIRICAL ANALYSIS OF REFORMS TOWARDS INDIVIDUAL TAXATION

We first study hypothetical revenue-neutral reforms towards individual taxation. These reforms have distributive effects only among married couples. The empirical analysis shows us the location of the line that separates the more "modern" couples that benefit from such a reform from the more "traditional" couples that are made worse off. Moreover, we check whether the conditions for political feasibility and welfare improvements in Theorem 4 are satisfied at different points in U.S. history. Second, we turn to real-world reforms towards individual taxation implemented in Sweden and the United Kingdom. We show how our setup can be used to analyze such reforms. Specifically, we check how much political support such reforms would generate in the U.S. today.

Calibration. In Appendix C.2.5, we explain in detail how we calibrate the revenue functions $\mathcal{R}_1$ and $\mathcal{R}_2$. Here we elaborate on what we assume about the elasticities that capture the behavioral responses to taxation. Our assumptions shown in Table 3.1 are guided by the empirical literature that finds stronger behavioral responses to taxation for secondary earners (see, e.g., Eissa and Hoynes (2004) and Bargain et al. (2014)) while acknowledging the variation of estimates (see, e.g., Blau and Kahn (2007), Saez et al. (2012), Neisser (2021)).

In our baseline scenario, we assume that intensive margin elasticities are constant over time and equal 0.25 for primary earners in couples, and 0.75 for secondary earners.[22] We also consider a scenario with elasticities that are higher than the ones in the baseline, and one with lower elasticities. We finally assume that the extensive margin elasticities are homogeneous across different types of couples, and decrease with income from 0.65 to 0.25 until the 90th percentile of the gross income distribution, and stays constant in the top decile (see Figure C.18).

Revenue-Neutral Reforms towards Individual Taxation: Political Feasibility. Figures 3.5 and 3.6 show how the population shares of reform winners and losers have changed over time. In these graphs, winners from a reform towards individual taxation are those couples, whose primary earnings are below the (green) line. In 1961, only a fifth of all married couples would have benefited from the reform. Couples with high secondary

[22]Note that even though elasticities are constant for primary and secondary earners, the average elasticity for couples can vary across the income distribution and across years since it is a weighted average based on the income shares of the primary and secondary earner (see Appendix Figure C.14).

Table 3.1. – Assumptions about Labor Supply Elasticities

	Primary Earner	Secondary Earner
Low Elasticity Scenario	0.15	0.35
Baseline Elasticity Scenario	0.25	0.75
High Elasticity Scenario	0.5	1.5

Note: This table displays our assumptions about the labor supply elasticities for primary and secondary earners in married couples. Assumptions are guided by the range of estimates found in the literature, e.g. Gustafsson (1992), Blundell and Macurdy (1999), Blau and Kahn (2007), Eissa and Hoynes (2004), LaLumia (2008), Kaygusuz (2010), Saez et al. (2012), Bargain et al. (2014), and Neisser (2021).

earnings were rare and hence a reform towards individual taxation would not have been politically feasible.[23] The behavioral responses of primary and secondary earners govern the slope of the green line. Larger elasticities of primary earners tilt the lines to the right and thus tend to decrease the number of reform winners. Under our baseline assumptions about behavioral responses to taxation, support has risen from approximately 23 percent in the 1960s to 55 percent today.[24] Even under the extreme assumption of an implausibly high elasticity of primary earnings to taxation, the reform is with 45 percent close to the majority threshold. Thus, while reforms towards individual taxation were not politically feasible in the past, they may become so if the trend continues.

Revenue-Neutral Reforms towards Individual Taxation: Welfare. By Theorem 4, a generic social welfare function approves a revenue-neutral reform towards individual taxation if $Y_1^g < r\ Y_2^g$. With the reverse inequality it is welfare-damaging. Figure 3.8 presents results for various social welfare functions under different assumptions about behavioral responses.[25] If welfare-evaluation dots locate above (resp. below) the respective green line, the reform is considered welfare decreasing (resp. welfare improving).

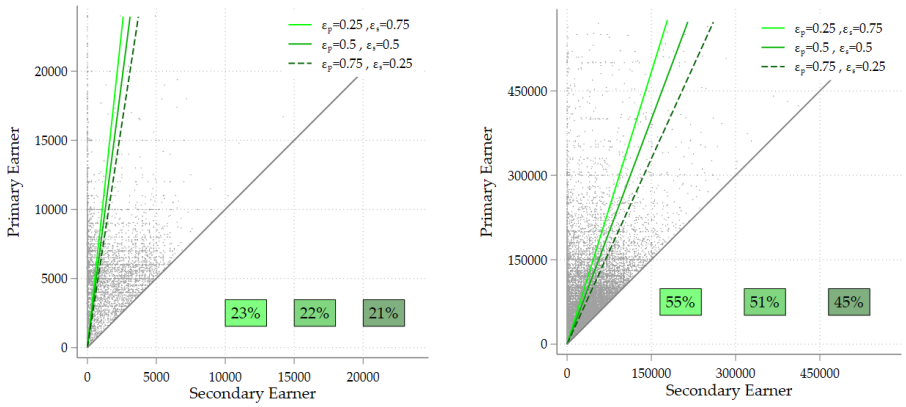
A striking feature is that a Rawlsian welfare function (with welfare weights concentrated on low income couples) and an Affirmative Feminist social welfare function (with weights that are increasing in the women's income share) are on different sides of the line that separates winners and losers. The reason is that among low-income couples the share of primary earnings tends to be high (see Figure 3.7). Therefore, only few low-income couples benefit from lower taxes on secondary earnings, and all are harmed by the higher taxes on primary earnings.

[23] In 1961, around sixty percent of couples had no secondary earnings at all. These couples lie exactly on the vertical axis of Figure 3.5 and represent a large fraction of reform losers.

[24] As of 2019, married couples represent 37 (54) percent of all tax units (individuals). Since singles are unaffected, around 30 percent of individuals would benefit from the reform, 24 percent would be made worse off, and 46 percent would remain unaffected.

[25] See Table C.12 in the Appendix for the formal specification of these welfare functions.

Figure 3.5. – Reform towards Individual Taxation: Political Economy



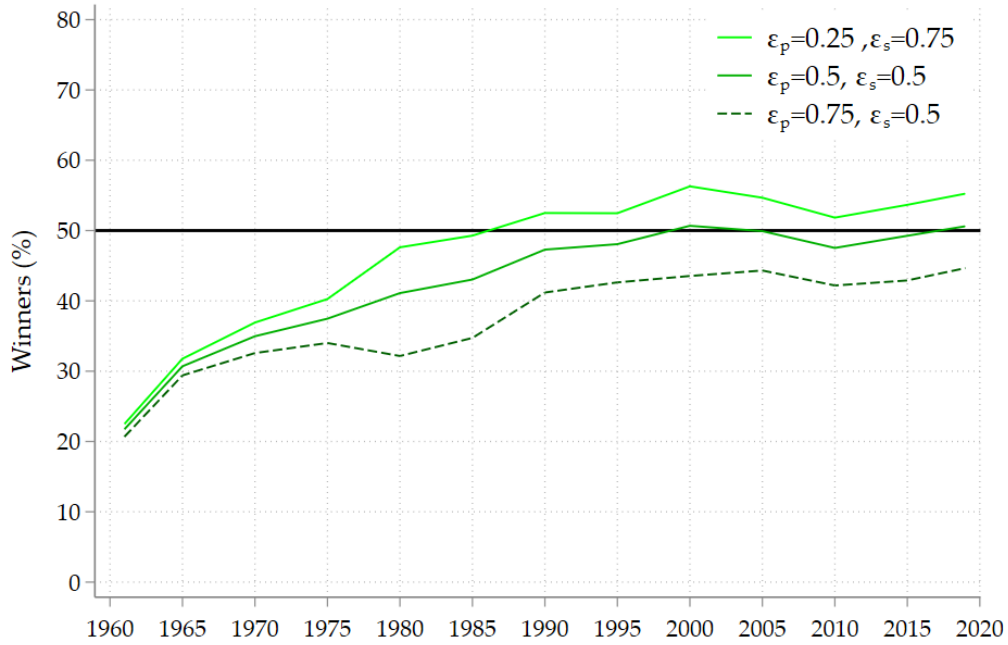
(A) 1961 (B) 2019

Note: This figure shows for 1961 and 2019, how the political support for a revenue neutral reform towards individual taxation among married couples varies with behavioral responses to taxation. Each gray dot represents a couple in the data, with the income of the primary (secondary) earner displayed on the vertical (horizontal) axis. Couples that lie below (above) the green line are winners (losers) from a reform towards individual taxation. The light green solid line refers to the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). All results are displayed including extensive margin responses. The figure also displays the respective share of couples than benefits from a reform towards individual taxation. Note that couples with no secondary earnings lie exactly on the vertical axis and constitute around 60 percent in 1961 and 25 percent in 2019. Figures for more years are displayed in Appendix Figure C.31. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Figure C.41 replicates this figure for the full adult population.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Discussion. The insight that reforms towards individual taxation may give rise to a conflict between the welfare of "the poor" and the welfare of "working women" is without precedence in the literature. It raises two questions. First, are there alternative reforms towards individual taxation that do not give rise to such a conflict? Second, are such conflicts empirically plausible? When we do an evaluation of actual – as opposed to hypothetical – tax reforms, do we also find such conflicts?

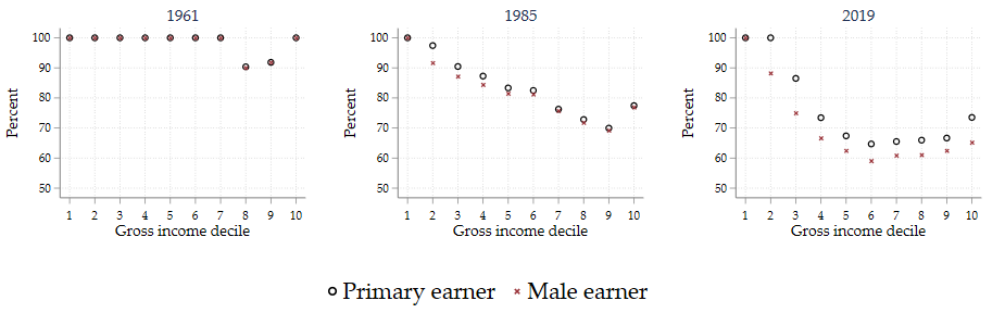
To answer the first question, we consider an alternative reform towards individual taxation. We suppose that marginal tax rates are lowered for all secondary earners, as before, but marginal tax rates are increased only for primary earners from the upper half of the income distribution. By construction, Rawlsian welfare will not decrease following such a reform. Poor couples with positive secondary earnings benefit, and poor couples without secondary earnings are not harmed. The reform does not collect as much revenue as one that taxes all primary earnings at a higher rate, with the implication that the tax rates on secondary earnings cannot be reduced as much. As shown in Appendix Figure C.34,

Figure 3.6. – Reform towards Individual Taxation: Share of Winners over Time



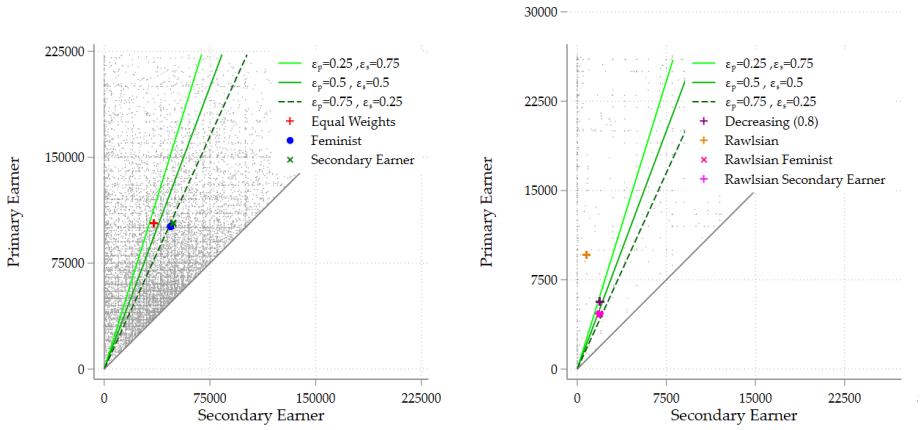
Note: This figure shows how the political support for a revenue neutral reform towards individual taxation among married couples evolved over time. All results are displayed including extensive margin responses. The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Figure C.42 replicates this figure for the full adult population.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure 3.7. – Median Share of Primary and Male Earner



Note: This figure shows the median income share of the primary earner in the couple by income decile. Earnings shares are computed on the basis of non-negative wage, business and farm income. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Figure C.43 replicates this figure for the full adult population.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure 3.8. – Reform towards Individual Taxation: Welfare (2019)



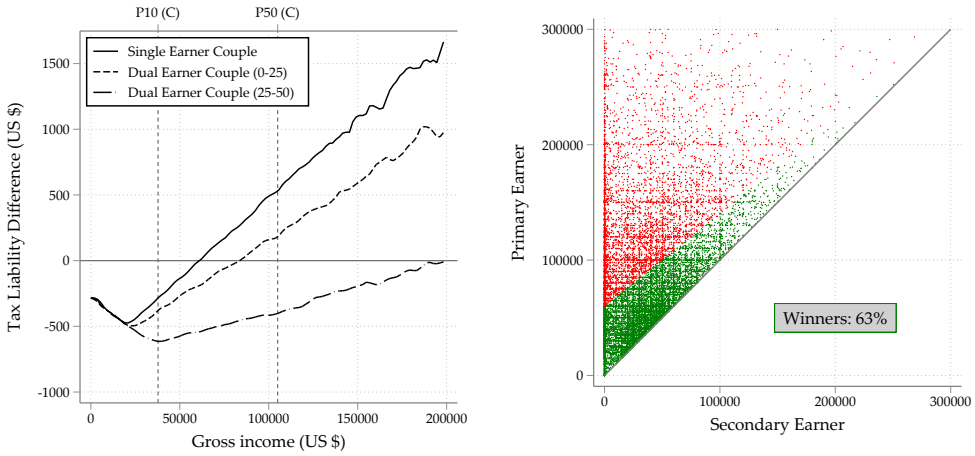
(A) Middle of Distribution (B) Bottom of Distribution

Note: This figure shows for the current tax system, how a reform towards individual taxation is evaluated from a welfare perspective under different exogenous welfare weights. Figure 3.8a (3.8b) displays welfare implications for welfare weights centered in the middle (bottom) of the income distribution. Each gray dot represents a couple in the data with specific income of the primary (secondary) earner displayed on the vertical (horizontal) axis. Couples that lie below (above) the green line are winners (losers) from a reform towards individual taxation. Welfare evaluations with different welfare weights are plotted as a colored dot the location of which is defined via the average welfare-weighted primary earnings (vertical axis) and the average welfare-weighted secondary earnings (horizontal axis). Welfare evaluations below (above) the green line indicate that the reform is welfare increasing (decreasing). The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (secondary) earner has an elasticity of 0.25 (0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). All results are displayed including extensive margin responses. For detailed information on welfare weight specification, see Table C.12. The specific percentile used for Rawlsian weights is P5 and $a = 0.8$ for decreasing welfare weights. Illustrations for other years are shown in Appendix Figures C.32 and C.33. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Figure C.44 replicates this figure for the full adult population.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

an affirmative feminist welfare function still goes up under such a reform. The reform is, moreover, politically feasible. There is one group that is harmed: couples from the upper part of the income distribution with low secondary earnings. The complementary group of reform beneficiaries accounts for more than 70 percent of the population.

The reforms that we discuss above are just two exemplary reforms out of a large set of conceivable reforms *towards individual taxation*. To answer the second question, we discuss how policymakers dealt with distributive conflicts in Sweden (1971) and the United Kingdom (1990). Specifically, we (i) analyze the shape of the respective reform, and (ii) evaluate whether a reform towards individual taxation of this shape would achieve majority support in the US as of today.

Figure 3.9. – Swedish Reform to Individual Taxation (1971) Applied to US (2019)



(A) Average Tax Liability Difference (B) Winners and Losers

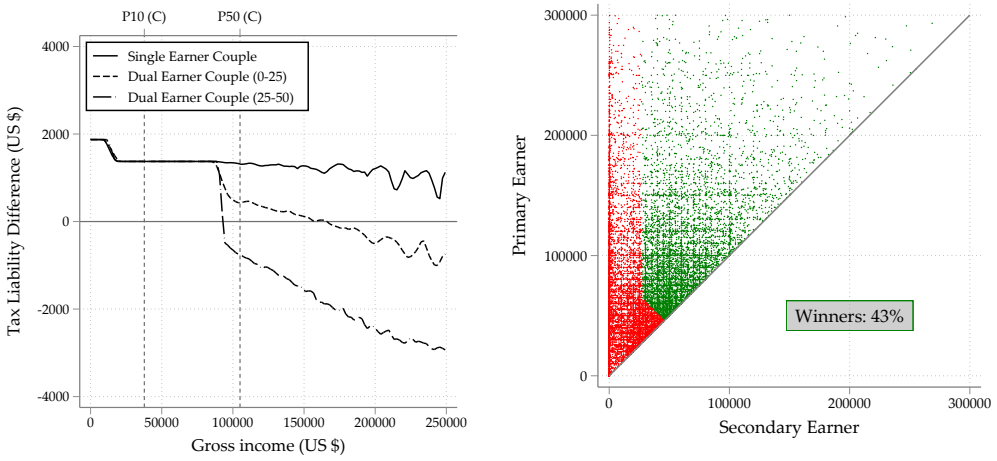
Note: The figure shows the average change in tax liabilities (Figure 3.9a) as well as winners and losers (Figure 3.9b) for a reform towards individual taxation that replicates the Swedish reform in the United States as of 2019. For details on the properties of the Swedish reform in 1971, and how we interpret it in our framework using primary and secondary earner revenue functions, see Appendix C.4.2. Revenue functions are estimated under both extensive and intensive margin behavioral responses and baseline values for intensive margin elasticities for primary (0.25) and secondary earners (0.75). Average tax liabilities across income levels are calculated using a local polynomial regression for three different couple types. Any revenue gain or loss is redistributed lump-sum across married couples. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM, CPS-ASEC, and Selin (2014).

The precise definitions of the reforms can be found in Appendix C.4.2. There, we also explain how we interpret these reforms through the lens of our framework, which requires to distill a reform direction h and reform intensity τ.

The reform in Sweden (1971) did not give rise to a conflict between the welfare of "the poor" and the welfare of "working women". On top of the move to individual taxation, the reform included a tax reduction at the bottom of the income distribution for single-earner couples. These reductions are paid primarily by rich single-earner households. There are also net contributors among rich dual-earner couples, but they are further up in the income distribution. Overall, the reform would generate large political support if done in the US today (63% of winners).

By contrast, the reform in the UK gave rise to a conflict between the welfare of "the poor" and the welfare of "working women". The reform included tax cuts which were larger for richer people and which led to a loss of overall tax revenue. Moreover, the tax cuts increased with secondary earnings. As a consequence, those with low-earnings were

Figure 3.10. – UK Reform to Individual Taxation (1990) Applied to US (2019)



(A) Average Tax Liability Difference (B) Winners and Losers

Note: The figure shows the average change in tax liabilities (Figure 3.10a) as well as winners and losers (Figure 3.10b) for a reform toward individual taxation that replicates the UK reform in the United States as of 2019. For details on the properties of the UK reform in 1990, and how we interpret it in our framework using primary and secondary earner revenue functions, see Appendix C.4.2. Revenue functions are estimated under both extensive and intensive margin behavioral responses and baseline values for intensive margin elasticities for primary (0.25) and secondary earners (0.75). Average tax liabilities across income levels are calculated using a local polynomial regression for three different couple types. Any revenue gain or loss is redistributed lump-sum across all married couples. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM, CPS-ASEC, and Stephens and Ward-Batts (2004).

made worse off. To benefit from the reform a couple had to be rich or have substantial secondary earnings. Overall, if such a reform were implemented in the U.S. today, it would generate more opposition than support (43% winners).

3.7 CONCLUDING REMARKS

This paper examines the joint taxation of married couples from a political economy perspective. We show empirically that, in the United States, marriage bonuses and penalties have remained largely unchanged since the 1960s when holding incomes constant. However, the number of individuals benefiting from joint taxation has steadily declined. We demonstrate that this trend has led to growing support for reforms towards individual taxation. As of today, such reforms are on the brink of achieving majority support. At the same time, we show that these reforms can have significant distributive consequences, potentially creating tensions between the welfare of "the poor" and the welfare of "working women." In the final part of the paper, we analyze how policymakers have addressed this potential conflict in practice. We evaluate the implementation of individual taxation

in Sweden and the United Kingdom. The Swedish reform mitigated the conflict by limiting losses for single-earner couples at the lower end of the income distribution, whereas the UK reform exacerbated it: to benefit from the UK reform, one had to be relatively wealthy or have high secondary earnings.

Chapter 4

Self-Interest, Stated Preferences, and the Taxation of Couples

This chapter is based on co-authored work with Felix Bierbrauer, Maximilian Blömer, and Andreas Peichl.

4.1 INTRODUCTION

This paper studies preferences over the tax treatment of couples in Germany. It uses two different approaches to measure preferences over tax policy and provides a systematic comparison. This is the paper's methodological contribution. The paper also relates to the recurrent debate about the tax treatment of married couples and families by providing evidence on "what people actually want." This is the substantive contribution.

The first approach is based on the assumption of utility-maximizing behavior. Individuals choose their earnings, trading off the benefits of disposable income and the costs of productive effort, also taking account of the consequences for intra-family allocation problems. When a tax reform allows an individual to realize higher utility, the individual is identified as a reform beneficiary. Otherwise, the individual is regarded as a reform loser. Thus, preferences over policies are identified with the material self-interest of market participants. This approach is used by a rich literature on the optimal taxation of couples (see, e.g., Ales and Sleet, 2022; Alves et al., 2024; Bierbrauer et al., 2024; Boskin and Sheshinski, 1983; Gayle and Shephard, 2019; Immervoll et al., 2011; Malkov, 2021). It is also used frequently in political economy analyses of redistributive taxation (see, e.g., Acemoglu et al., 2008; Bénabou and Ok, 2001; Bénabou, 2000; Bierbrauer et al., 2025; Bierbrauer and Boyer, 2015; Bierbrauer et al., 2021; Meltzer and Richard, 1981; Piketty, 1995; Roberts, 1977). The second approach is based on a direct elicitation of policy preferences. It uses survey experiments to analyze how people think about policies, how well-they are informed about them, what their worldviews and probabilistic beliefs are, what notions of fairness resonate with them, how they rank specific policy proposals and also how malleable these policy preferences are. This literature has been growing in recent years and explored various domains of public policy such as climate policy, migration policy, monetary policy, or redistributive taxation (see, e.g., Alesina et al., 2023; Binetti et al., 2024; Dechezlepretre et al., forthcoming; Stantcheva, 2021).

In this paper, we use both approaches to study preferences over the tax treatment of married couples in Germany. Germany, like France and the US, is a country in which married couples are taxed based on their joint income.[1] The consequence is that primary earners (mostly men) and secondary earners (mostly women) face the same marginal tax rate. This is at odds with a basic prescription of optimal tax theory: The empirical fact that secondary earners show stronger behavioral responses to taxation implies that they

[1]This stands in contrast to many other countries that transitioned from joint to individual taxation, such as Sweden in the 1970s and the United Kingdom in 1990. While most countries today officially implement individual taxation, many still incorporate elements of joint taxation in their tax systems, such as spousal allowances and tax credits (see, e.g. Bierbrauer et al., 2025).

should be taxed at a lower rate (Boskin and Sheshinski, 1983). Taxing them at the same rate is, moreover, detrimental for female labor market participation (Bick and Fuchs-Schündeln, 2017). This might be considered problematic both from a gender balance perspective and with a view on the scarcity of labor in an aging society. To what extent are these concerns reflected in the policy preferences that we measure? How large is the support for a move away from joint taxation? While these are specific questions that motivate our analysis, we shed light also on a broader theme: To what extent are the policy preferences elicited in survey experiments aligned with the policy preferences used in workhorse models of optimal tax theory and political economy. With a progressive income tax, single-earner couples are the ones who benefit the most from joint taxation. When we rely on material self-interest, we count them among the opponents of a reform towards individual taxation, and, if the population share of these couples is sufficiently large, we conclude that such a reform is politically infeasible. Empirically, however, this conclusion may be wrong. It is conceivable that, e.g., concerns for gender balance or views of fairness towards unmarried couples with children are so strong that they dominate material self-interest. Thus, it is possible that the support for a reform that abandons joint taxation is smaller or stronger than predicted by an analysis that focuses on material self-interest. Our analysis sheds light on which of these possibilities is empirically more plausible.

Material Self-Interest. Our measure of policy preferences based on material self-interest uses an approach developed in Bierbrauer et al. (2025). We use the most transparent version of this setup: Individuals in married couples choose earnings to maximize their joint surplus, defined as the difference between disposable income and their costs of productive effort. A tax reform affects a couple's tax burden and it provokes behavioral responses that affect overall tax revenue. We focus on revenue-neutral reforms to individual taxation. Such a reform lowers the marginal tax rate on secondary earnings and increases the marginal tax rate on primary earnings, with the changes in marginal tax rates calibrated to the behavioral responses to taxation in such a way that the reform is revenue-neutral. If the status quo has progressive and joint taxation, then such a reform is beneficial for couples with secondary earnings sufficiently close to primary earnings. Couples with a large discrepancy between primary and secondary earnings are worse off. Bierbrauer et al. (2025) develop sufficient statistics formulas that can be used to evaluate such reforms. As an input, these formulas require data on a status quo income distribution and a status quo tax system. They also invoke estimates of the elasticities that capture the behavioral responses to taxation. If all this is available, the approach can

be used to identify reform winners and losers from a reform. This information can also be used to check whether the reform would receive majority support and/or whether it would improve measures of social welfare.

We bring this approach to data from the German Socioeconomic Panel (GSOEP) and the ifo microsimulation model. Thereby, we can assess at the tax unit level, who would benefit and who would lose if a revenue-neutral reform towards individual taxation was carried out in Germany. We find, by and large, that the population share of reform beneficiaries equals the population share of reform losers. Put differently, if there was a referendum on a revenue-neutral reform towards individual taxation and if individuals would vote purely based on material self-interest, this would be a very close race with "modern couples" supporting the reform and "traditional couples" opposing it.

We also find that a revenue-neutral reform towards individual taxation would be rejected by a Rawlsian measure of social welfare. The reason is that single-earner couples are most prevalent in the bottom deciles of the income distribution. It would be supported, however, by a Feminist welfare measure, i.e., a welfare measure with weights that increase in the income share of the woman. Thus, given Germany's status quo, concerns beyond material self-interest could actually break the tie between the opponents and the proponents of a reform towards individual taxation. We therefore complement our analysis based on material self-interest with a survey experiment to uncover the extent to which German's actually share these concerns.

Stated Policy Preferences in a Survey Experiment. We cooperated with *forsa*, one of the most prolific survey institutes in Germany, and administered a survey of a representative sample of around 8,000 adult citizens of the German population. We elicited attitudes towards the tax treatment of couples in Germany. Specifically, we elicited both preferences towards specific policies and attitudes towards general principles of taxation. To ensure that respondents understand the effects of certain policies, we both explained the policies in general terms, but also integrated a factorial vignette study design in which respondents were presented with concrete trade-offs about the consequences of the status quo policy and of reforms of the system.[2]

Across the different elicitation methods, we find that material self-interest is an important driver of stated policy preferences, but that it does not fully explain them: For instance,

[2]Factorial survey experiments are a well established method in the social sciences to assess preferences and beliefs in a way that reflect real-world complexities and mitigates experimenter demand effects and social desirability bias (Auspurg et al., 2017; Jasso and Webster, 1999; Wiswall and Zafar, 2018). For an instructive overview, see Auspurg and Hinz (2014).

among single-earner couples and couples with a rather unequal within couple earnings distribution, between 60 and 70 percent oppose a move towards individual taxation. If only material self-interest mattered, the number would have to be 100 percent. Among dual earner couples, there is less opposition; 50 percent oppose the reform. If only material self-interest mattered, there would be zero support in this group.

One can explain the divergence from material self-interest by reference to different social welfare concerns that are also explored in the first approach. First, consistent with the notion that concerns for gender balance are more prominent among women, female spouses are more likely to approve a reform that introduces individual taxation, holding material self-interest constant. Also, respondents of political parties of the progressive spectrum, i.e, those parties that put a larger weight on gender-equality, are more likely to approve a reform of joint taxation.

Another explanation for the divergence from material self-interest are genuine competing principles of just taxation that relate to horizontal equity concerns and incentives of taxation. For instance, we find that one of the most popular and least controversial principles is that *couples with the same joint income should pay the same amount of taxes*. Also, the preferred jointness in marginal tax rates between spouses is not zero, indicating that respondents think that marginal tax rates of one spouse should, to some degree, be positively related to the earnings of the other spouse.

The bottom line is that stated preferences in the survey experiment reveal less support for a reform towards individual taxation than the analysis based on material self-interest. While there is no majority support for a strong reduction of marriage bonuses, higher benefits for families with children are very popular, both among respondents with and without children. This is also confirmed by the attitudes towards general principles of taxation, in which respondents across the political spectrum express their consent for a preferential tax treatment of families with children.

Related Literature. Our paper integrates two broad strands of the literature. First, we relate to the literature on the taxation of couples. This literature examines the design of tax systems in contexts with two-person households, typically composed of primary and secondary earners. A portion of this literature focuses on characterizing optimal tax systems from a welfare perspective (Golosov and Krasikov, 2025; Holter et al., 2023; Kleven et al., 2009), while other studies trace out the implications of couple taxation for labor market participation, consumption, and savings decisions of households (Bick and

Fuchs-Schündeln, 2017; Borella et al., 2023a; Guner et al., 2012).[3] Our paper introduces a political economy perspective to this literature. Specifically, we adapt the methodology of Bierbrauer et al. (2025) to the German context, assessing the political feasibility of reforms transitioning from joint to individual taxation. However, in contrast to Bierbrauer et al. (2025), we scrutinize whether the predicted degree of political support based on a framework of material self-interest is also realized when people are asked directly about their preferences for reforms that modify the tax treatment of couples.

This brings us to the second strand of the literature, which elicits citizens' attitudes towards specific policies and estimates the desired degree of redistribution, either using incentivized experiments (Andreoni and Miller, 2002; Fisman et al., 2007) or large-scale surveys (Stantcheva, 2021). Some work describes policy preferences in specific domains like pension design (Schuetz et al., 2023), climate policies (Dechezlepretre et al., forthcoming; Woerner et al., 2024), trade (Stantcheva, 2022), immigration (Alesina et al., 2023), monetary policy (Binetti et al., 2024), and tax policy (Ballard-Rosa et al., 2017; Engelmann et al., 2023; Fisman et al., 2020). Other studies analyze the distributional preferences of citizens with reference to more abstract notions of fairness (Almås et al., 2020; Cappelen et al., 2013; Cappelen et al., 2010).[4] Our contribution to this body of literature is twofold. To the best of our knowledge, our paper is the first to elicit policy preferences in the domain of couples taxation, as well as attitudes towards different normative concepts implicitly embedded in the design of the taxation of couples. Second, by contrasting the approach based on material self-interest with stated preferences, our approach holds relevance beyond the specific context of couple taxation. Specifically, we demonstrate how surveys can complement the standard toolbox of economists and identify where they yield results congruent with or divergent from neoclassical economic analyses.

Both the theory-led construction of policy preferences and the measurement of policy preferences in survey experiments can be regarded as second-best solutions. Policy preferences that are revealed by political choices would be first best. Measuring preferences over alternative designs of public policy in this way would require choice data from referenda or other forms of direct democracy. Such data rarely exists, however. Political behavior mostly takes the form of voting for a political party or its representatives.

To be clear, the construction of policy preferences on the assumption that the private choices of individuals (what to consume, how much to work etc.) are utility-maximizing

[3]For a recent detailed overview of this literature, see Bierbrauer et al. (2024), who provide novel insights on the characteristics of welfare- and Pareto-improving tax reforms in the domain of couple taxation.

[4]For overviews of the literature on the elicitation of fairness preferences using experimental and survey methods, see Gaertner and Schokkaert (2012), Almås et al. (2023), and Almås et al. (2024b).

invokes a revealed-preference assumption. When we see the earnings of couples in the data we interpret them as an optimal response to a labor-leisure trade-off. This data, however, does not reveal anything about policy trade-offs. From private choices, we cannot learn, say, whether an individual would be willing to accept higher taxes in return for more gender balance in the overall tax system. We only learn whether or not an individual's labor-leisure trade-off improves with a tax reform. This is what is captured by policy preferences that reflect material self-interest. Since private choices cannot reveal policy preferences, survey experiments are useful to obtain a more complete view of how individuals rank alternative policy options. A key contribution of this paper is to combine both approaches for a comprehensive analysis of preferences over the taxation of couples. This combination is unprecedented by previous analyses of policy preferences using survey experiments.

Outline. The remainder of the paper is organized as follows. Section 4.2 describes the system of joint taxation in Germany and provides information about the different data sources. Section 4.3 and Section 4.4, respectively, explain the methodology and results for our analyses of policy preferences based on material self-interest, and stated preferences, respectively. Section 4.5 concludes.

4.2 POLITICAL BACKGROUND AND DATA

4.2.1 Institutional Background

Status Quo. In Germany, married spouses have the option to file their tax return jointly (and almost all couples do) which implies that the tax liability only depends on the sum of their incomes. Whereas other countries like the United States and France also feature such a system of joint taxation, the German income tax explicitly links the tax schedules for singles and couples through a system of income splitting (*Ehegattensplitting*).[5] Income splitting implies that a couple is taxed as if it consisted of two singles who each contribute fifty percent to the couple's joint income, i.e.

$$T_m(y_1, y_2) = 2 \times T_s\left(\frac{y_1 + y_2}{2}\right)$$

[5]France features a family splitting system, in which the number of children enter the splitting factor. In the United States, singles and married couples face completely separate tax schedules. As shown in Bierbrauer et al. (2025), the link between separate tax schedules for singles and couples can be expressed via an income-dependent splitting function.

where T_m (T_s) is the tax function for married couples (singles) with primary (secondary) earnings y_1 (y_2). This has the following implications: First, who contributes how much to the joint income does not affect the tax liability. Single-earner couples and dual earner couples with the same joint income also face the same tax liability. Second, since the schedule T_s is progressive, spouses with unequal earnings obtain a marriage bonus, i.e., they pay less taxes than an unmarried couple with the same earnings. The marriage bonus increases, moreover, in the joint income, holding fixed the income shares of the primary and the secondary earner. It also increases in the income share of the primary earner, holding fixed the couple's total income. In this system, both spouses face the same marginal tax rate, i.e., the tax rate that applies to one additional unit of income is the same for the primary and the secondary earner. The consequence is that – relative to individual taxation – the primary (secondary) earner faces a lower (higher) marginal tax rate. This also implies that individual earnings incentives change with marriage.

The Discourse. We briefly summarize key arguments in favor and against joint taxation that are put forward in the public and in the academic discourse. Joint taxation can be justified with an appeal to horizontal equity: Couples with the same income should face the same tax burden (couples neutrality, see, e.g., Bittker, 1975). It can be criticized for its implications on earnings incentives: It is not responsive to the differences in the behavioral responses of primary and secondary earners. Social welfare and female labor market participation could be improved through a reform that lowers tax rates on secondary earnings (see, e.g., Bick and Fuchs-Schündeln, 2017). Thereby joint taxation also contributes to the persistence of traditional role models. The unequal treatment of married and unmarried couples can also raise fairness concerns. The diversity of living arrangements challenges the special status of marriage. In particular, married couples with and without children receive tax benefits which are not available to unmarried parents (see, e.g., Wissenschaftlicher Beirat beim Bundesministerium der Finanzen, 2018).

Reform Options. In our survey, we elicit preferences over various reform proposals that are "on the table" in Germany's public discourse. One proposal (referred to as *Realsplitting*) is to limit marriage bonuses in the following way: Spouses can choose an amount x so that the primary earner's tax liability is $T_s(y_1 - x)$ and the secondary earner's tax liability is $T_s(y_2 + x)$, where x must not exceed €13,805 (corresponding to the maximum amount of alimony allowance). Relative to the status quo, this limits the marriage bonus when the spouse's incomes differ by more than €27,610. Another proposal is to have marriage bonuses that increase in the number of children, referred to as family

splitting (*Familiensplitting*) – the system in place in France. Under this system a couple's tax liability would be given by

$$T_m(y_1, y_2, c) = (2 + 0.5c) \times T_s \left(\frac{y_1 + y_2}{2 + 0.5c} \right) ,$$

where c is the number of children. Individual taxation (*Individualbesteuerung*),

$$T_m(y_1, y_2) = T_s(y_1) + T_s(y_2) ,$$

is a third alternative. This would result in higher taxes for married couples with unequal incomes.[6]

4.2.2 Data

Our analysis of policy preferences based on material self-interest combines household data on earnings from the German Socioeconomic Panel (GSOEP v35, data for the year 2017, surveyed in 2018) with the ifo microsimulation model. To analyze stated policy preferences we developed a *forsa* survey module that elicited attitudes towards the tax treatment of couples in the year 2021. Tables D.1 and D.2 in the appendix display selected demographics for the two data sources.

German Socio-Economic Panel. The German Socio-economic panel (GSOEP) is an annual survey among 20,000 households in Germany covering a wide range of topics including household composition, and labor market participation. We use a version of GSOEP (v35) that has been prepared for the utilization of the ifo microsimulation model.[7] The model allows us to extract average and marginal tax rates for actual and hypothetical households. For the analysis, we focus on married couples of the working age population, i.e., married couples, in which both spouses are between 25 and 55 years old. All information, including both income data, as well as the income tax and transfer system regulations are based on the year 2017.

Forsa Survey Module. We elicit attitudes towards the tax treatment of couples through a special survey module administered by *forsa*, a professional survey company. Data col-

[6] While a full switch to individual taxation would (most likely) not be feasible for constitutional reasons in Germany, it is still an interesting benchmark. For details, see the report by the scientific advisory board of the Germany Federal Ministry of Finance (Wissenschaftlicher Beirat beim Bundesministerium der Finanzen, 2018).

[7] For details on the microsimulation model, see Blömer and Peichl (2020).

lection took place in the form of an online survey in the aftermath of the German federal election in 2021. The survey was conducted as part of the *forsa.omninet* panel. The panel is a stratified sample of the German population aged 18 and above. Our module was answered by around 8,000 respondents. To address potential selection concerns, all our estimates use survey weights targeted to German census data.[8] Apart from standard demographic characteristics, the survey also contains information on party preferences.

The module on the tax treatment of couples consists of three parts. In the first part, respondents are asked about their support of *specific reforms*. In the second part, respondents are asked about their attitudes towards *general principles of couple taxation*. In the third part, respondents are asked to determine tax liabilities when confronted with *household pairs* that vary along multiple dimensions. This part also includes information on the tax liability and on the marginal tax rates of primary and secondary earners under the status quo and under different reform options (real splitting, family splitting, individual taxation). These figures are based on the ifo microsimulation model and are calculated under the assumption that any change in tax revenue is redistributed lump-sum.[9] This simplifying assumption implies that all individuals are equally affected by changes in overall tax revenue, to the effect that changes in tax liabilities are the only source of heterogeneity among couples.

4.3 MATERIAL SELF-INTEREST

Our analysis of policy preferences that reflect material self-interest uses a framework developed in Bierbrauer et al. (2025) and Bierbrauer et al. (2023b). In its most general version, couples engage in Nash bargaining to determine who contributes how much to the joint income, and possibly also the allocation of household duties. The Nash bargaining outcomes depend on the tax function that translates the spouse's earnings into a joint disposable income. A tax reform alters this tax function. Thereby it changes earnings incentives. It also changes tax liabilities. Bierbrauer et al. (2025) characterize the implications that alternative reform directions have for overall tax revenue, individual welfare and various measures of social welfare by means of sufficient statistics. These sufficient statistics are most easily interpreted under a range of ancillary assumptions.[10]

[8] Although the *forsa.omninet* uses online data collection, participants are recruited via phone as a stratified sample based on quota sampling. For those panelists who have no internet access, TV screens are linked to forsa by a set-top-box that allows responding to the survey.

[9] For details on the estimated aggregate fiscal effects of different reform options and distributional implications based on the ifo microsimulation model, see Appendix D.1.4.

[10] Quasi-linear preferences, behavioral responses only at the intensive margin, iso-elastic costs of productive effort, the couple's disposable income is a public good for the spouses.

For expositional clarity, this is what we use here. Moreover, we focus on one specific class of reform directions, referred to as revenue-neutral reforms towards individual taxation. In the following we describe the sufficient statistics that (i) enter the construction of such a reform, (ii) are used to determine its political feasibility and (iii) are used to analyze its implications for social welfare.

Revenue-Neutral Reforms towards Individual Taxation. The construction relies on a simpler class of reforms. Consider a reform that raises marginal tax rates on primary earnings by τ whenever y_1 belongs to a range of incomes that starts at an income level y_{1s} and has length ℓ. Outside this range, marginal tax rates do not change. As formally shown in Bierbrauer et al. (2025), for τ and ℓ small such a reform's impact on tax revenue can be summarized by

$$\mathcal{R}_1(y_{1s}) = \underbrace{1 - F_1^y(y_{1s})}_{\text{Mechanical Effect}} \underbrace{- y_{1s}\, f_1^y(y_{1s})\; e_1\, \mathrm{E}\left[\frac{T_m'(y_m^0)}{1 - T_m'(y_m^0)} \mid y_1^0 = y_{1s}\right]}_{\text{Behavioral Effect}}. \tag{4.1}$$

Here, F_1^y is a cumulative distribution function that describes the (marginal) distribution of primary earnings in the status quo and f_1^y is the corresponding density. The primary earner's Frisch elasticity of labor supply is assumed to be constant across the income distribution and denoted by e_1. $T_m'(y_m^0)$ is the marginal tax rate that applies to a couple with joint earnings of y_m^0 in the status quo, and the expectations operator conditions on primary earnings being equal to y_{1s}. The mechanical effect captures that the reform increases the tax liability of couples with primary earnings larger than y_{1s}. The behavioral effect captures that couples who are confronted with a higher marginal tax rate reduce primary earnings.[11]

The function $y_1 \mapsto \mathcal{R}_1(y_1)$ describes how the revenue effect varies with the starting point of the bracket under consideration. Using the mathematics of functional derivatives, one can characterize the revenue implication of a reform that raises marginal tax rates slightly on all primary earnings as

$$\bar{\mathcal{R}}_1 \quad := \quad \int \mathcal{R}_1(y_1)\; dy_1\;.$$

The revenue implications of reforms that alter the marginal tax rates on secondary earnings can be characterized along the same lines. This yields $\bar{\mathcal{R}}_2 \;:=\; \int \mathcal{R}_2(y_2)\; dy_2$ for a reform that slightly changes the marginal tax rates on all secondary earnings. Using the mathematics of functional derivatives one more time, one can show that a reform that

[11] In the empirical application, we use an extended version of the formula that also takes into account extensive margin responses. For details, see Appendix D.1.2.

raises the marginal tax rate on all primary earnings by τ_1 and lowers the marginal tax rate on all secondary earnings by τ_2, for τ_1 and τ_2 sufficiently close to zero is revenue neutral if and only if

$$\frac{\tau_1}{\tau_2} \quad = \quad \frac{\bar{\mathcal{R}}_2}{\bar{\mathcal{R}}_1} \, .$$

Political Economy. Consider a revenue-neutral reform that raises the marginal tax rate on all primary earnings by τ_1 and lowers the marginal tax rate on all secondary earnings by τ_2. As formally shown in Bierbrauer et al. (2025), a couple with primary earnings of y_1^0 and secondary earnings of y_2^0 benefits from such a reform if $\tau_1 \, y_1^0 \, < \, \tau_2 \, y_2^0$; that is, if the additional taxes on primary earnings are dominated by the reduced taxes on secondary earnings so that the joint tax burden goes down. Equivalently,

$$y_1^0 \quad < \quad \frac{\bar{\mathcal{R}}_1}{\bar{\mathcal{R}}_2} \, y_2^0 \, . \tag{4.2}$$

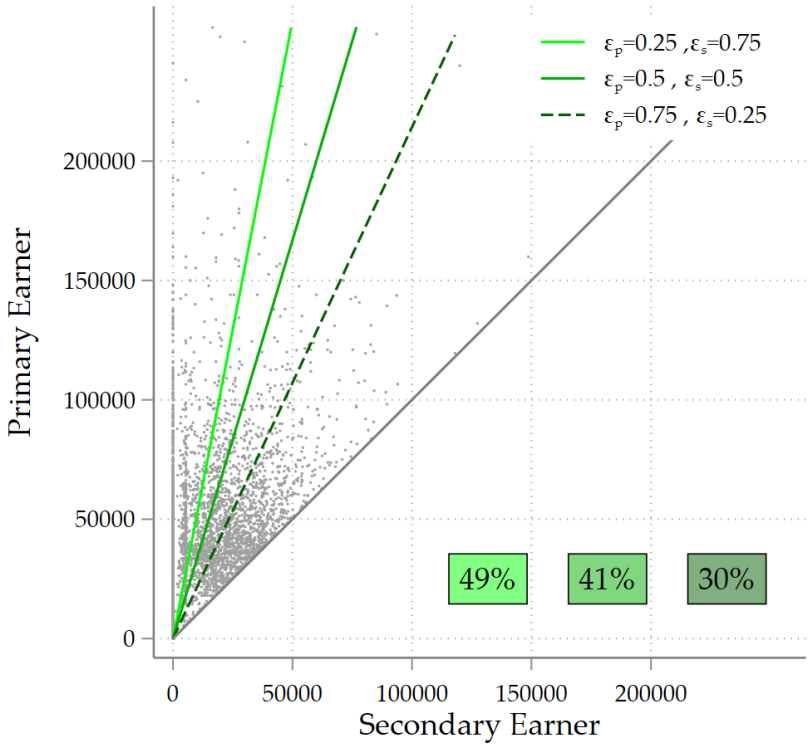
A revenue neutral-reform towards individual taxation is politically feasible if the population share of couples for which this inequality holds is larger than the population share of couples for which the reverse inequality holds.

All ingredients to equations (4.1) and (4.2) can be estimated using information on the earnings of married couples from GSOEP, the marginal tax rates under the status quo tax system from the ifo microsimulation model, and assumptions about the elasticities of primary and secondary earnings. For the latter part, we follow the literature and assume in the baseline that the behavioral responses to taxation are stronger for secondary earners, i.e. $e_1 = 0.25$ and $e_2 = 0.75$.[12] However, we also consider variations in which primary earners face the same or higher elasticities than secondary earners.

Figure 4.1 visualizes the results by plotting the joint distribution of primary and secondary earnings. The three green lines separate winners and losers (for different elasticity values): couples (gray dots) that lie below the green line are winners from a reform towards individual taxation. The support for the reform depends on what is assumed about the labor supply elasticities of primary and secondary earner's. The support varies between 30% and 49%. Hence, in all scenarios that we consider it is bounded from above by fifty percent. In the most plausible scenario (i.e. where $e_1 < e_2$), the support is close to fifty percent (49%).

[12]In particular, our assumptions are guided by the range of estimates found in the literature, e.g. Blundell and Macurdy (1999), Blau and Kahn (2007), Eissa and Hoynes (2004), LaLumia (2008), Saez et al. (2012), Bargain et al. (2014), and Neisser (2021).

Figure 4.1. – Reform towards Individual Taxation, Political Feasibility



Note: This figure shows for 2017, how the political support for a revenue-neutral reform towards individual taxation among married couples varies with behavioral responses to taxation. Each gray dot represents a couple in the data with specific income of the primary (secondary) earner displayed on the vertical (horizontal) axis. Couples that lie below (above) the green line are winners (losers) from a reform towards individual taxation. The light green solid line refers to the baseline elasticity scenario in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). All results are displayed including extensive margin responses. The figure also displays the respective share of couples than benefits from a reform towards individual taxation. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Gross income consists of wage and self-employment income.
Source: Own calculation based on GSOEP v35.

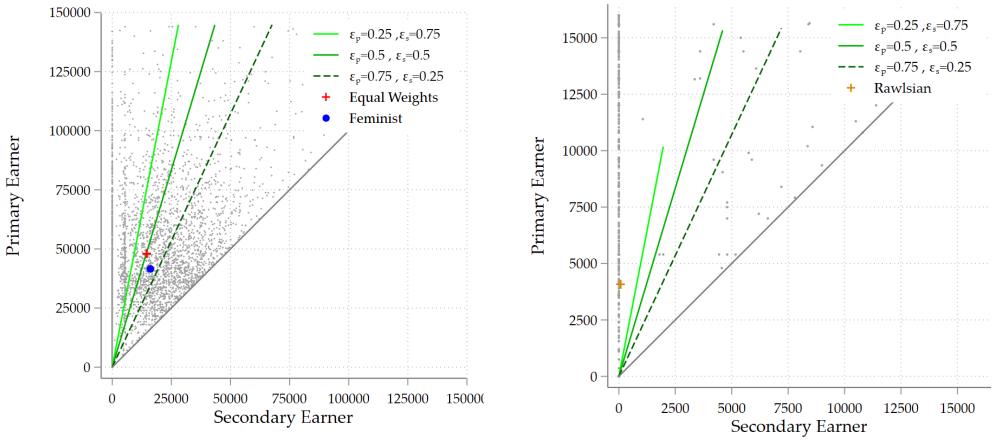
Implications for Social Welfare. We use additive social welfare functions to evaluate revenue-neutral reforms towards individual taxation. As shown more rigorously in Bierbrauer et al. (2025), a revenue-neutral reform towards individual taxation improves social welfare when a weighted average of the couples' tax burdens goes down. This condition holds when

$$\mathbf{E}\left[\mathbf{g}_m(\cdot)\, y_1^0\right] \quad < \quad \frac{\bar{\mathcal{R}}_1}{\bar{\mathcal{R}}_2}\, \mathbf{E}\left[\mathbf{g}_m(\cdot)\, y_2^0\right] \;, \tag{4.3}$$

where the welfare weights $\mathbf{g}_m(\cdot)$ depend on the characteristics of a couple. They can be specified in alternative ways. One possibility is to have the same weights for all couples,

with the implication that welfare is simply a measure of total economic surplus, i.e. total output minus total effort costs. This welfare measure indicates whether or not a reform is efficiency-enhancing. A Rawlsian welfare measure concentrates weights at the bottom of the income distribution. We use it to checks whether or not a reform is beneficial for "the poor". Finally, we refer to a welfare measure as "Feminist" when it has weights that increase in the woman's income share.

Figure 4.2. – Reform towards Individual Taxation, Welfare



(A) Middle of distribution (B) Bottom of distribution

Note: This figure shows for 2017, how a reform towards individual taxation is evaluated from a social welfare perspective under different specifications of welfare weights. Figure 4.2a (4.2b) displays welfare implications for welfare weights that are centered in the middle (bottom) of the income distribution. Welfare evaluations with different welfare weights are plotted as a colored dot. The location of a welfare dot is defined via the average welfare-weighted primary earnings (vertical axis) and the average welfare-weighted secondary earnings (horizontal axis). Welfare evaluations below (above) the green line indicate that the reform is welfare increasing (decreasing). The light green solid line illustrates the result under the baseline elasticity scenario in which the primary (secondary) earner has an elasticity of 0.25 (0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). All results are displayed including extensive margin responses. For detailed information on welfare weight specification, see Appendix Table D.3. The specific percentile used for Rawlsian weights is P10. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Gross income consists of wage and self-employment income.
Source: Own calculation based on GSOEP v35.

Figure 4.2 illustrates the results. Whether a revenue-neutral reform towards individual taxation is efficiency-enhancing (red cross) depends on whether the labor supply elasticity of primary earners is assumed to be smaller than the elasticity of secondary earners (which is what the empirical literature suggests). This is a reflection of the inverse elasticities logic in the given context. The reform clearly improves Feminist welfare as indicated by the blue dot and it leads to a loss of Rawlsian welfare (orange cross). The reason for the

latter result is that the share of single-earner couples is particularly high at the bottom of the income distribution (see Appendix Figure D.1). These couples are losers from a reform towards individual taxation, which is picked up by the Rawlsian welfare measure.

The policy preferences analyzed in this section are special in a variety of ways: First, they do not include distributive considerations, fairness concerns or views on the status of marriage. These are potential drivers of policy preferences beyond material self- interest. They are covered in the subsequent section on stated preferences in a survey experiment. Second, they are dependent on the specification of an economic model. As shown above, when elasticities change, political feasibility and welfare implications change too. Policy preferences measured in our survey experiment are not calibrated to the fine details of the economic environment.

4.4 STATED PREFERENCES

Our analysis of stated preferences consists of three parts. First, we assess the approval of four selected reform options that have been discussed in the run-up to the federal election in 2021 by political parties and in the media. Second, we assess the approval of more general principles of household taxation and show how they relate to the support of different reform options from the first part. Finally, we employ a factorial survey design, in which we present respondents with household pairs that vary randomly across multiple characteristics such as income, the number of persons in the household, and marital status. We then ask participants whether they would like to change the households' tax liabilities or their marginal tax rates. In each of the three parts, we will analyze, how well stated preferences fit to the predictions based on the approach of material self-interest.

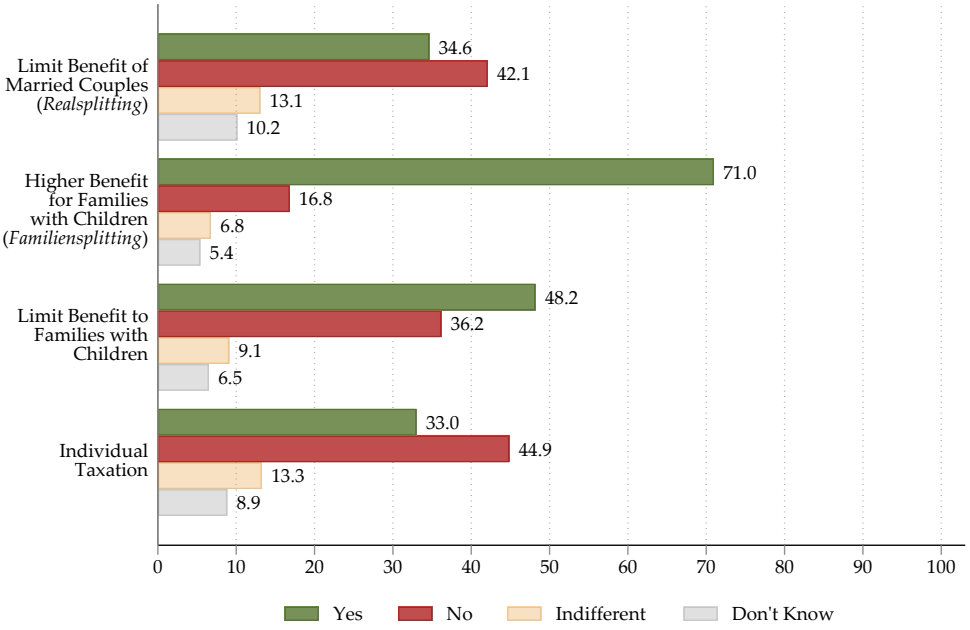
4.4.1 Specific Policy Reforms

In the first part of the survey module, respondents are informed about the nature of joint taxation in Germany.[13] Subsequently, they are asked about which of the reforms (real splitting, family splitting, individual taxation) they endorse.[14]

[13]In particular, they are informed that there is income splitting ("Ehegattensplitting") for married couples in Germany which implies that married couples usually pay less taxes than an unmarried couple with the same joint income. For details, see the original German questionnaire in Appendix D.2 and the corresponding translation in Appendix D.3.

[14]We gave respondents four different response options, i.e. "Yes", "No", "Don't care", "Don't know". Before answering, respondents could read brief explanations of the different reforms and the status quo (see Appendix D.3).

Figure 4.3. – Approval of Reforms



Note: This figure displays the approval rates of different reform options regarding reforms of the system of joint taxation in Germany. We use survey weights provided by *forsa*. The German terms for the respective policies are displayed in italics within parentheses.
Source: Own calculations based on *forsa* survey module.

Figure 4.3 shows the aggregate support for the different reform options. The least preferred reform is a complete switch towards individual taxation which is disapproved by almost 45 percent of the respondents. Limiting the benefit of married couples is also not popular, only 35 percent of respondents approve such a reform. In contrast, a reform that provides a more beneficial tax treatment for couples with children receives large support (71 percent). Limiting the benefit of joint taxation exclusively to couples with children is again much less popular (48 percent support).

These answers have to be interpreted with caution. Even though the reform descriptions contained information about a revenue-neutral rebate of changes in tax revenue, it immediately stands out that the most popular reform is one that involves tax cuts for families. The less popular reforms all imply that married couples lose tax benefits that they have in the status quo. The answers to these "simple" policy questions may therefore interpreted as showing that tax cuts are more popular than tax increases. To address this concern, the factorial survey presented in Section 4.4.3 will make trade-offs more transparent.

Support among Married Couples. A key insight from the framework based on material self-interest is that married couples with a large gap between primary and sec-

ondary earnings are harmed by a revenue-neutral reform towards individual taxation while spouses with a small gap benefit. Stated policy preferences show a similar pattern, since the opposition against (support for) a reform that introduces individual taxation declines (increases) with a more equal within-couple income distribution (see Figure 4.4a). Among single-earner couples and couples with an unequal distribution of earnings, close to 70 percent reject the reform that introduces individual taxation. At the same time, only 50 percent of dual earner couples with a very equal distribution reject the reform.

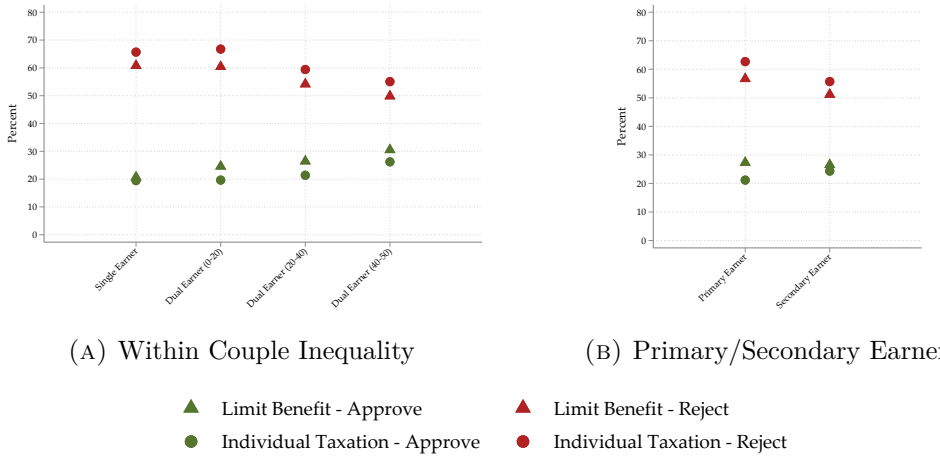
While the relative differences in support are consistent with the theoretical predictions, the level of support is substantially lower than the analysis based on self-interest would suggest. A possible explanation is that, in our theoretical analysis, increases in overall tax revenue are redistributed lump-sum. Thus, in our theory, couples benefit from a reform towards individual taxation even if their own tax burden increases somewhat. Again, one explanation for the stronger opposition among the participants of the survey experiment is that they do not take these revenue implications into account, possibly because they are not as salient to them as the changes in the distribution of tax liabilities.

Another insight from the framework based on material self-interest is that if spouses engage in Nash bargaining, their preferences over various reform options are aligned (Bierbrauer et al., 2025). Consequently, we would not expect any differences between primary and secondary earners. This is confirmed by Figure 4.4b, showing that there are only marginal differences between the two types of earners in the approval of different reforms that abolish joint taxation.

The insights of Figure 4.4 are purely descriptive. It might well be that the observed preference pattern is not a confirmation of the importance of material self-interest, but rather driven by demographic or party-alignment differences that are correlated with the within-couple income distribution. To address this concern, we estimate the effect of the within-couple income distribution and the effect of the earner type of the respondent in a multivariate regression framework. In particular, we employ an ordered logistic regression framework, in which we regress the ordered degree of support for different reforms (Reject reform - Indifferent - Approve reform) on a broad range of observable characteristics at the level of the couple (income level, within couple income distribution, presence of children) and the level of the respondent (earner type, sex, age, education, political party preference). This allows us to assess the effect of specific characteristics on political support, holding other attributes constant.

Figure 4.5 displays the results and shows that the observations from the descriptive analyses are confirmed by the multivariate analysis. In particular, the figure displays the effect

Figure 4.4. – Approval of Reforms by Married Couples



(A) Within Couple Inequality (B) Primary/Secondary Earner

Note: This figure displays the approval (disapproval) rates of the reform towards for individual taxation for all married couples across different within household income distributions (Figure 4.4a) and for primary and secondary earner respondents (Figure 4.4b) as green (red) markers. Circles show results for the switch to individual taxation and triangles show the reform to real splitting. Respondents who are indifferent or do not answer the question are excluded from the analysis. We use survey weights provided by *forsa*. *Source:* Own calculation based on *forsa* survey module.

of various characteristics in terms of odds ratios. Higher odds ratios for a categorical variable defining a specific group imply that the respective group is more likely to approve the reform compared to the baseline group. For instance, compared to single-earner couples, the odds of endorsing a reform of joint taxation are around two times as large for dual-earner couples with very equal incomes, holding all other factors constant. In contrast, we do not find any statistically significant differences in the approval odds for primary and secondary earners. This suggests that for questions about concrete reforms, respondents self-interest from a couple perspective seems to play a prominent role.

In addition, and consistent with self-interest, the support for limiting the benefit of married couples comes with an income gradient while a complete elimination of joint taxation does not. A complete elimination of joint taxation would affect couples at all income levels equally. In contrast, limiting the benefit for married couples would primarily hit rich couples. As shown in Figure 4.5a, respondents in such couples are less likely to approve a reform.

Despite the consistency with predictions based on material self interest, the analysis also points to some factors that play a role for stated support while they are irrelevant from the perspective of material self-interest. For instance, women are more likely to approve a reform that introduces individual taxation. The data does not allow for a direct interpretation of this effect, but it is in line with the idea that women in couples

have a stronger social preference for gender balance in the work place and thus tend to act in line with the social preferences based on a Feminist social welfare function (see Section 4.3).

Support in the Population at Large. Figure 4.6 is the analog of 4.5 that is based on the responses of all survey participants, including, in particular, unmarried respondents. It also includes in addition the analysis of reforms that focus on families with children. Attitudes towards *Individual Taxation* vary by household type in a way that is in line with self-interest: Respondents are more willing to give up the benefits of others than to give up their own benefits, e.g., singles are more likely to support a reduction of marriage bonuses than married individuals. Their odds of approving a switch to individual taxation are almost four times higher. Interestingly, the largest approval rates for reforms towards individual taxation are found among divorced and separated respondents.[15] Unmarried couples take an intermediate position between married couples and divorced respondents. This suggests that expected benefits from joint taxation in the future might also play a role in the reform approval decision. However, the effect of the contemporaneous material benefit seems to outweigh these potential gains in the future.[16]

There is less variation in the approval towards other reforms coming with *Higher benefits for families with children* or reforms that *Limit the Benefit to Families with Children*. For these reform options, there is again positive correlation with self interest, the approval odds of respondents with children are between three and four times higher compared to respondents without children (see Figures 4.6b and 4.6c).

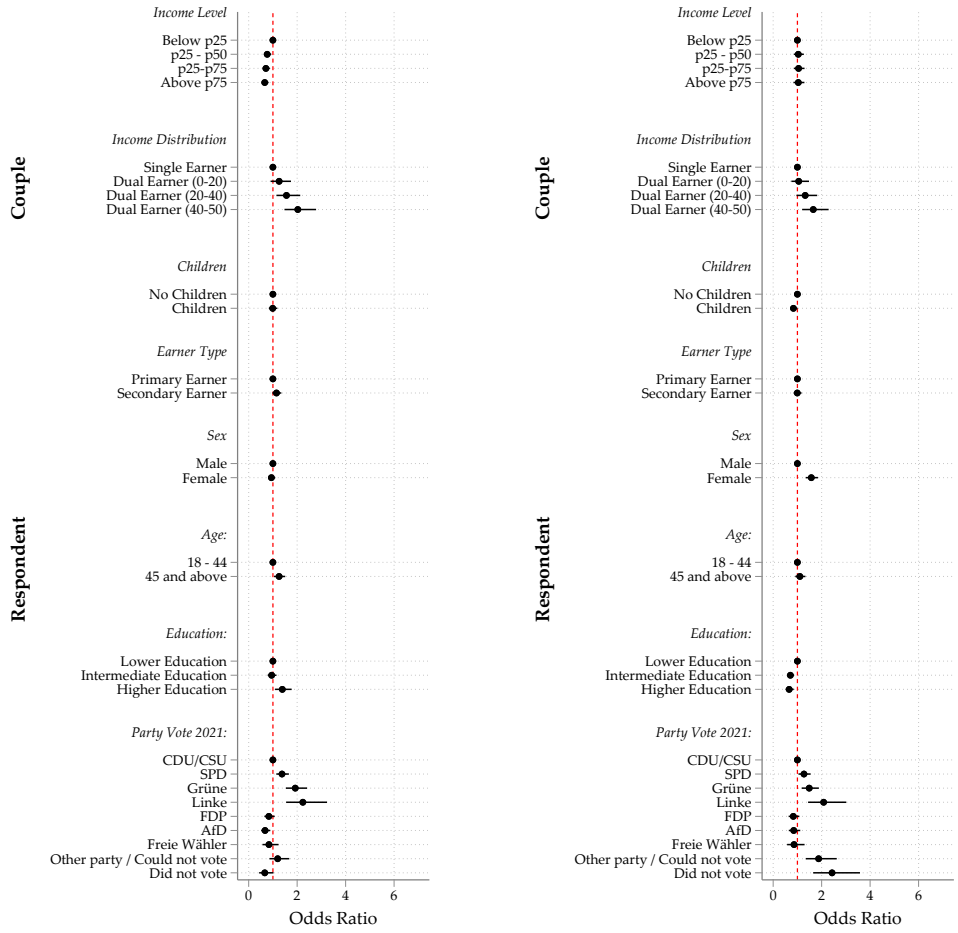
Stated preferences over reform options vary in a systematic way with voting behavior in the German federal election of 2021 (according to the self-reported vote). The variations are broadly in line with the proposals of the respective parties in the election campaign.[17] In particular, voters of the more left-leaning parties that tend to emphasize gender equality more (Linke, SPD, Grüne) are more likely to express support for reforms that reduce

[15] In the group of divorced respondents, the support for individual taxation is 55 percent and thus 33 percentage points higher than among married couples (see Appendix Figure D.2). This suggests that respondents evaluate consequences of income splitting differently before and after the end of a marriage and give potential adverse effects a larger weight.

[16] Also note that a reform of joint taxation might come with transitional or grandfathering regulations for those couples that are already married. This introduces an additional level of horizontal equity concerns that unmarried couples are be particularly concerned about. In Appendix Figure D.7, we show that grandfathering regulations are very popular among married couples, since they protect their current benefits from the tax treatment of couples while unmarried couples disagree to a large extent such transitional rules.

[17] We illustrate party positions on the tax treatment of married couples with reference to quotes from their manifestos in Appendix Tables D.4 and D.5.

Figure 4.5. – Approval of Reforms by Married Couples, Heterogeneity



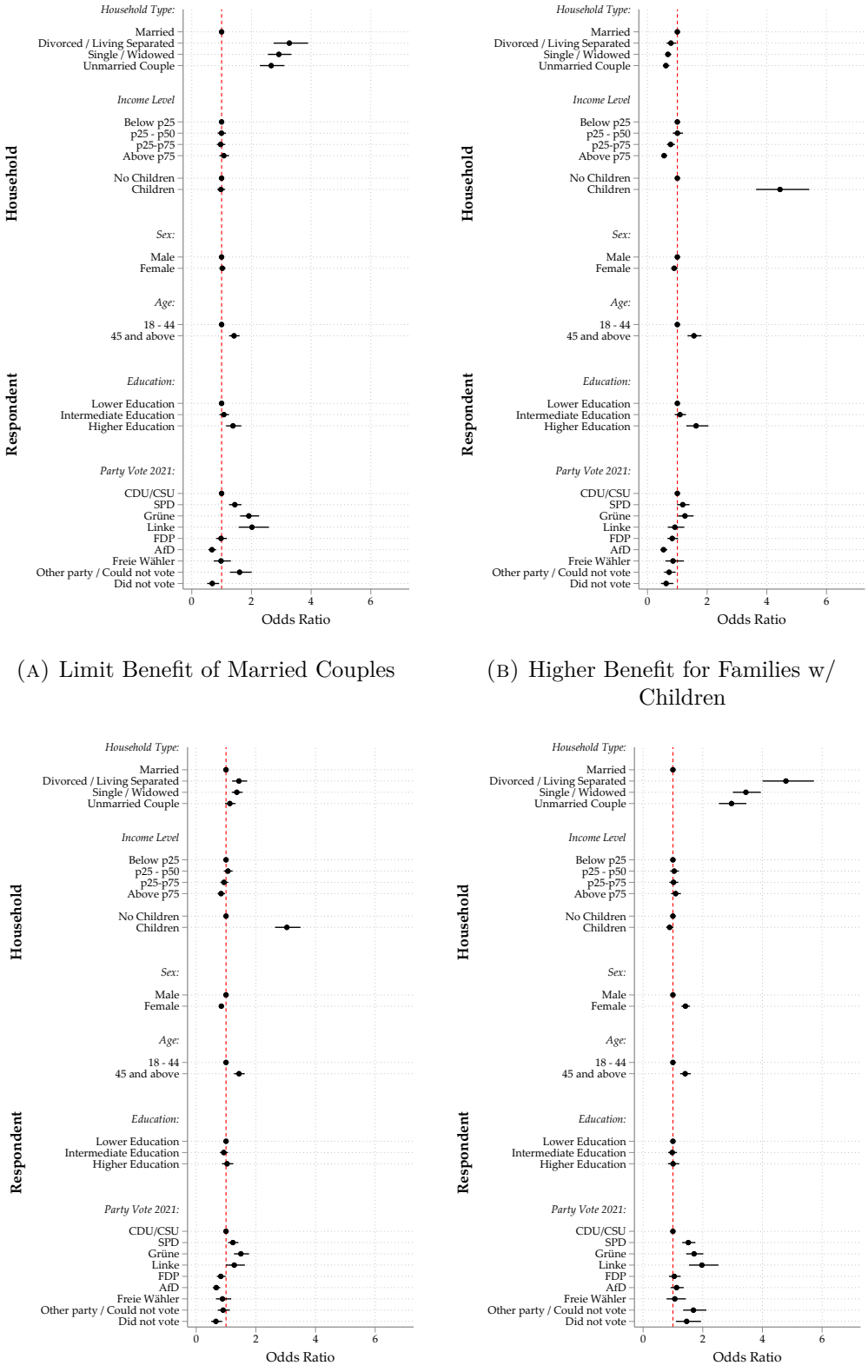
(A) Limit Benefit of Married Couples (B) Individual Taxation

Note: This figure displays odds ratios (transformed coefficients) and related 95 percent confidence intervals from an ordered logit regression for the four different reform options focusing on married couples. We exclude "Don't know" responses in both the dependent and independent variables. For illustrative purposes, the baseline category for each categorical variable is also shown as the first respective category. We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

marriage bonuses than the voters of other parties (CDU, FDP).[18] However, it is important to note that this regression analysis shows differential odds, but not absolute levels of support. Even among the voters of the more left-leaning parties, majority support for a reduction of marriage bonuses is lacking. (see Appendix Figure D.8). Our analysis in

[18]This classification is based on a "gender equality check" of the party election proposals for the German Federal election 2021 published by the National Council of German Women's Organizations (2021). The AfD was not analyzed in this document.

Figure 4.6. – Approval of Reforms, Heterogeneity



(A) Limit Benefit of Married Couples

(B) Higher Benefit for Families w/ Children

(C) Limit Benefit to Families w/ Children

(D) Individual Taxation

Note: This figure displays odds ratios (transformed coefficients) and related 95 percent confidence intervals from an ordered logit regression for the four different reform options. We exclude "Don't know" responses in both the dependent and independent variables. For illustrative purposes, the baseline category for each categorical variable is also shown as the first respective category. We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

Section 4.3 provides a possible explanation. Such a reform would harm many couples in the bottom deciles of the income distribution. Plausibly, the voters of left-leaning parties are particularly sensitive to this effect.

4.4.2 Principles of Couples Taxation

The previous section showed that self-interest is an important explanation for political support of reform options. In addition, value judgments are a potential drivers of policy preferences. To see which value judgments are relevant when it comes to the tax treatment of couples, we asked individuals whether they support principled statements on a five-point scale ranging from "Full Disagreement" to "Full Agreement". In total, we provided 11 statements that put emphasis on different aspects of the tax treatment of couples ranging from horizontal equity ("A married couple should pay the same amount of taxes as an unmarried couple when both couples have the same total income.") to incentive concerns ("Marriage should not change the incentives for income generation for one of the partners.").[19]

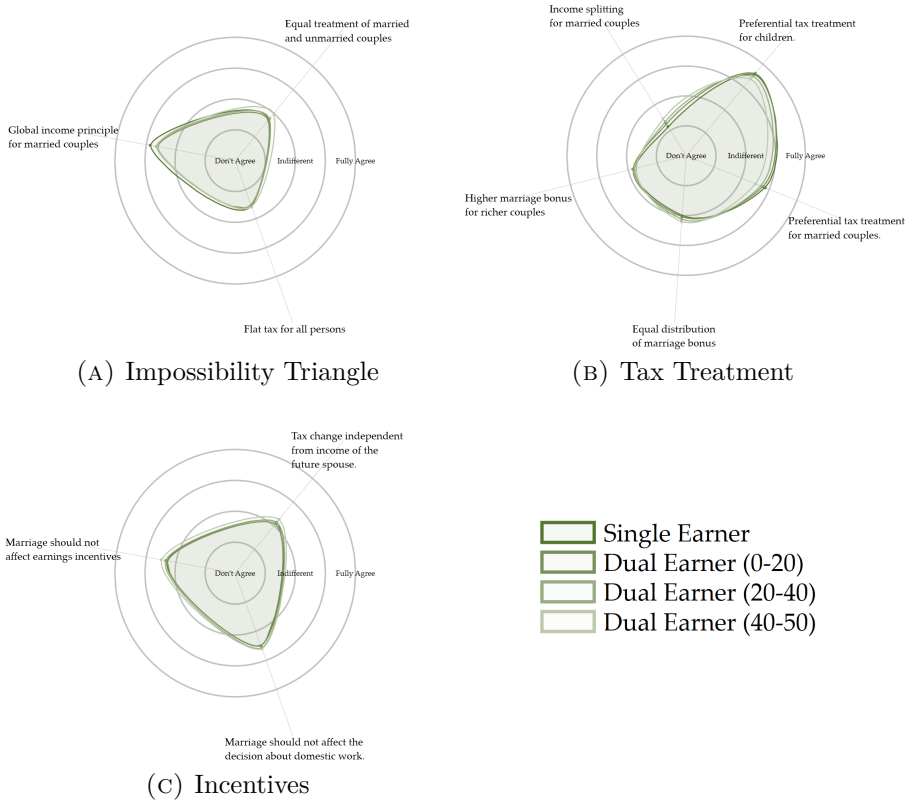
We organize these 11 principles into three different categories and visualize the average responses of respondents in Figure 4.7, focusing on different types of married couples.

The first three principles are mutually inconsistent and we refer to them as forming an "impossibility triangle": (i) an equal treatment of married and unmarried couples (marriage neutrality), (ii) a progressive system with increasing marginal tax rates (progressivity), and (iii) equal tax liabilities for couples with the same joint income (couples neutrality).[20] As shown in Figure 4.7, *independent of how they are affected by the tax treatment of couples*, respondents in the survey tend to agree more on (iii) than on (ii) or (i). The fact that there are only minor differences in the evaluations between different types of married couples implies that normative evaluations of horizontal equity concerns are rather independent from the implications of material self-interest. In addition, while some of these principles, e.g., the global income principle, are approved uniformly across the political spectrum , the horizontal equity between married and unmarried spouses is more important for supporters of center-left political parties (see Figure 4.8).

[19]For details on the questionnaire and exact statements, see Appendix D.2.

[20]In the lax law literature, this is referred to as the so-called "marriage tax trilemma" (see Bittker, 1975, for a first discussion of this trilemma): It is mathematically impossible to design a system that neither encourages nor penalizes marriage (marriage neutrality), taxes higher income individuals at higher rates (progressivity), and imposes the same tax liability across all married couples with the same income (couples neutrality). An income tax system can incorporate any two of these features, but it cannot have all three.

Figure 4.7. – General Principles, by Married Couple Type



(A) Impossibility Triangle

(B) Tax Treatment

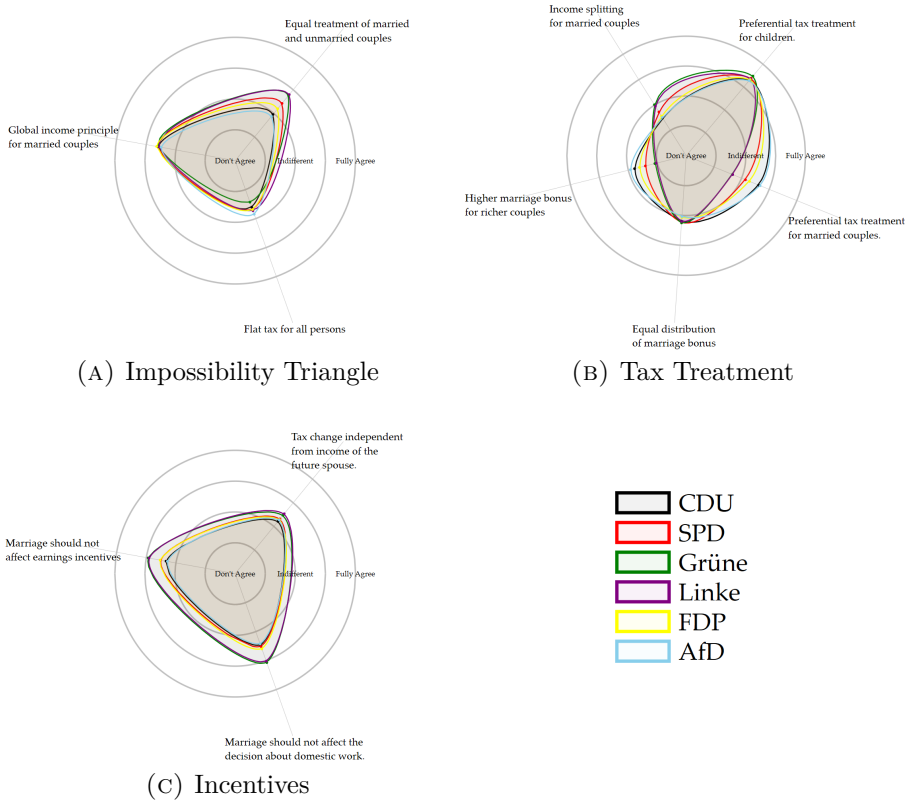
(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation for all married couples by different within household income distributions. The first three principles cover the impossibility triangle principles (see Figure 4.7a). Figure 4.7b displays results for detailed principles of the tax treatment of couples while the last three principles capture principles specified in terms of different incentive margins (see Figure 4.7c). Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question. *Source:* Own calculations based on *forsa* survey module.

The second set of principles refers to the tax liabilities of different household types. In line with the results on the popularity of higher benefits for families with children, respondents agree strongly with the principle that there should be a preferential tax treatment for families with children. In line with self-interest, married couples on average tend to agree with the preferential tax treatment for married couples, . However, there is strong disagreement on this principle across party lines. While supporters of the radical left party as well as the green party rather tend to reject the preferential tax treatment for married couples, supporters of all other political parties are rather in favor of this principle.

Finally, the list of principles also included statements about how the taxation of couples should (not) shape incentives along different margins. In general, respondents tend to

Figure 4.8. – General Principles, by Political Party Preference



(A) Impossibility Triangle

(B) Tax Treatment

(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation by political party preference measured as the stated party vote in the German federal election in 2021. We focus on the six major parties represented in the German parliament. The first three principles cover the impossibility triangle principles (see Figure 4.8a). Figure 4.8b displays results for detailed principles of the tax treatment of couples while the last three principles capture principles specified in terms of different incentive margins (see Figure 4.8c). Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question. We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

agree that marriage should not affect earnings incentives, decisions about domestic work and the decision of whom to marry. Again, supporters of the left and the green party agree more strongly to these principles.

In the appendix, we provide evidence on how the support regarding these principles varies across different household types (Appendix Figure D.9), and is affected by whether the respondent is a primary/secondary earner (Appendix Figure D.10), or a male/female respondent (Appendix Figure D.11). Broadly, the patterns that we find are reflective of self-interest. For instance, single households and unmarried couples are more likely to favor an equal treatment of married and unmarried couples, while married couples show strong agreement for a preferential tax treatment of married couples. However, some principles show unanimous support like the preferential tax treatment of children

and the global income principle. In line with the results from Section 4.4.1, the answers of primary and secondary earners are, by and large, similar. However, there are some gender differences. For instance, women show a stronger concern for horizontal equity between married and unmarried couples whereas men show on average more support for a preferential tax treatment of married couples. We also find that the support to general principles is broadly in line with the support of specific policies in the previous analysis (see Appendix Figures D.12-D.15).

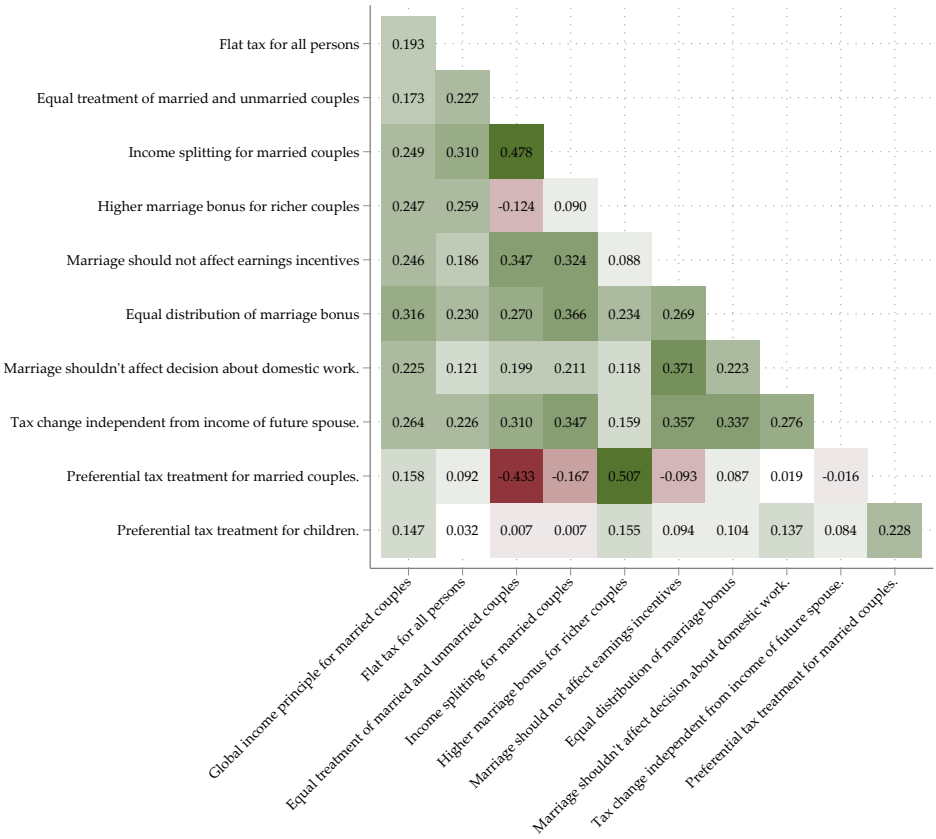
Inconsistencies in the Support of Principles. The average support of different principles of couples taxation reveals that respondents might not necessarily be aware of the trade-offs and the potential impossibility of reconciling multiple principles at once. This is illustrated by Figure 4.9 that displays the correlation coefficient between the different principles. While respondents show consistency for some pairs ("Preferential treatment for married couples" displays a negative correlation with "Equal treatment for married and unmarried couples"), other correlations reveal inconsistencies. For instance, respondents that are in favor of a preferential tax treatment of married couples are at the same time more likely to think that marriage should not lead to changes in earnings incentives. The subsequent section addresses this concern and elicits respondents' preferences through a factorial survey design that makes trade-offs explicit by informing respondents about tax liabilities of different household types in the status quo and that, moreover, visualizes the impact of different reform options.

4.4.3 Using Household Pairs to Illustrate Reform Options

While the analysis of stated preferences in Section 4.4.1 and Section 4.4.2 did not make the trade-offs between different reforms explicit, we now turn to the analysis of a factorial design that forces respondents to take a stance on these trade-offs. In the factorial design, we present each respondent with pairs of households, i.e., the vignettes, that vary randomly across different characteristics. When a respondent is confronted with a particular pair of households, she is given the chance to change the households' tax liabilities and marginal tax rates for both the primary and secondary earner. We also illustrate the implications of different reform options from Section 4.4.1 (individual taxation, family splitting, real splitting) for the specific households and ask them again about their support for the respective reform.[21]

[21]One might be worried that choices in hypothetical vignette experiments are just cheap talk, since they don't have real world consequences. However, recent work involving very similar vignette design shows that preferences elicited in hypothetical situations deliver stable results that are in line with an incentivized experiment (Hufe and Weishaar, 2025).

Figure 4.9. – General Principles, Correlations



Note: This figure displays the pairwise correlation coefficients for different principles of couple taxation. Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question. We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

Construction of Household Pairs. The household pairs are constructed based on a random combination of four different characteristics, i.e., the household type, the number of children, gross income, as well as the income share of primary earner. Each characteristic can take multiple expressions that are shown in Table 4.1. The number of expressions yields a set of 84 different household types.[22] An exemplary household pair is visualized in Figure 4.10.

Question Sequence. Each respondent evaluates three different household pairs. For every household pair, we proceed in three steps.

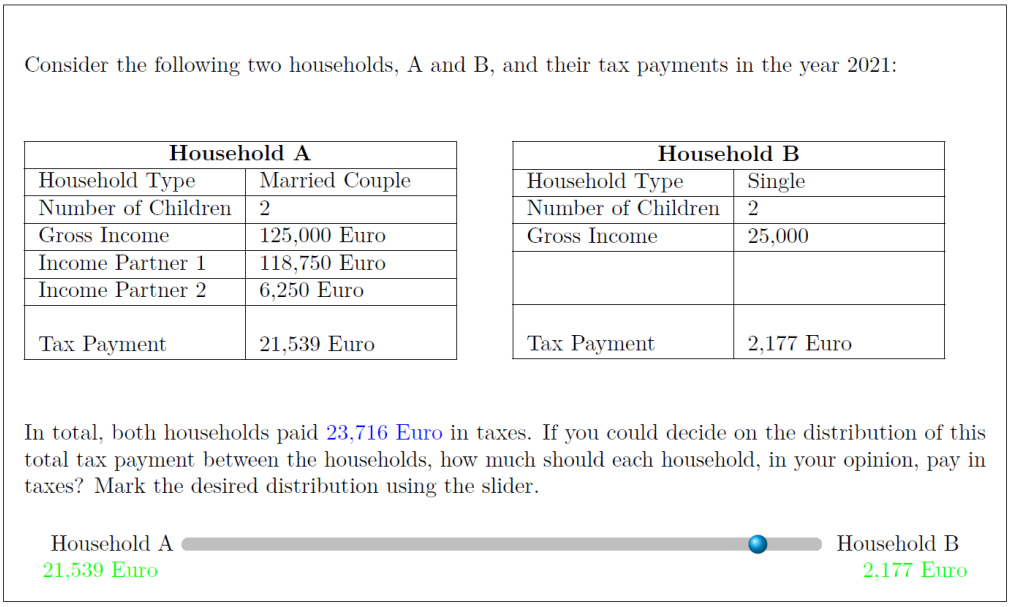
[22]In the selection of the number of characteristics and respective expressions, researchers face a trade-off between precision and complexity. A greater number of vignette characteristics allows for better control of potential confounding factors. However, it also increases the cognitive load on participants when making distributional choices (Auspurg and Hinz, 2014).

Table 4.1. – Characteristics and Levels in Vignettes

Characteristics	Number of Expressions	Expressions
Household Type	3	Single / Unmarried Couple / Married Couple
Number of Children	2	0 children / 2 children
Gross Income	6	25,000 / 50,000 / 75,000 / 100,000 / 125,000 / 150000
Income Share Primary Earner	3	95 / 75 / 55

Notes: This table displays the variables and the (number of) expressions used in the vignette study. The income share of the primary earner is only included for couples. For details on the questionnaire, see Appendix D.2.

Figure 4.10. – Example of Question



Notes: This figure illustrates an exemplary household pair presented to respondents. The characteristics vary randomly according to the expressions defined in Table 4.1. The figure illustrates a translated version of the question. The original statements in German can be found in Appendix D.2.

1. We show respondents the characteristics of the household and the **tax liability** under the status quo.[23] We then ask respondents to state how they would split the tax liability if they could decide on the distribution of the total tax burden.

2. We show respondents the characteristics of the household and the **marginal tax rates** of the primary and secondary earner under the status quo. We then ask respondents to state how they think additional income of the primary and secondary earner should be taxed.

[23]The status quo tax liabilities have been calculated using the ifo microsimulation model according to the law as of 2021.

3. We show respondents the characteristics of the household and display **how reforms affect tax liability and marginal tax rates**. To make the reforms revenue neutral, we assume that any gain or loss in tax revenue is rebated lump-sum at the individual per capita level. We then ask respondents on whether they would support the corresponding reform.

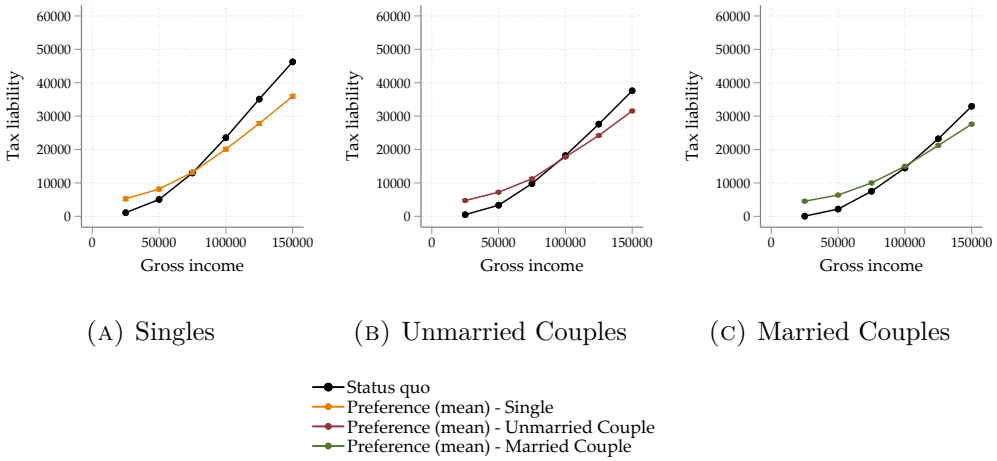
Results: 1) Tax Liability. To analyze the preference regarding the tax liability of different household types, we contrast the status quo tax liability with the average tax liability assigned by respondents.

For this purpose, we transform the household pair data set to a stacked household data set and estimate a fully saturated regression model, i.e., a regression of the status quo (preferred) tax liability on the full set of interactions of household characteristics (see Table 4.1). We then estimate the marginal effects of specific variables based on the predictions from this model, thereby averaging over other variables in the model. We present results for (i) the preferred tax progressivity, (ii) the preferred tax treatment of married couples, and (iii) the preferred tax treatment of households with children.

Regarding tax progressivity, Figure 4.11 shows that respondents state a preference for a *less* progressive tax system, since for both singles, unmarried couples and married couples, the average preferred tax liability is larger (smaller) at the bottom (top) of the income distribution.

Taking trade-offs into account, respondents are also in favor to reduce the of the status quo benefit of married couples, since the preferred tax liability difference between married couples and unmarried couples is smaller than under the status quo (see Figure 4.12). Yet, respondents indicate that married couples should still pay less taxes than unmarried couples with the same income, that is, the preferred marriage bonus is still positive but smaller than under the status quo. The difference between the preferred and the actual marriage bonuses is exclusively driven by the large marriage bonuses that are realized by "traditional couples", i.e., couples with an unequal within-couple income distribution, and especially for those in the upper parts of the income distribution (see Figure 4.12d). This is in line with our results on the support for various principles of taxation in Section 4.4.2: While there is an agreement with a preferential tax treatment for couples, there is less agreement for marriage bonuses that increase in gross earnings. Indeed, respondents' preferred marriage bonus are in line with the status quo for couples with a more equal within-couple income distribution (see Figures 4.12b and 4.12c).

Figure 4.11. – (Preferred) Tax Progressivity, by Vignette Household Type



(A) Singles (B) Unmarried Couples (C) Married Couples

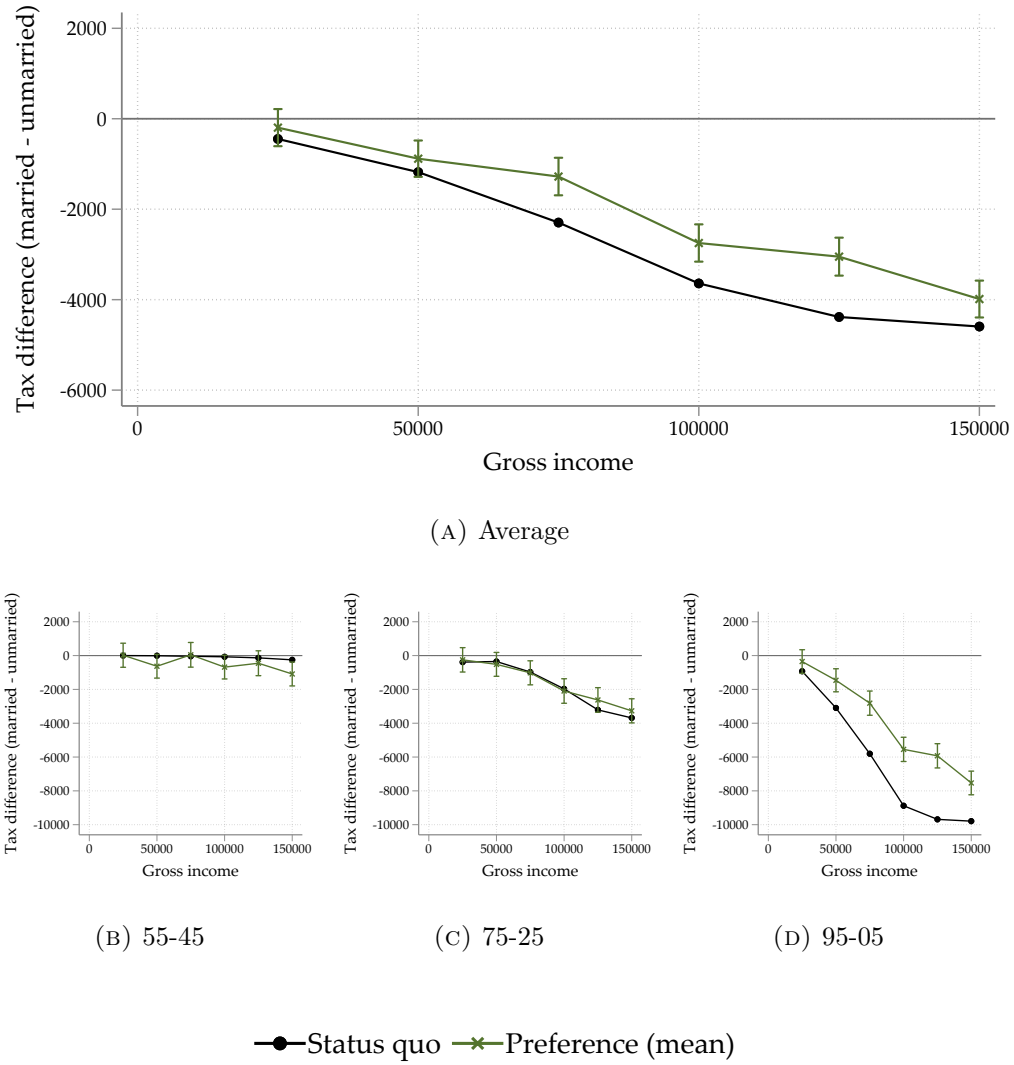
Note: This figure displays the status quo (gray dots) and average preferred (colored dots) tax liability across different gross incomes and separately for single households (orange dots), unmarried couples (red dots) and married couples (green dots). For the average preferences, 95 percent confidence intervals are included. Results are based on the predicted average of a fully saturated linear regression model including all interactions of household type, gross income, primary earner income share, and the number of children (see Table 4.1). We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

The previous sections revealed that respondents show strong support favoring families with children in the tax code. Figure 4.13 reveals that this support also holds taking into account that households without children pay more taxes in this case. However, the preferred benefits for a household with children are not larger than under the status quo. For single vignette households, the benefit of children is even slightly lower than under the status quo. This suggests that respondents misperceived the current tax treatment of children in the tax code when opting for larger benefits for families with children, or that they overlooked the trade-offs that higher benefits for children imply.

Self-Interest. In Appendix Figure D.17, we show that the behavior of respondents in the distribution of tax liabilities to different household pairs is again in line with predictions based on material self-interest. In particular, while married respondents agree on average to the marriage bonus in the status quo, non-married respondents on average want to reduce the marriage bonus. Also, while respondents with children have the desire to increase benefits for vignette households with children, respondents without children want to reduce these benefits compared to the status quo.

Results: 2) Marginal Tax Rates. Apart from changing the tax liability, we also gave respondents the possibility to change marginal tax rates of primary and secondary

Figure 4.12. – (Preferred) Marriage Bonus



(A) Average

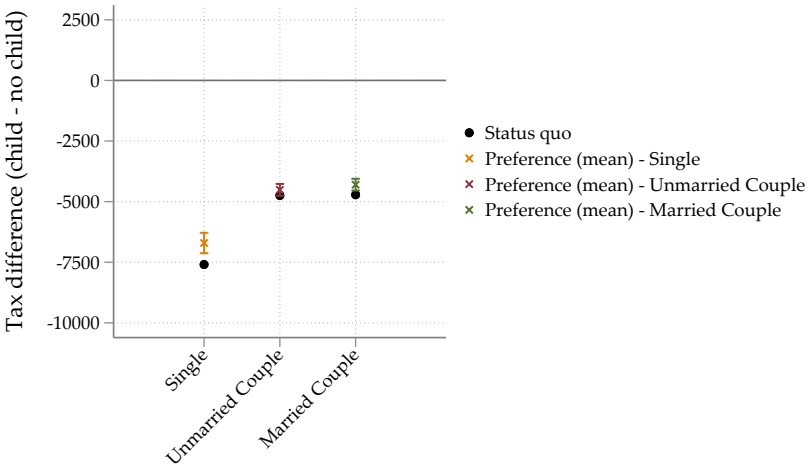
(B) 55-45

(C) 75-25

(D) 95-05

Note: This figure displays the status quo (gray dots) and average preferred (colored dots) tax liability difference between an unmarried and a married couple across different gross incomes. Figure 4.12a presents the average (preferred) tax liability difference while Figures 4.12b-4.12d differentiate between very equal distribution of incomes within a couple (55-45) and very unequal distribution of incomes within a couple (95-05). Results are based on the predicted average of a fully saturated linear regression model using all married and unmarried couples and including all interactions of household type, gross income, the primary earner income share and the number of children (see Table 4.1). We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

Figure 4.13. – (Preferred) Tax Benefit for Children



Note: This figure displays the status quo (gray dots) and average preferred (colored dots) tax liability difference between a tax unit with two children and a tax unit with zero children across different household types. Results are based on the predicted average of a fully saturated linear regression model including all interactions of household type, gross income, the primary earner income share and the number of children (see Table 4.1). We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

earners. In particular, we gave them the possibility to change them independently of each other. Thus, the respondents were given the possibility to move away from joint taxation and towards a differential tax treatment of primary and secondary earnings. For this purpose, we first informed respondents about how much extra taxes the primary and secondary earner would have to pay if their gross income increased by 100 euro respectively. Subsequently, they were asked separately, how much additional taxes the primary and secondary earner *should* pay on these next 100 euro.

We are particularly interested in the extent to which respondents want to change the tax system's degree of jointness, i.e., the interdependence of the marginal tax rates on primary and secondary earnings.[24] Empirically, jointness can be characterized via the correlation between the marginal tax rate of the primary and secondary earner. Table 4.2 shows the actual and preferred pairwise (PW) correlation coefficients. For vignettes with married couples, the preferred correlation is lower than the actual correlation suggesting that respondents prefer a lower degree of jointness indicated by a preferred average correlation of 0.6.

As shown by Figure 4.14, the preferred jointness varies in two dimensions. First, at a given income level of the married couple, the preferred jointness declines with a higher

[24]Note that jointness is positive in the status quo. If one spouse earns more, this increases the marginal tax rate for both, and hence the marginal tax rate on the earnings of the other spouse.

primary earner income. Second, at a given primary earner income share, the preferred jointness declines with the income level of the couple. However, in contrast to prescriptions from theoretical results on optimal taxation, e.g., Bierbrauer et al. (2024), the preferred correlation at the top is still substantial and not close to zero.

Table 4.2. – MTR Correlations, Status Quo vs. Preferences

Couple Type	Distribution of Income	Actual PW-Correlation	Preferred PW-Correlation	Difference
Unmarried Couple	Average	0.22 (0.007)	0.47 (0.008)	0.25
Married Couple	Average	0.76 (0.004)	0.61 (0.007)	-0.15

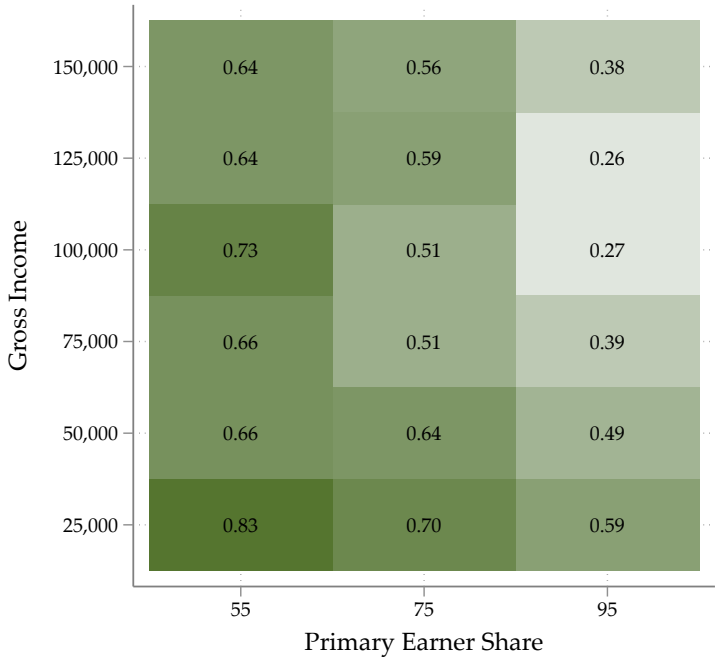
Note: This table presents correlations between primary and secondary earner marginal tax rates under the status quo and according to respondents preferences. We provide correlations separately for unmarried couples and married couples. We use survey weights provided by *forsa*.
Source: Own calculation based on *forsa* survey module.

Self-Interest. In Appendix Figure D.18, we show that self-interest plays a limited role for the assessment of marginal tax rate differences between primary and secondary earners in married couples. Across different types of respondents, primary earners in vignette couples are assigned a marginal tax rate that is between 3 and 4 percentage points higher than the one for secondary earners.

Results: 3) Support for Reforms. After showing respondents the consequences of three different tax reforms (limited marriage bonuses for married couples, more tax advantages for families with children, individual taxation) for different household types, we elicited the the support for these reforms a second time. We elicited the support whenever respondents had seen the reform consequences for a specific pair of households. Thus, respondents were given the possibility to update their approval or disapproval with every household pair they saw.

Figure 4.15 displays the share of respondents keeping (switching) between approval and disapproval of a reform compared to their answers to the " simple" questions at the very beginning of the survey (see Section 4.4.1). Seeing the consequences of a reform at the household level, increases the support for limited marriage bonuses somewhat (55 vs. 48 percent support). For individual taxation, while the aggregate support barely changes (44 vs. 43 percent support), 41 percent of respondents change their opinion compared to the question at the beginning of the survey, although in opposite directions. This stands in contrast to the support for higher benefits for families with children. For this reform, 35

Figure 4.14. – Preferred Jointness



Note: This figure displays the preferred correlation between the marginal tax rates of primary and secondary earners for different vignette income levels and different primary earner income shares.
Source: Own calculations based on *forsa* survey module.

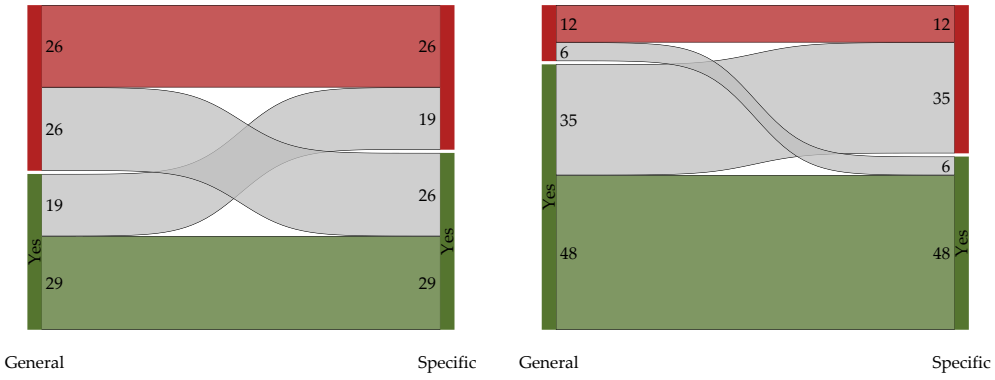
percent of respondents that initially agreed to the reform change their opinion and now oppose it. In the aggregate, this reduces the support for such a reform from 83 percent to 54 percent.

Self-Interest. In Appendix Figure D.19, we show how the change or persistence in reform support after information provision through the factorial survey is related to self-interest. For this purpose, we classify *forsa* respondents into winners and losers of the reform by simulating the respective change in tax liability at the individual level using the ifo microsimulation model in combination with the detailed information on household characteristics in the *forsa* survey. We further account for the fiscal effects of reforms by redistributing revenue gains and losses at the individual per-capita level.[25]

For realsplitting (individual taxation), the analysis illustrates that 32 (31) percent of individuals consistently act in line with self-interest both before and after the factorial survey. For family splitting, only 24 percent consistently act in line with self-interest before and after the reform. However, after respondents have seen the consequences of

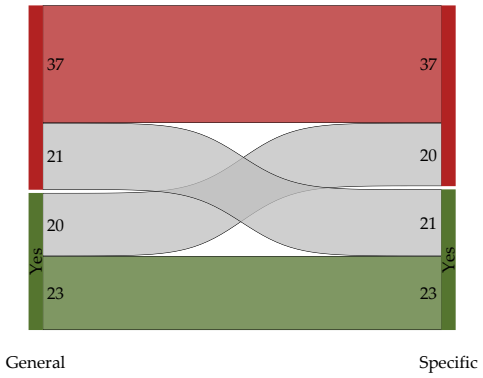
[25]For details on the estimated fiscal effects of the reforms, see Appendix D.1.4.

Figure 4.15. – Approval of Reforms, Changes through Information



(A) Limit Benefit of Married Couples (*Realsplitting*)

(B) Higher Benefit for Families w/ Children (*Familiensplitting*)



(C) Individual Taxation

Note: This figure displays how the support for three reform options concerning the taxation of couples changes after the respondents are provided with detailed information on the status quo and the consequences of specific reforms. The information is provided through three different household pairs. We exclude respondents that answer "Don't know" or "Indifferent" in either the general elicitation or the specific elicitation, since the different response levels do not allow for a consistent comparison. Each respondent is included three times in the data, since specific support was elicited for each of the three household pairs shown to respondents separately. We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

the reform through the household pairs, 28 percent switch sides in line with self-interest and start to oppose the reform. Therefore, the large shift in the support for family splitting comes primarily from individuals that are considered losers of the reform.

4.5 CONCLUSION

The taxation of couples is a widely debated topic among academics and policy makers alike. Given the vast theoretical and empirical literature pointing to the adverse effects of joint taxation – both in terms of efficiency and for gender equality – it is a puzzle that joint taxation is still present in many industrialized countries including Germany, the United States, and France.

This paper sheds light on the puzzle by providing evidence that the persistence of joint taxation (income splitting) in Germany can be explained by political economy considerations. We make use of two complementary approaches. First, we construct preferences over tax policies from an assumption of utility-maximizing behavior. We then employ sufficient statistics based on *material self-interest* to show that the share of winners based on from a reform towards individual taxation is close to but below fifty percent. We also show that different welfare objectives judge the desirability of such a reform in different ways. The welfare assessment depends on the trade-off between the consequences for "the poor" who tend lose and the consequences for working women who tend to gain from it.

Second, we employ a large-scale representative survey among the German population to elicit *stated preferences* and attitudes towards couples taxation. Main predictions from the analysis of support based on material self-interest are confirmed across different elicitation methods. Married couples with unequal distribution of earnings oppose reforms towards individual taxation while single earner couples agree to such reforms. This implies that reforms towards individual taxation are controversial across household types. In sum, we find less support for a reform towards individuals than in our analysis based on self-interest, even after respondents are informed in detail about the consequences of these tax reforms for different household types. Most respondents further prefer a tax system with positive jointness, i.e. a tax system so that the tax rate on the earnings of one spouse tends to increase in the earnings of the other spouse. This is consistent with their opposition to individual taxation.

Chapter 5

Just Cheap Talk? Investigating Fairness Preferences in Hypothetical Scenarios

This chapter is based on co-authored work with Paul Hufe. A previous version has been published as a CESifo Working Paper. See Hufe and Weishaar (2025) for the full reference.

5.1 INTRODUCTION

There is expanding literature in economics and other social sciences that investigates which inequalities are seen as unfair by people (Alesina et al., 2018; Almås et al., 2020; Andre, forthcoming; Cappelen et al., 2007; Jasso and Webster, 1999; Konow, 2000). In these papers, fairness preferences are typically elicited using incentivized experiments or non-incentivized surveys. Researchers considering the choice between these two research designs face a trade-off. On the one hand, experiments combine stylized representations of real-world situations with payout-relevant choices of respondents. On the other hand, survey questions often mirror real-world contexts more closely; however, respondents' answers have no consequences in the real world. Therefore, survey-based methods are often considered unreliable predictors of actual behavior. This raises the question of whether researchers can employ non-incentivized surveys to analyze fairness preferences or whether such answers must be considered "cheap talk." In this paper, we address this question by using a representative sample of the US adult population to test whether answers to hypothetical questions align with those from an incentivized experiment.

Our survey tool integrates core functionalities of impartial spectator experiments (Almås et al., 2024a; Almås et al., 2020; Andre, forthcoming; Cappelen et al., 2013; Konow, 2000; Konow et al., 2020) with the methodological advantages of factorial surveys (Auspurg et al., 2017; Gaertner and Schwettmann, 2007; Jasso and Webster, 1999; Konow, 1996).[1] The questions in our survey tool show respondents pairs of hypothetical persons that are described in terms of observable characteristics, i.e., their gender, age, educational attainment, parental background, working hours, and labor market earnings. Respondents then distribute earnings between these two persons based on the given information.

In impartial spectator designs, choices are not payoff-relevant for the respondents themselves. Yet, answers in a hypothetical task might still differ from their incentivized analogs. First, hypothetical distribution tasks are prone to different biases, including experimenter demand effects and social desirability biases (Stantcheva, 2023). For instance, in the absence of real-world consequences, respondents may tend towards more equal income allocations to comply with perceived expectations from the researcher. Such demand effects are less of a concern when real money is at stake (Haaland et al., 2023). Second, even without such systematic biases, hypothetical distribution tasks may be prone to measurement error, e.g., if non-incentivized respondents are less attentive when completing the tasks.

[1]For detailed reviews on experimental and survey-based evidence on fairness preferences, see Almås et al. (2023), Almås et al. (2024b), and Gaertner and Schokkaert (2012).

Our survey tool differs from prior literature on fairness preferences which has largely focused on the extent to which individuals reward rather abstract concepts such as "luck," "productivity," "hard work," and "talent" (e.g., Cappelen et al., 2010; Mollerstrom et al., 2015). In contrast to these studies, our survey tool allows us to elicit fairness preferences that can be directly mapped to observable labor market inequalities, e.g., gender gaps (Blau and Kahn, 2017), returns to hours (Kuhn and Lozano, 2008), education premia (Harmon et al., 2003), and intergenerational persistence (Roemer and Trannoy, 2016). However, the focus on these real-world inequalities also makes it prohibitively costly to elicit the relevant preferences in an experimental design where respondents' choices are consequential for the earnings of actual persons. Therefore, it is important to test whether hypothetical distribution tasks deliver credible results that align with the "gold standard" of incentivized experiments.

To address this question, we collected data from 1,602 adults in the United States between October and November 2022. The sampling was designed to be representative of various demographic characteristics such as gender, age, education, employment status, and region of residence. The survey modules, as well as core analyses, were pre-registered via the Open Science Framework (OSF), No. DV3KP.

We validate our survey tool along three dimensions. First, we test whether the distributional choices of respondents are different if they are payoff-relevant. For this purpose, we run an experiment with a between-subject design. All respondents answer a survey where they face a selection of tasks from our survey tool. Respondents in the treatment group are informed that one of the persons shown to them is a real person and that the decision made by a randomly chosen respondent will determine the monthly earnings of this individual. Thus, in contrast to the control group, they know that each choice may have substantial financial consequences for a real person. This design allows us to test whether fairness preferences in our hypothetical tool are consistent with the "gold standard" of an incentivized experiment (Bauer et al., 2020; Enke et al., 2022; Falk et al., 2023). Second, we test whether the distributional choices are stable over time. For this purpose, we employ a within-subject design and run an obfuscated follow-up one week after the baseline survey.[2] In particular, we invite respondents to another survey, where they again face a selection of tasks from our survey tool. Some tasks are repeated from the baseline wave, allowing us to calculate intertemporal correlations. This design enables us to test the stability of fairness preferences in our hypothetical tool and gives crucial information on measurement error in the elicited preference data (Gillen et al.,

[2]This choice of the time lag is consistent with several recent articles that validate survey-based measurement tools using time-lags between one and two weeks (Bauer et al., 2020; Enke et al., 2022; Falk et al., 2023; Fallucchi et al., 2020).

2019; Stantcheva, 2023). Lastly, next to the methodological validation of the survey design, we conduct a suggestive substantive analysis. Specifically, we describe the nature of fairness preferences identified through our survey and the heterogeneity of fairness views within the US population. This analysis has several caveats since the survey design was premised on methodological validation. Nevertheless, the substantive analysis provides an important cross-check on whether our hypothetical survey tool recovers preferences that are consistent with previous studies on fairness preferences in the US (Almås et al., 2020; Fisman et al., 2023; Konow et al., 2020).

Our results can be summarized as follows. First, the distributional choices of respondents are not affected by making them relevant to the earnings of real persons. The point estimates for the "real-person treatment" are small and insignificant at conventional levels of statistical significance. This conclusion remains unaffected when considering treatment effects on the distribution of choices and treatment effects within various population subgroups. Second, the distributional choices of respondents are relatively stable over time. The average (intra-respondent) intertemporal correlation of distributional choices is 0.56, which lies in the range of test-retest correlations of other preference survey modules (Enke et al., 2022). Furthermore, the intertemporal correlation is slightly higher in the incentivized group, suggesting that incentives have a small positive effect on reducing measurement error in the elicited preferences. Third, we find that the nature of the recovered fairness preferences is broadly consistent with previous studies on the US. Inequality acceptance ranges between Gini coefficients of 0.30 and 0.53 (e.g., Almås et al., 2020), the majority of respondents adopt intermediate fairness positions that are influenced by discretionary variables such as education and working hours (e.g., Konow, 2000), and the distributional choices of different population subgroups are consistent with self-serving biases (Costa-Font and Cowell, 2014). In summary, the results from our validation suggest that the proposed hypothetical survey tool recovers fairness preferences that are consistent with incentivized choices, stable, and reasonable in light of the existing literature.

This paper contributes to two strands of the literature. First, we contribute to the literature on fairness preferences. There is a large literature in economics and other social sciences trying to understand the nature and anatomy of fairness preferences in different population groups (Almås et al., 2020; Andre, forthcoming; Cappelen et al., 2007; Gaertner and Schokkaert, 2012; Jasso and Webster, 1999; Konow, 2000; Starmans et al., 2017). In this paper, we validate a vignette-based survey tool that allows researchers to investigate fairness preferences in a flexible and cost-efficient way. Therefore, this study provides a crucial step to strengthen the methodological toolkit for investigating fairness preferences in applied research. Furthermore, there is a growing literature investigating the upstream determinants and downstream consequences of fairness preferences (Adriaans,

2023; Alesina et al., 2018; Andersen et al., 2023; Fehr et al., forthcoming). These studies often rely on survey-based measures of fairness preferences. Our results provide encouraging news for such research designs as the consistency of hypothetical and incentivized choices suggests that survey-based measures are not systematically biased compared to their incentivized analogs. Second, we contribute to a growing methodological literature that validates survey-based measurement tools in various domains, including risk, time, competition, and social preferences (Bauer et al., 2020; Enke et al., 2022; Falk et al., 2023; Fallucchi et al., 2020). To the best of our knowledge, our study is the first validation of an impartial spectator task under high monetary stakes.

The remainder of the paper is organized as follows. Section 5.2 describes the survey tool and provides information on the data collection. In Section 5.3, we present results for the effects of the "real-person treatment." Section 5.4 describes the stability of fairness preferences. Section 5.5 provides a suggestive comparison of the recovered fairness preferences relative to existing literature. Section 5.6 concludes the paper.

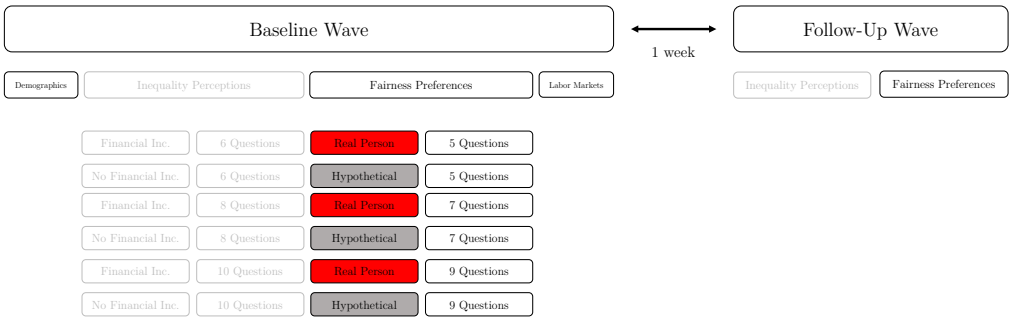
5.2 SURVEY TOOL AND DATA COLLECTION

Survey Structure. Figure 5.1 provides an overview of the survey used in our analysis. The survey is structured into two waves, each consisting of multiple modules. In the first module of the baseline wave, we elicit the demographic characteristics of respondents. The second and third modules measure inequality perceptions and fairness preferences. The final module contains additional questions about the labor market. The two modules of the follow-up wave mirror the perceptions and preference modules of the baseline wave. In this paper, we focus exclusively on fairness preferences.[3]

Preference Module. The module on fairness preferences consists of multiple questions that follow the design of factorial survey experiments. Factorial survey experiments are well-established tools in the social sciences to assess preferences and beliefs (Auspurg et al., 2015; Auspurg et al., 2017; Fisman et al., 2020; Jasso and Webster, 1997; Jasso and Webster, 1999; Wiswall and Zafar, 2018). In such experiments, respondents evaluate multiple hypothetical scenarios that vary at random in pre-defined characteristics. The random variation of characteristics has two main advantages. First, the design can replicate the complexities of the real world. In particular, respondents are forced to make trade-offs and weigh the importance of different real-world attributes against each other

[3]The perceptions modules are designed to assess respondents' perceptions of inequality in the labor market. All treatments in this module are independent of the treatments in the preference module, allowing us to analyze these data in isolation.

Figure 5.1. – Survey Structure

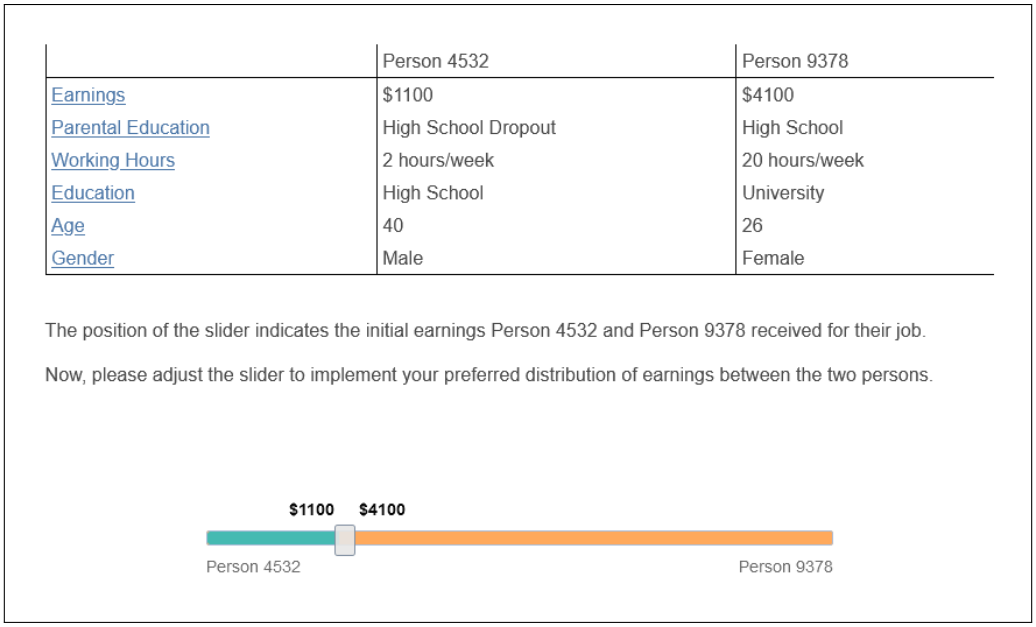


Note: This figure visualizes the structure of the survey with two waves (baseline, follow-up). Each wave consists of multiple modules. The modules on inequality perceptions are blurred out since they are not covered in this paper. The main treatment group (control group) is highlighted in red (gray).

when making their choices. Second, the simultaneous variation of characteristics mitigates experimenter demand effects and social desirability biases—concerns that are particularly relevant in the domain of fairness preferences (Auspurg and Hinz, 2014).

In our survey module, respondents receive information on the observational characteristics of two persons—see Figure 5.2 for an example. We describe these persons in terms of six characteristics, i.e., their gender, age, own education, parental education, working hours, and labor market earnings. The choice of these characteristics was guided by the following considerations. First, we selected the dimensionality of the vignettes by acknowledging the trade-off between precision and complexity in the design of factorial experiments. On the one hand, a higher number of vignette characteristics has the benefit that more confounding factors can be held constant. On the other hand, increasing the number of characteristics also puts a larger cognitive burden on participants when making distributional choices. With six dimensions, we opted for an intermediate level of complexity (Auspurg and Hinz, 2014). Second, we then selected characteristics that are both relevant for understanding earnings inequality and widely available in standard household survey data (Bick et al., 2022; Goldin, 2014; Lemieux, 2006; Magnac and Roux, 2021; Mazumder, 2005). We are aware that the resulting selection of characteristics is somewhat discretionary. Therefore, we also validate our selection ex-post by asking respondents about characteristics they consider particularly important when making distributional choices. Appendix Figure E.1 shows that five out of the six selected characteristics are chosen by at least ten percent of respondents. Furthermore, many of our chosen characteristics correlate strongly with other characteristics that participants find important, e.g., age (our selection) and work experience (respondents selection).

Figure 5.2. – Fairness Preference Elicitation, Exemplary Question



Note: This figure provides an example of a question screen in the fairness preference module. Each question shows the characteristics of two persons in six dimensions (earnings, gender, age, education, parental education, working hours) in a table format. Each of the two persons has been allocated a random person identifier from 1 to 9999. The ordering of characteristics in the table is randomized at the respondent-question level. A slider allows respondents to select their preferred distribution of earnings between the two persons. The chosen allocation is also shown numerically above the slider.

Each characteristic can take multiple expressions. For instance, the characteristic of education can take three values, i.e., *High School Dropout*, *High School*, or *University*. The potential expressions for each characteristic are shown in Table 5.1. The order of characteristics is randomized at the respondent-question level to ensure that results are not driven by order effects (Day et al., 2012).

Combining all potential expressions yields a set of 720 profiles ($2 \times 2 \times 3 \times 3 \times 4 \times 5$) and a set of 258,840 ($\frac{720 \times 719}{2}$) unique unordered profile pairs. In the following, we will refer to each unique profile as a *vignette person* and each unique profile pair as a *vignette*. We show respondents a selection of vignettes determined via a random draw from the full set. Based on the presented information, respondents can use a slider to adjust the initial earnings and to implement their preferred earnings distribution in the vignettes. All respondents answered the same selection of vignettes, with some variation in the number of vignettes per respondent—see also our discussion on the "length treatment" below.

Treatment 1: "Real-Person Treatment." To assess whether hypothetical questions in factorial surveys recover fairness preferences consistent with incentivized experiments,

Table 5.1. – Fairness Preference Elicitation, Characteristics and Expressions

Characteristic	Number	Displayed Values for Expression
Gender	2	Male / Female
Age	2	(26 , 35 , 40) / (50, 55, 59)
Education	3	High School Drop-Out / High School / University
Parental Education	3	High School Drop-Out / High School / University
Working Hours	4	(2, 5) / (20, 27, 31) / (39, 40) / (48, 51, 59)
Earnings	5	\$1,100 / \$2,700 / \$4,100 / \$5,900 / \$11,400

Note: This table shows the characteristics displayed in questions of the fairness preference module (column 1), the number of coarse expressions within each characteristic (column 2), and the displayed values for each expression (column 3). We employ a second randomization for age and working hours to display exact values instead of ranges. For each of the two age range groups (25-44, 45-65) and each of the four working hours range groups (1-9, 10-34, 35-44, More than 44), we draw from integers in the respective range.

we randomize respondents into two groups. Respondents in the control group complete a series of hypothetical distribution tasks. Respondents in the treatment group complete the same series of tasks. However, they are informed that one of the vignette persons is a real person. Furthermore, they are informed that the choice of one respondent will be selected to determine the monthly earnings of this person. Respondents know that the real person is between 25 and 65 years old, is a resident of the United States, and works in a job to earn money. Importantly, this information does not allow respondents to distinguish the real person from any other vignette person. They also know that the total earnings of the real person consist of two parts: (i) a fixed payment of \$500 and (ii) a flexible payment that can be changed by the respondent.

The research team hired a real person in August 2022—see Appendix Table E.1 for the characteristics of the real person as displayed in their vignette. The hired person was informed that the fixed-term contract would have a duration of one month and that the exact amount of their earnings would be determined by another individual; however, they did not know the exact process of how this happens. The person only knew their total earnings would be \$500 or above. To determine the potential earnings of the real person, we proceeded in two steps. First, we allocated the real person a monthly earnings value by randomly drawing from the set of potential earnings displayed in Table 5.1. Second, we randomly matched the real person with another (hypothetical) vignette person. This two-step procedure fixed the volume of earnings in the vignette of the real person at \$5,200. Therefore, including the fixed payment of \$500, the upper bound of potential earnings for the real person was \$5,700. This upper bound would be realized if the decisive respondent allocated all the vignette earnings to the real person. Importantly, the vignette with the real person was presented alongside all other vignettes, and the identity of the real person was concealed from respondents. Consequently, respondents also faced situations in which

the potential earnings implications were even higher. The average earnings volume in the displayed vignettes—and therefore the average upper bound of potential payments to the real person from the respondents' perspective— was \$11,420 in the main vignettes of the preference module. Furthermore, we ensured the salience of the real-world consequences through a training task. Specifically, we trained respondents on an example vignette and highlighted the potential earnings consequences for the real person after respondents had made their distributional choices.

Our incentivization is based on high stakes with a low probability of implementation. Alternatively, we could have used an incentive scheme based on lower stakes but a higher probability of implementation. Due to our focus on real-world characteristics, the former path appears more natural since it allows us to focus on payouts that reflect realistic monthly wages drawn from the US working population. However, we contend that a generalization of our results to alternative incentive schemes is an interesting path for future research.

Treatment 2: Length Treatment. To assess the sensitivity of our conclusions to the length of the survey, we vary the number of vignettes in the survey module. All respondents made at least five distributional choices; however, 1/3 of respondents received 2 or 4 additional vignettes, respectively. For the main validation, we focus on the first five questions answered by all respondents and use the variation from the length treatments in robustness analyses. The assignment to the length treatments was independent of the allocation to the "real-person treatment." Therefore, we obtain six groups of approximately equal size that vary in their exposure to the "real-person treatment" and the survey length (Figure 5.1).

Baseline Wave. We administered the baseline wave of the survey to 1,602 adult citizens of the United States. Data collection took place between October and November 2022. Respondents were contacted through the survey provider *Dynata* and received a participation payment depending on the expected survey length.[4] The mean (median) completion time was 24 (19) minutes for the baseline wave (Appendix Figure E.2).

Respondents were targeted to match the population along five dimensions (gender, age, education, employment status, and region of residence). In Panel A of Table 5.2, we compare our sample to the American Community Survey (ACS) regarding the targeted

[4]For the baseline wave, respondents were able to earn between \$0.20 and \$1.50. For the shorter follow-up wave, the payment varied between \$0.10 and \$1.20. The varying participation payment was used by *Dynata* to obtain responses from demographic segments of the population that are more difficult to reach.

characteristics. In general, we match the data well. Our sample has a slight under-representation of people with low education. In addition, we received over-proportional (under-proportional) responses from mid-western states (southern states). Panel B of Table 5.2 further shows that our sample is also broadly representative regarding other observable characteristics like ethnicity and income.

We take various steps to ensure the quality of survey answers. First, we included an attention check at the beginning of the preference module. Respondents who failed this attention check were screened out directly and were not part of the sample. Second, we asked a training question after explaining the tasks in the preference module. Around 72% of respondents passed this question on the first try. In robustness checks, we show that our results are not sensitive to excluding respondents who did not pass the training question on their first attempt.

Follow-Up Wave. We invited respondents to a follow-up wave one week after they completed the baseline wave. Among others, the follow-up wave consists of a fairness preference module with six questions. As in the baseline wave, respondents faced a selection of distribution tasks based on the vignettes from our survey module. Three of the questions are repetitions from the baseline wave, which were presented to the respondents in an obfuscated way. This feature allows us to assess the stability of fairness preferences over time.

The follow-up obtained a response rate of around 44%, and around 90% of respondents answered within two weeks. The resulting sample is slightly older but otherwise broadly comparable to our baseline sample regarding observable demographics (Appendix Table E.2). The mean (median) completion for the follow-up wave was 14 (9) minutes (Appendix Figure E.2).

Table 5.2. – Demographics, Comparison to ACS

Panel A: Hard Quota Demographics

Variable	Survey (%)	ACS (%)
Gender (N=1589)		
Male	48.58	48.66
Female	51.42	51.34
Age (N=1602)		
18 - 24	12.98	11.89
25 - 34	13.98	17.85
35 - 44	16.29	16.48
45 - 54	15.61	15.97
55 - 64	18.04	16.63
Older than 64	23.10	21.18
Education (N=1602)		
Lower Education	4.31	11.43
Middle Education	39.95	27.58
Higher Education	55.74	60.99
Employment Status (N=1602)		
In Labor Force	61.92	61.97
Unemployed	5.74	2.83
Not in Labor Force	32.33	35.19
Region (N=1602)		
North-East	19.35	17.43
Mid-West	33.52	20.76
West	23.91	23.76
South	23.22	38.05

Panel B: Other Selected Demographics

Variable	Survey (%)	ACS (%)
Ethnicity (N=1598)		
White / Caucasian	81.29	73.60
Black / African American	9.45	12.47
American Indian / Alaska Native	1.31	0.82
Asian / Asian American / Native Hawaiian / Pacific Islander	4.07	6.08
Other	3.88	7.03
Hispanic (N=1600)		
Hispanic	10.06	16.40
Not Hispanic	89.94	83.60
Income (N=1599)		
Did not work	22.33	32.83
$0 - $25,000	19.45	24.06
$25,000 - $40,000	17.07	13.27
$40,000 - $59,000	14.20	9.94
$59,000 - $88,000	12.26	9.83
More than $88,000	14.70	10.00

Note: This table compares the sample of our baseline survey to the American Community Survey (ACS, 2019). Panel A (B) shows hard quota demographics (other selected demographics). Sample sizes vary across variables since we omit small answer categories in gender ("Non-binary"), ethnicity ("Prefer not to answer"), hispanic ("Prefer not to answer"), and income ("Prefer not to answer"). Appendix Table E.2 presents the full set of demographics.
Source: Own calculations based on survey responses and American Community Survey (ACS, 2019).

5.3 EFFECTS OF THE "REAL-PERSON TREATMENT"

In this section, we investigate whether the potential for real-world implementation affected the distributional choices of respondents. First, we present methodological checks on the randomization and the anonymity of the real person. Second, we present the treatment effects on distributional choices. Third, we investigate potential heterogeneities by population subgroups. Fourth, we present robustness analyses. All analyses in this section are pre-registered unless noted otherwise.

Balancing. To give our estimates a causal interpretation, the treatment assignment must be uncorrelated with any respondent characteristics that may predict their distributional choices. Therefore, we test the balance of respondents' socio-demographic characteristics between the treatment and the control group. In particular, we regress the treatment status $Treat_i$ of respondent i on K pre-specified individual characteristics denoted by x_i^k:

$$x_i^k = \alpha^k + \beta^k \ Treat_i + \varepsilon_i^k. \tag{5.1}$$

Table 5.3 presents the results. In this table, we show sample sizes, point estimates of β^k, mean outcomes of the control group, and p-values associated with β^k. We account for multiple hypothesis testing by correcting for the family-wise error rate using the step-down procedure of Romano and M. Wolf (2005) and Romano and M. Wolf (2016).

For all considered characteristics, we find point estimates that are close to zero and small compared to the control mean. Furthermore, we cannot reject the null hypothesis of $\beta^k = 0$ at conventional levels of statistical significance. These results suggest that our randomization was successful and that we can give our treatment effects a causal interpretation.[5]

Anonymity of the Real Person. Out of all the vignettes faced by respondents, only one choice can determine the earnings of the real person. Respondents are not told which decision is hypothetical, and therefore, they are encouraged to treat each vignette as if it were payoff-relevant. The equal consideration of vignettes, however, would be threatened if respondents could identify the real person. In this case, the treatment would have no bite for vignettes that do not involve the real person (Andre, forthcoming).

[5]We pre-specified a joint balancing test that included a treatment in the preference module and a treatment in the perceptions module. We depart from the pre-analysis plan since we focus on the preference module in this paper. However, the pre-specified joint balancing test leads to the same conclusion as the balancing test presented above (Appendix Table E.3).

Table 5.3. – Balancing Tests

Demographics	Binary Split	N	Point Estimate	Control Mean	RW p-values
Gender	Male vs. Female	1589	-0.007	0.5176	0.987
Age	$<$ 45 years vs. $\geq$ 45 years	1602	-0.050	0.5925	0.241
Ethnicity	White/Caucasian vs. Other	1598	0.006	0.1842	0.987
Education	$\leq$ HS vs. $>$ HS	1602	0.037	0.5387	0.543
Employment	(Self)Employed vs. Other	1602	-0.018	0.3900	0.938
Hours	$<$ 35 hours vs. $\geq$ 35 hours	1468	-0.019	0.5249	0.938
Income	$<$ \$40,000 vs. $\geq$ \$40,000	1599	-0.003	0.4128	0.987

Note: This table presents results of the balancing test outlined in equation (5.1). The table shows point estimates for the coefficient of interest, the mean of the control group (first group in binary split), and associated heteroskedasticity robust p-values adjusted for multiple hypothesis testing using the Romano and M. Wolf (2016) step-down procedure with 1000 bootstrap draws. Sample sizes differ across demographic characteristics due to the exclusion of responses such as "Prefer not to answer" (see Table 5.2).
Source: Own calculation based on survey responses.

To address this concern, we asked respondents to identify the real person at the end of the preference module. To facilitate this task, we allowed respondents to look up all previous vignettes. Furthermore, we strongly incentivized the identification of the real person by paying out \$5 for a correct answer. Despite these incentives, only about 5% of respondents guessed correctly (Appendix Figure E.3). This number is close to the number that would be obtained if all respondents were to guess by chance.[6] This result suggests that we successfully preserved the anonymity of the real person and supports the assumption that respondents considered all choices equally likely to be payoff-relevant.

Treatment Effects on Distributional Choices. We estimate average treatment effects through ordinary least-squares using the following model:

$$\Delta y_{ij} = \alpha^j + \beta^j \, Treat_i + \mathbf{x}_i' \, \boldsymbol{\delta}^j + \varepsilon_{ij}, \tag{5.2}$$

where Δy_{ij} is the difference in money allocated to vignette persons A and B by respondent i in vignette j. $Treat_i$ is the binary treatment indicator. Per our pre-analysis plan, we also include the vector $\mathbf{x}_i'$ to control for demographic variables that are found to be unbalanced. Due to the successful randomization into treatment, this vector is empty (see Table 5.3).

Table 5.4 displays the results for each of the five vignettes that were answered by all respondents. We present point estimates of β^j, means of the outcome variables in the control group, and a set of three different p-values associated with β^j: (i) uncorrected

[6]Depending on the length treatment, the preference module contains five, seven, or nine questions. Therefore, if all respondents guessed randomly, we would expect that 7.6% ($= \frac{1}{3} \times [\frac{1}{10} + \frac{1}{14} + \frac{1}{18}]$) identify the real person correctly.

Table 5.4. – Allocations, "Real-Person Treatment"

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	1602	114.52	1038.02	0.324	0.322	0.841
Q2	1602	8.37	-2298.01	0.931	0.932	0.932
Q3	1602	211.77	2036.38	0.591	0.598	0.841
Q4	1602	-256.17	-4072.08	0.321	0.325	0.841
Q5	1602	-264.02	6184.43	0.343	0.337	0.841

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculation based on survey responses.

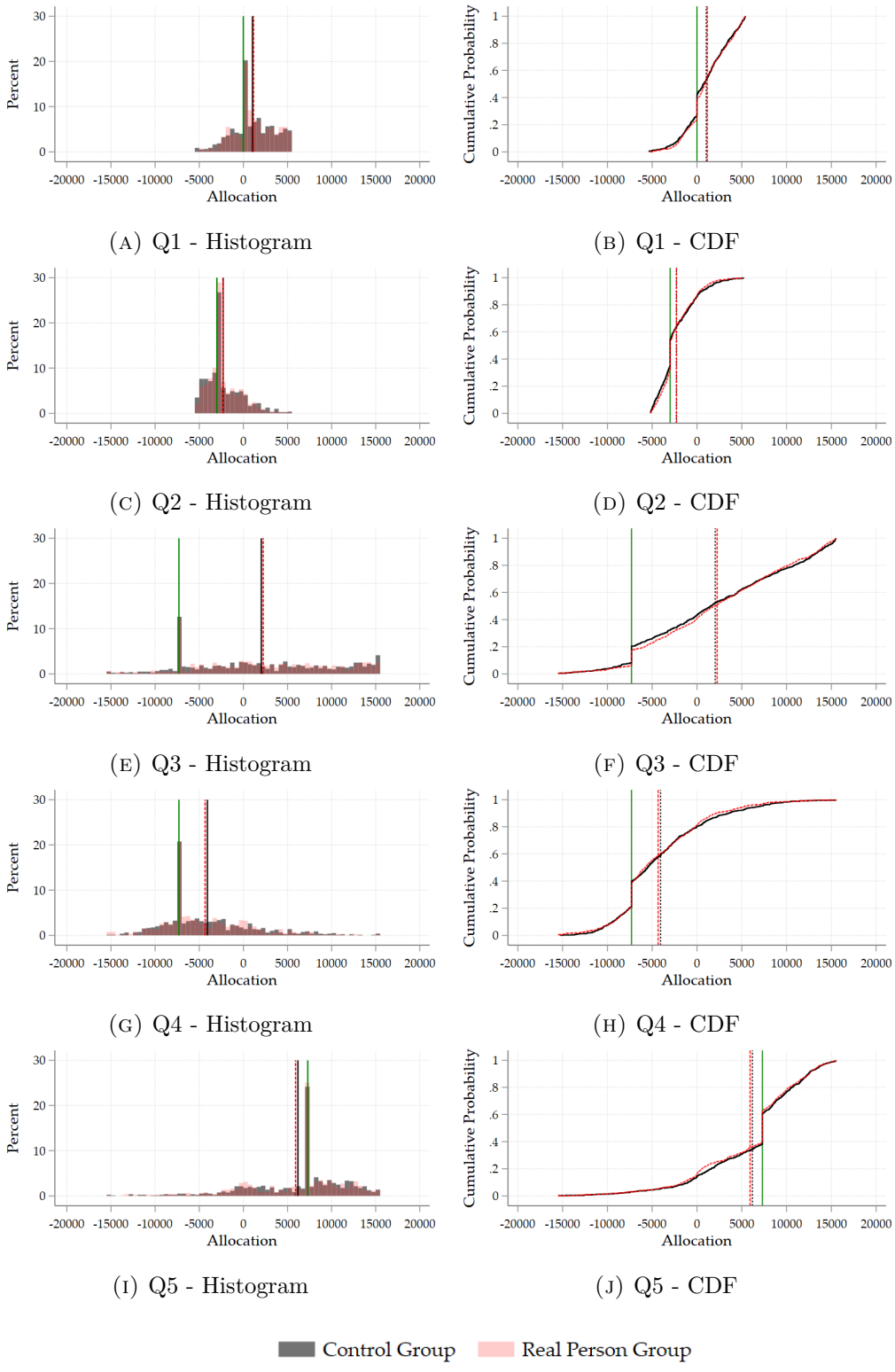
analytical model p-values, (ii) p-values based on bootstrapped standard errors, (iii) p-values that account for multiple hypothesis testing through the step-down procedure of Romano and M. Wolf (2005) and Romano and M. Wolf (2016).

Point estimates of β^j are small and range between \$8 and \$264. This range corresponds to 0.4%–11.0% of the control mean. For none of the five considered vignettes, we can reject the null hypothesis that average choices are equal across treatment and control groups. Appendix Table E.4 presents results from non-pre-registered analyses using alternative scalings of the outcome variable. In particular, we repeat the previous analysis by replacing Δy_{ij} with the income share allocated to person A in the vignette. In this case, we find treatment effects that range between 0.1 and 1.1 percentage points. This corresponds to 0.29%–2.25% of the control mean without any significant differences between the treatment and control groups.

The average treatment effects may conceal offsetting treatment responses at different parts of the outcome distribution. To assuage this concern, we plot distributions of Δy_{ij} separately for treatment and control groups (Figure 5.3). The left panels show frequency distributions, whereas the right panels show corresponding cumulative distribution functions for each considered vignette. Visual inspection shows that the distributions largely overlap. This diagnosis is confirmed by Kolmogorov–Smirnov tests, according to which we cannot reject the null hypothesis that the implemented distributions are the same in treatment and control groups (Appendix Table E.5). These results show that our null findings concerning average treatment effects are not the result of offsetting responses at different parts of the outcome distribution.

Heterogeneity Analysis. Different population groups might react differently to the potential real-life consequences of a distributional choice. The absence of heterogeneous

Figure 5.3. – Distribution of Allocations, "Real-Person Treatment"



(A) Q1 - Histogram (B) Q1 - CDF

(C) Q2 - Histogram (D) Q2 - CDF

(E) Q3 - Histogram (F) Q3 - CDF

(G) Q4 - Histogram (H) Q4 - CDF

(I) Q5 - Histogram (J) Q5 - CDF

Note: This figure displays how distributional choices vary between treatment and control groups. We focus on the first 5 questions that are answered by all respondents. The left panel displays histograms using fixed bins of \$500. The right panel displays cumulative distribution functions. Average choices of the treatment (control) group are represented by dashed red (gray) vertical lines. Solid green vertical lines visualize the status quo distribution of labor market earnings.
Source: Own calculation based on survey responses.

treatment effects, however, is important since many analysts might be interested in using hypothetical questions to assess differences in fairness preferences across population groups. If different population groups had a different propensity to reveal their preferences in hypothetical settings, analyses might erroneously detect group differences in fairness preferences, or the absence thereof, in such settings.

To test for heterogeneous treatment effects, we use equation (5.2) in split sample analyses for 14 pre-specified demographic sub-groups. For every population subgroup, we test one hypothesis for each of the five vignettes, i.e., we test 70 hypotheses in total. We summarize the resulting information in Table 5.5 by showing the number of hypotheses that are rejected at the 5% and 10% levels of statistical significance, respectively. As expected, without considering multiple hypothesis testing, some statistically significant differences emerge. For example, when considering model p-values, two out of the 70 tested hypotheses are found to be different from each other at the 5%-level. These differences vanish once we correct for multiple hypothesis testing.[7] This result suggests that the similarity of distributional choices between hypothetical and incentivized scenarios holds across a broad range of population subgroups.

Robustness. We implement a series of robustness checks to analyze the sensitivity of the previous findings.

First, one may be worried that our results are driven by low-quality answers from inattentive respondents. To address this concern, we implemented a control question to screen out inattentive respondents at the beginning of the survey. In addition, we pre-specified two alternative sample selection criteria that allow us to filter out low-quality responses. On the one hand, we excluded respondents who did not pass the training question in the preference module on the first try. On the other hand, we excluded respondents with extreme values in the response time distribution, i.e., we dropped respondents above the 90th percentile and below the 10th percentile of the module-specific response time distribution. Table 5.6 shows that none of the two restrictions alters the results substantially, suggesting that our results are not driven by low-quality responses from inattentive respondents.

Second, one may conjecture that treatment effects become more pronounced with the length of the survey since the potential real-life consequences encourage respondents to stay attentive for a longer time. We can test this conjecture by looking at differences in

[7]The adjustment for multiple hypothesis testing is made at the level of the population subgroup, i.e., we adjust for the fact that we test five hypotheses within each demographic subgroup.

Table 5.5. – Allocations, "Real-Person Treatment", by Demographic Subgroups

Variable	Value	N	Rejected Hypotheses at 5% (10%)		
			Model p-value	Resample p-value	RW p-value
Gender	Male	772	0 (1)	0 (0)	0 (0)
	Female	817	0 (0)	0 (0)	0 (0)
Age	$<$ 45 years	693	0 (0)	0 (0)	0 (0)
	$\geq$ 45 years	909	1 (2)	0 (2)	0 (0)
Ethnicity	White/Caucasian	1299	0 (1)	0 (1)	0 (0)
	Other	299	0 (0)	0 (0)	0 (0)
Education	$\leq$ HS	709	1 (1)	1 (1)	0 (0)
	$>$ HS	893	0 (0)	0 (0)	0 (0)
Employment	(Self)Employed	992	0 (0)	0 (0)	0 (0)
	Other	610	0 (0)	0 (0)	0 (0)
Working Hours	$<$ 35 hours	712	0 (0)	0 (0)	0 (0)
	$\geq$ 35 hours	756	0 (0)	0 (0)	0 (0)
Earnings	$<$ \$40,000	941	0 (0)	0 (0)	0 (0)
	$\geq$ \$40,000	658	0 (0)	0 (0)	0 (0)

Note: This table presents results of the regression analysis outlined in equation (5.2) for a subsample of respondents based on a particular demographic characteristic. We present the number of respondents, the number of rejected hypotheses according to heteroskedasticity robust model p-values, resample p-values, and p-values adjusted for multiple hypothesis testing. In total, 5 hypotheses are tested for each subsample. Detailed information on regression results are shown in Appendix Tables E.8 - E.21.
Source: Own calculations based on survey responses.

treatment effects across the length treatment groups (see Figure 5.1). In particular, we run split sample analyses for these three groups. We summarize the resulting information in Table 5.7 by showing the number of hypotheses that are rejected at the 5% and 10% levels of statistical significance, respectively. We cannot detect any significant treatment effects once multiple hypothesis testing is accounted for, irrespective of survey length. This result suggests that the answer quality of hypothetical survey modules on fairness preferences does not deteriorate with survey length.

Third, one may suspect that respondents exert more effort on the task if their decision has real-life consequences. Although we do not find any differences in choices in our setting, increased effort could lead to differential choices in settings that are more complicated than ours. To investigate this possibility, we use response times at the question level as a noisy measure for unobserved effort in the task and analyze how response times vary between treatment and control groups. Results in Table 5.8 show that there are

Table 5.6. – Allocations, "Real-Person Treatment", Alternative Samples

Panel A: Directly Passed Training Question

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	1154	168.87	1308.25	0.237	0.244	0.686
Q2	1154	31.76	-2574.27	0.774	0.777	0.949
Q3	1154	112.27	3320.52	0.813	0.813	0.949
Q4	1154	-147.78	-4508.78	0.613	0.629	0.935
Q5	1154	-290.95	6884.88	0.361	0.339	0.774

Panel B: Exclude Response Time Outliers

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	924	162.99	1394.95	0.317	0.321	0.766
Q2	924	109.22	-2699.89	0.364	0.375	0.766
Q3	924	-270.11	4030.09	0.606	0.607	0.854
Q4	924	-30.23	-4623.23	0.925	0.933	0.933
Q5	924	-477.60	7159.91	0.169	0.177	0.572

Note: This table presents results of the regression analysis outlined in equation (5.2) for different restricted samples. Panel A displays results for the sample of respondents that passed the training question of the fairness preference module at the first try. Panel B presents results for a sample that excludes respondents with high (above p90) and low (below p10) response times of the fairness preference module. We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications. See Appendix Tables E.6-E.7 for the corresponding balancing and Kolmogorov-Smirnov tests.
Source: Own calculation based on survey responses.

Table 5.7. – Allocations, "Real-Person Treatment", by Survey Length Subgroups

		Rejected Hypotheses at 5% (10%)		
Module Length	*N*	Model p-value	Resample p-value	RW p-value
5 Questions	534	0 (1)	0 (1)	0 (0)
7 Questions	536	0 (0)	0 (0)	0 (0)
9 Questions	532	1 (1)	1 (1)	0 (0)

Note: This table presents results of the regression analysis outlined in equation (5.2) for a subsample of respondents based on the survey module length, i.e, whether respondents answered five, seven, or nine questions in the preference module. We present the number of respondents, the number of rejected hypotheses according to heteroskedasticity robust model p-values, resample p-values, and p-values adjusted for multiple hypothesis testing. In total, we test five, seven, and nine hypotheses for each subsample depending on the number of questions. Detailed information on regression results are shown in Appendix Tables E.22 - E.24.
Source: Own calculations based on survey responses.

Table 5.8. – Preferences, Response Time (Min.), "Real-Person Treatment"

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	1585	0.00	1.00	0.892	0.904	0.989
Q2	1585	0.00	0.46	0.936	0.940	0.989
Q3	1585	0.00	0.50	0.837	0.839	0.989
Q4	1585	-0.01	0.37	0.501	0.555	0.936
Q5	1585	0.01	0.36	0.729	0.743	0.981

Note: This table presents results of the regression analysis outlined in equation (5.2) using the response time in minutes as the dependent variable. For every question, we focus on response times below the 99th percentile of the question-specific response time distribution. We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculation based on survey responses.

virtually no differences between treatment and control groups, suggesting that the effort of respondents does not decrease when facing hypothetical instead of incentivized scenarios.[8]

In this section, we have shown that there are no systematic differences in distributional choices between hypothetical and incentivized scenarios. This conclusion holds for average allocations, distributions of allocations, and within various demographic subgroups. Furthermore, this conclusion is robust to various sensitivity checks, such as excluding low-quality responses. In summary, the results suggest that hypothetical vignettes capture the same fairness preferences as their incentivized analogs.

5.4 STABILITY OF FAIRNESS PREFERENCES

In this section, we investigate whether the survey module captures genuine fairness preferences that are stable over time. In particular, we use the longitudinal variation between the baseline wave and the follow-up wave of the survey. The follow-up wave consists of a fairness preference module with six questions, three of which are repetitions from the baseline survey. To avoid respondents anchoring their responses on their answers in the baseline survey, we obfuscate the repeated questions by mixing them in random order with novel questions that have not been shown to respondents previously. We present results in three steps. First, we present intertemporal correlations based on the pooled

[8]To limit the influence of extreme outliers, we focus on respondents whose response time for a particular question is below the 99th percentile of the question-specific response time distribution. In Appendix Table E.25, we repeat the exercise without this restriction. Treatment effects increase due to single outliers in the treatment and control groups. However, none of the differences is statistically significant, and our general conclusion remains unaffected.

follow-up sample. Second, we investigate whether these intertemporal correlations vary by treatment status in the baseline survey. Third, we present robustness analyses. We registered the follow-up survey in our pre-analysis plan. However, since the survey provider expressed considerable uncertainty about the likely response rates, we did not pre-specify the associated analyses presented in this section.

Intertemporal Correlations. We estimate intertemporal correlations through ordinary least-squares using the following model:

$$\Delta y_{ij,t} = \alpha + \sigma \, \Delta y_{ij,t-1} + \varepsilon_{ij}, \tag{5.3}$$

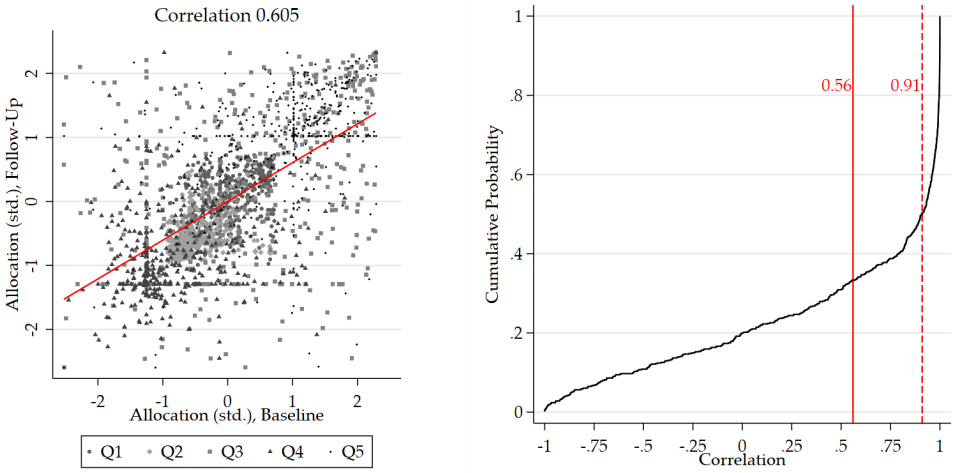
where Δy_{ij} is again the difference in allocations to vignette persons A and B by respondent i in vignette j in the baseline wave $(t-1)$ and the follow-up wave (t), respectively. In all estimations, we standardize $\Delta y_{ij,t}$ and $\Delta y_{ij,t-1}$ on the estimation samples such that they have a mean of zero and a standard deviation of one. As a result, estimates of σ can be interpreted as intertemporal correlation coefficients.

Figure 5.4a plots the raw standardized data of $\Delta y_{ij,t}$ against $\Delta y_{ij,t-1}$, with the fitted line indicating the point estimate of σ. The intertemporal correlation is estimated at 0.61, suggesting sizable stability of distributional choices over time.

Figure 5.4b visualizes the cumulative distribution of intertemporal correlations at the individual level. For each respondent, estimates of σ are based on the three repeated questions from baseline and follow-up. More than 80% of respondents display a positive correlation, and more than 65% have a correlation of 0.50 or higher. The mean (median) correlation across respondents is around 0.56 (0.91). These high intra-respondent correlations reaffirm our conclusion that the recovered distributional choices are fairly stable over time for most respondents in our sample.

Effect of "Real-Person Treatment." The estimated intertemporal correlations in equation (5.3) may be attenuated by measurement error in $\Delta y_{ij,t-1}$, i.e., the distributional choices in the baseline wave. Therefore, we can use estimates of σ to assess whether incentivized survey questions increase the signal-to-noise ratio in the recovered fairness preferences. If measurement error in $\Delta y_{ij,t-1}$ was less pronounced in incentivized scenarios,

Figure 5.4. – Stability of Fairness Preferences, Correlation



(A) Correlation (B) CDF

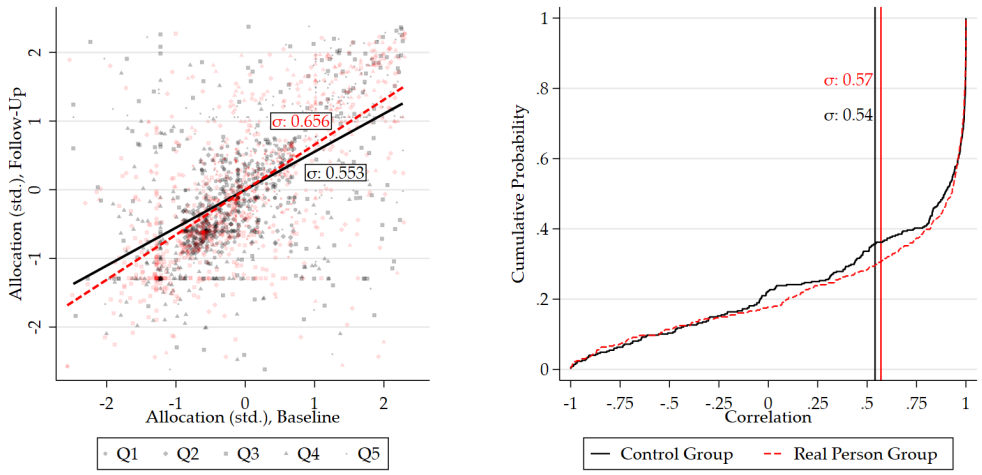
Note: This figure displays the intertemporal correlation between the baseline and follow-up wave (Figure 5.4a) and the cumulative distribution function of within-respondent correlations (Figure 5.4b). In Figure 5.4a, variables are standardized on the full sample, and the line indicates the line of best fit from a linear regression. In Figure 5.4b, variables are standardized at the individual level, and the solid (dashed) line indicates the mean (median) correlation across respondents.
Source: Own calculation based on survey responses.

σ would be significantly higher in the treatment group than in the control group. Such a finding would suggest that incentivized survey modules yield less noisy estimates of fairness preferences.[9]

In Figure 5.5a, we replicate Figure 5.4a by splitting our sample into the treatment and control groups from the baseline wave. Estimates of σ are slightly higher in the treatment (0.66) than in the control group (0.55). The difference of 0.10 is statistically significant at the five percent level (p-value=0.02). In Figure 5.5b, we show that the difference in stability is less pronounced when considering correlations at the individual level. The average intra-individual correlation is still slightly higher in the treatment (0.57) than in the control group (0.54). However, the difference of 0.03 is not statistically significant at conventional levels of statistical significance (p-value=0.49). These patterns suggest that incentivized survey modules may yield slightly less noisy estimates of fairness preferences. However, these gains are relatively moderate and may be quickly outweighed by the benefits of an unincentivized survey, e.g., lower cost, the potential to target broader population samples, etc.

[9]In the follow-up wave, all questions were hypothetical. Since we use $\Delta y_{ij,t}$ as outcomes in equation (2), the associated (classical) measurement error will not bias our estimates of σ.

Figure 5.5. – Stability of Fairness Preferences, "Real-Person Treatment"



(A) Correlation (B) CDF

Note: This figure displays the intertemporal correlation between the baseline and follow-up wave (Figure 5.5a) and the cumulative distribution function of within-respondent correlations (Figure 5.5b) separately for treatment and control groups. In Figure 5.5a, variables are standardized at the group level, and solid lines indicate lines of best fit from a linear regression. In Figure 5.5b, variables are standardized at the individual level, and solid lines indicate mean correlations across respondents.
Source: Own calculation based on survey responses.

Robustness. We again implement a series of robustness checks to analyze the sensitivity of the previous findings. These robustness checks are summarized in Table 5.9.

First, we check whether intertemporal correlations change when excluding low-quality answers from inattentive respondents who do not pass the training question on the first try. Intertemporal correlations increase slightly but remain very close to our full sample estimate. In an alternative test, we exclude respondents in the tails of the response time distribution. This sample restriction has virtually no effect on the estimated intertemporal correlations.

Second, we check whether the estimated intertemporal correlations are especially driven by individuals who always leave the slider close to its original position. In particular, we exclude observations where respondents leave the vignette slider within a two-sided five percentage point band around the initial earnings distribution in the baseline and the follow-up. Indeed, there seems a slight drop in intertemporal correlations when excluding these respondents. However, we also emphasize that the implemented test is likely too stringent. On the one hand, we exclude respondents who leave the slider unaltered in bad faith. On the other hand, we also exclude respondents with genuine libertarian

Table 5.9. – Stability of Fairness Preferences, Correlation, Robustness

	Full Sample	Restricted Sample		
		Training Question	Response Time	Not at Status Quo
Aggregate	0.605 (2127)	0.644 (1710)	0.609 (1707)	0.541 (1206)
Q1	0.375 (424)	0.389 (338)	0.391 (338)	0.413 (295)
Q2	0.328 (426)	0.335 (346)	0.298 (340)	0.436 (199)
Q3	0.407 (425)	0.390 (334)	0.399 (337)	0.356 (326)
Q4	0.332 (427)	0.337 (341)	0.312 (352)	0.317 (193)
Q5	0.382 (425)	0.377 (351)	0.335 (340)	0.497 (193)

Note: This table displays intertemporal correlations between the baseline and follow-up waves for the full sample and three restricted samples. The first restriced sample focuses on respondents that passed the training question of the fairness preference module in the baseline wave at the first try. The second restricted sample excludes respondents with high (above p90) and low (below p10) response times of the fairness preference module in the baseline wave. The third restricted sample excludes respondents whose allocated shares are at most 5 percentage points away from the status quo distribution of earnings. Variables are standardized on the sample used in the corresponding regression. Sample sizes are shown in parenthesis.
Source: Own calculations based on survey responses.

preferences. Therefore, we interpret the still substantial intertemporal correlation as a positive signal that we can recover stable preferences in areas further away from initial income positions.

Third, all previous conclusions hold when calculating intertemporal correlations at the level of individual questions. In our previous discussion, we especially focused on intertemporal correlations at the individual level. This is the appropriate level of analysis since factorial survey designs mostly use intra-respondent variation across multiple vignettes to identify the relevant preferences (Wiswall and Zafar, 2018). However, depending on the design, researchers may want to infer preferences from fewer vignettes per individual than in our setting. In Table 5.9, we, therefore, assess the extreme case where preferences would be identified based on a single question only. In this case, the preference signal is more noisy, translating into lower intertemporal correlations. Nonetheless, even in the extreme case of using only one vignette, the correlations are still substantial, ranging from 0.33 to 0.41 in the full sample.

In this section, we have shown that the distributional choices are relatively stable over time. This conclusion is robust to various sensitivity checks, among others, excluding low-quality responses. The presence of incentives slightly decreases the noise in elicited preferences. This decrease in noise, however, is fairly moderate and may be quickly outweighed by the potential benefits of running an unincentivized survey. In summary, the results suggest that hypothetical distribution tasks can yield high-quality data on stable fairness preferences.

5.5 NATURE OF FAIRNESS PREFERENCES

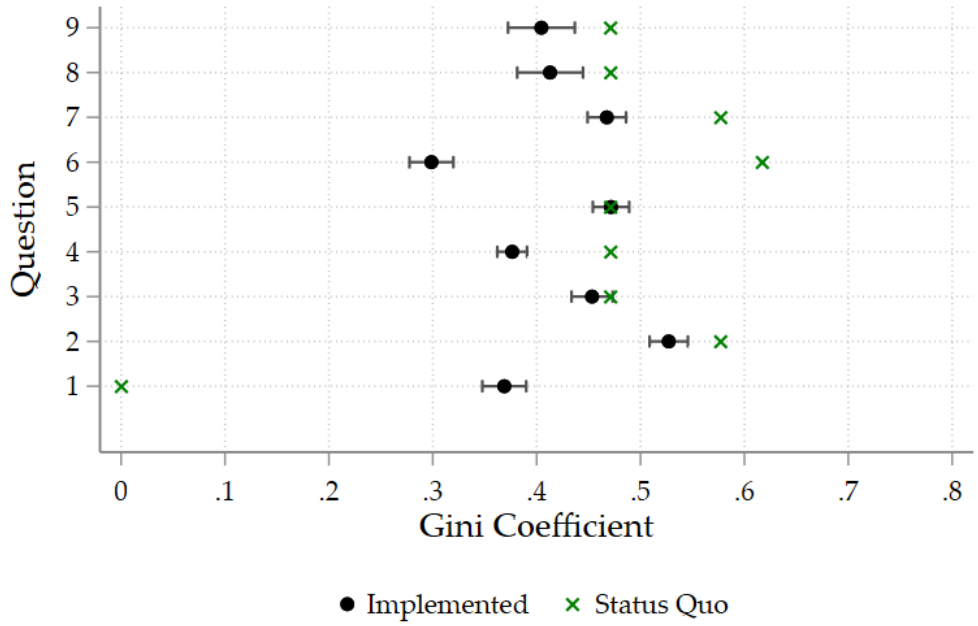
In this section, we accompany the main methodological validation of the previous sections by giving some suggestive insights into the nature of elicited fairness preferences. To be sure, this analysis comes with caveats. The primary purpose of this paper is to assess the measurement of fairness preferences in hypothetical settings as compared to the "gold standard" of incentivized experiments. Therefore, we made several methodological choices that prevent a full substantive analysis of the recovered preferences. For example, to maximize statistical power to detect differences between the treatment and control group, we show all respondents the same randomly selected subset of vignettes. Consequently, vignette characteristics are not equally represented, and correlations may exist among them. These features may affect respondents' willingness to tolerate inequality and how they incorporate different vignette characteristics into their choices. Therefore, we view the following analysis as a suggestive test of whether the recovered preferences are broadly consistent with findings from the existing literature.

With these caveats in mind, we will present the results of this section as follows. First, we analyze the level of inequality implemented by respondents. Second, we will analyze the prevalence of different fairness types in our sample. Lastly, we show the sensitivity of fairness preferences to different characteristics of the evaluated vignette persons. In all analyses, we will focus on unincentivized scenarios from the control group. However, our conclusions remain unaffected when focusing on the incentivized sample—see Appendix Figures E.4, E.6 and Appendix Table E.26 for replications of the main exhibits of this section based on the treatment group. The analyses of this section are exploratory. Therefore, they have not been registered in our pre-analysis plan.

Implemented Inequality. Figure 5.6 compares implemented inequality by respondents to the initial inequality separately for each vignette. The implemented Gini coefficients show substantial variation across vignettes (0.30–0.53). For comparison, Almås et al. (2020) use a representative sample of American respondents to show that they would implement a Gini coefficient of 0.35 (0.54) if the income-generating process were purely based on luck (merit). This suggests the range of implemented inequality across different scenarios in our setting is plausible.

In Appendix Figure E.5, we furthermore illustrate how inequality acceptance varies across respondents with different socio-demographic characteristics. Respondents who are more inequality-accepting tend to be older, more educated, and work longer hours. Those with a lower inequality tolerance tend to be female and non-white. Again, these patterns are

Figure 5.6. – Gini Coefficient



Note: This figure compares implemented Gini coefficients (gray dots) to initial Gini coefficients (green crosses) in each vignette. Gray bars indicate 95 percent confidence intervals. Gini coefficients are calculated at the respondent level as $\frac{|x-y|}{x+y}$ where x (y) is the amount allocated to vignette person A (B). We focus exclusively on respondents in the control group, i.e., those respondents who faced hypothetical scenarios. Appendix Figure E.4 replicates the analysis for respondents in the treatment group.
Source: Own calculations based on survey responses.

broadly consistent with existing literature. For instance, the findings of Almås et al. (2020) indicate that women and individuals with lower educational attainment are less accepting of inequality compared to men and those with higher education. The authors interpret these patterns in the light of potential self-serving biases in fairness preferences. It is reassuring that our survey replicates these patterns as well.

Fairness Types. Experimental literature has focused on estimating the prevalence of different fairness types that can be mapped to fairness principles in the philosophical literature—see Almås et al. (2024b) for a recent overview. On one end of the spectrum is the *egalitarian position*. Egalitarians consider all inequalities unfair, regardless of how these inequalities come about. Therefore, the egalitarian position prescribes an equal income distribution in any distributive situation. At the opposite end of the spectrum is the *libertarian position*. Libertarians consider all inequalities fair regardless of how these inequalities come about (Nozick, 1974). Therefore, the libertarian position prescribes a distribution of income that corresponds to the initial distribution in any distributive situation. Between these two extreme positions, there are several *intermediate positions*,

such as the responsibility-sensitive positions proposed by Arneson (1989), Cohen (1989), Dworkin (1981a), and Dworkin (1981b). These intermediate positions advocate for distinguishing between different sources of inequality, such as discretionary choices, ability, preferences, or circumstantial factors.

We estimate the prevalence of the egalitarian position by calculating the share of respondents who implement equal splits in all vignettes. Similarly, we estimate the prevalence of the libertarian position by calculating the share of respondents who accept initial inequality in all vignettes. When calculating these shares, we allow for "trembling hand" mistakes (Choi et al., 2007). For our baseline estimates, we use two-sided five percentage point bands around the egalitarian and libertarian answers to a vignette and allow respondents to be outside of the corresponding band for at most one vignette without repercussions on their classification as egalitarians or libertarians. We estimate the prevalence of the intermediate position as the remaining share of respondents who are not classified as egalitarians or libertarians.

Table 5.10 shows the results, where the highlighted areas represent our baseline estimates. Around two percent of respondents are classified as egalitarian, whereas around nine percent are classified as libertarians. The remaining 89% percent of respondents adopt intermediate positions. Therefore, most respondents adopt fairness positions that vary with the characteristics of the respective vignette. We note that this conclusion does not vary with the leniency with which we accept "trembling hand" mistakes. Even in the most lenient specifications where we allow for two-sided ten percentage point bands and two inconsistent answers, the share of respondents adopting intermediate positions is still 70%. Furthermore, we note that conditional on the adopted rule for "trembling hand" mistakes, the presented estimates for the prevalence of egalitarian and libertarian positions should be interpreted as upper bounds. We only presented respondents with a limited selection of five to nine vignettes. Therefore, in additional questions, the number of divergences from the egalitarian and libertarian positions can stay constant at best but not decrease. The estimated shares of egalitarians and libertarians in the US are smaller than the corresponding shares estimated in Almås et al. (2020). Their estimates classify 15% and 29% of the US population as egalitarians and libertarians, respectively. This difference may be rationalized by variations in how different preference types are identified or by the increased richness of the distributional scenarios in our setting.[10] Since the

[10] Almås et al. (2020) identify egalitarians as respondents who distribute resources equally in a situation where initial inequality is driven by productivity. They identify libertarians as respondents who do not redistribute at all when initial inequality is purely based on luck.

Table 5.10. – Preference Types

	Share Egalitarians (%)			Share Libertarians (%)		
Max. abs. difference (pp)	2	5	10	2	5	10
Allow for 0 inconsistent answers	0.00	0.51	1.73	5.68	6.41	7.65
Allow for 1 inconsistent answers	0.00	**1.46**	3.75	7.89	**9.47**	13.84
Allow for 2 inconsistent answers	0.89	2.71	7.69	10.75	13.90	22.22

Note: This table presents shares of egalitarians (libertarians) according to consistent choices in all questions of the baseline wave. We also vary the leniency of the classification by allowing for 0, 1, 2 answers that are inconsistent with egalitarian (libertarian) choices. In the baseline (highlighted estimates), we allow for a deviation of +/-5 percentage points and inconsistent choices in one question only. We focus exclusively on respondents in the control group, i.e., those respondents who faced hypothetical scenarios. Appendix Table E.26 replicates the analysis for respondents in the treatment group.
Source: Own calculations based on survey responses.

vignettes provide multidimensional information on the earnings-relevant characteristics of the recipients, respondents can express positions that deviate from the polar cases of egalitarian/libertarian fairness preferences in more nuanced ways.

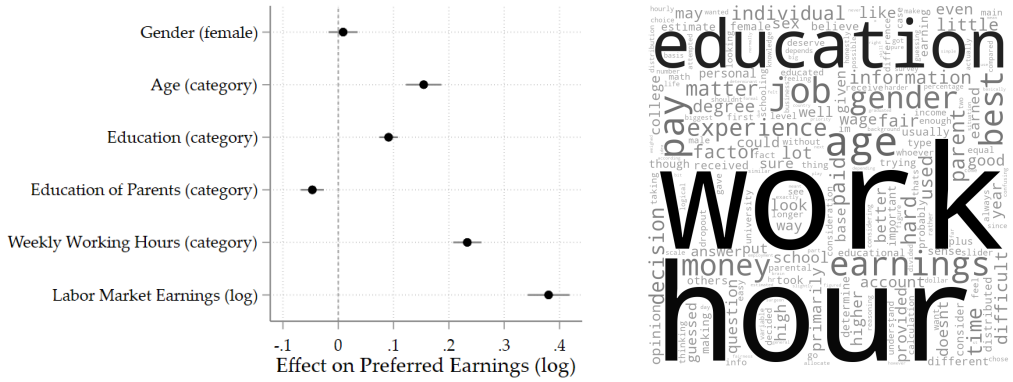
Importance of Vignette Person Characteristics. In the last step, we analyze the impact of particular earnings-relevant characteristics on the fairness preferences of respondents. To this end, we transform our data as follows. We create a data set where each row represents one person m from vignette j. Then, we replicate these data for each respondent i who made a distributional choice for vignette j and include the corresponding income allocations $y_{im(j)}$ as the outcome variable of interest. Stacking these data, we obtain a panel data set with multiple observations for each vignette person $m(j)$ and each respondent i.

We then estimate the following model via ordinary least-squares:

$$\begin{aligned} \ln y_{im(j)} = {} & \beta_1\, gender_{m(j)} + \beta_2\, age_{m(j)} + \beta_3\, educ_{m(j)} \\ & + \beta_4\, educpar_{m(j)} + \beta_5\, hours_{m(j)} + \beta_6\, \ln earn_{m(j)} \\ & + \theta_{[earn_{A(j)}+earn_{B(j)}]} + \epsilon_{im(j)}. \end{aligned} \tag{5.4}$$

The right-hand side variables in the first two lines of equation (5.4) represent the six vignette characteristics considered in our fairness preference module. The associated coefficients $\beta_1 - \beta_6$ capture the linear effect of personal characteristics on fair earnings conditional on the remaining vignette characteristics. We control non-parametrically for the total sum of vignette earnings $earn_{A(j)} + earn_{B(j)}$ by including corresponding fixed

Figure 5.7. – Fairness Preferences, Importance of Profile Characteristics



(A) Regression Analysis (B) Text Analysis

Note: This figure displays how respondents take the vignette characteristics into account for their distributional choices. Figure 5.7a displays the point estimates and 95 percent confidence intervals from equation (5.4). Figure 5.7b displays a word cloud from a text analysis of the open-ended question in the fairness preference module. At the end of the module, people were asked about how they came up with their distributional choices and to describe their reasoning in their own words. Based on the text corpus from the open-ended answers, we used natural language processing techniques to rank the frequency of specific terms. We focus exclusively on respondents in the control group, i.e., those respondents who faced hypothetical scenarios. Appendix Table E.27 replicates the analysis for different transformations of the outcome variable, and under the inclusion of respondent fixed effects. Appendix Figure E.6 replicates the analysis for respondents in the treatment group.
Source: Own calculation based on survey responses.

effects. Thereby, we account for the fact that vignette persons who are paired with high-earning persons in their vignette may mechanically receive higher income allocations. Standard errors are clustered at the respondent level.

The interpretation of the estimated coefficients comes with two important caveats. First, we cannot control how respondents associate the displayed vignette characteristics with other unobserved characteristics, such as productivity or job performance. Therefore, $\beta_1 - \beta_6$ are composite parameters that capture the reward for a certain characteristic, and the reward for unobserved factors correlated with it while holding all other observed characteristics constant. Second, we disregard non-linear effects across the ordinal vignette characteristics. The methodological choices mentioned at the beginning of this section limit the available variation in our data and prevent us from relaxing this stringent functional form assumption.

Figure 5.7a displays the coefficients from regression equation (5.4) along with the corresponding 95% confidence intervals. On the one hand, the results indicate that fair labor market earnings increase with vignette characteristics relevant to labor market performance and that are (partially) under the control of individuals. For example, there are strong positive effects of education and weekly working hours on fair income allocations.

This result suggests that fairness preferences in the US are at least partially consistent with normative theories that emphasize the role of discretionary choices in determining fair income shares, (e.g., Konow, 2000). Similar conclusions can be drawn for age and initial earnings, if interpreted as proxy indicators for relevant labor market experience and on-the-job performance, respectively.[11]

On the other hand, the results show that non-discretionary vignette characteristics are not rewarded in fair income allocations. For example, conditional on the other vignette characteristics, the point estimate of gender cannot be distinguished from zero, suggesting that US residents perceive adjusted gender gaps in labor market earnings as unfair. Furthermore, the respondents in our sample assign higher (lower) earnings to individuals with lower (higher) parental education. This finding could be rationalized by the fact that respondents are willing to compensate people for a disadvantaged socio-economic family background. Note, however, that our analysis only presents relative effects of vignette characteristics. Therefore, we cannot distinguish whether respondents compensate individuals for a disadvantaged background (low parental education) or penalize individuals coming from an advantaged family background (high parental education).

Given the abovementioned caveats, we interpret these findings as suggestive. However, in Appendix Table E.27, we show that all previous conclusions are robust to alternative specifications. In particular, we use preferred income shares or preferred income ratios as outcome variables of interest. While the magnitude of coefficients changes, the direction of effects remains unaltered. Also, our results do not change qualitatively when controlling for respondent fixed effects, i.e., when we only use within-respondent variation to estimate preferred rewards. In addition, while our baseline analysis only uses responses from respondents with hypothetical tasks, the measured rewards are similar in the real-person treatment group (Appendix Figure E.6).

To substantiate the quantitative evidence, we also use natural language processing techniques to present findings from a text analysis (Ferrario and Stantcheva, 2022). At the end of the baseline survey, we asked respondents how they made their distributional choices and allowed them to describe their reasoning in an open-text field. Figure 5.7b visualizes the text analysis in a word cloud, highlighting the frequency of observed terms. The word

[11]By controlling for initial labor market earnings, we ensure that we measure the desired reward of all other characteristics, holding initial labor market earnings constant. For instance, a positive coefficient on education implies that conditional on initial market earnings, higher-educated people are considered more deserving than lower-educated people. In a full roll-out of our factorial survey, initial labor market earnings would be uncorrelated with all other vignette characteristics and we could replicate our analyses excluding initial labor market earnings. However, due to the methodological focus of our data collection, we operate with a limited number of vignettes where the characteristics of interest are correlated with each other. To account for such correlations, we include initial labor market earnings in all our estimations.

cloud shows that respondents put a strong emphasis on working hours, earnings, and the education of the vignette persons when making their distributional choices. The emphasis on these characteristics, therefore, echoes the results from our quantitative analysis.

In this section, we have shown that our survey module recovers fairness preferences that are broadly consistent with the existing literature. This conclusion holds for the degree of inequality acceptance, the prevalence of fairness types, and the characteristics determining the extent of fair income allocations. Since the data collection was designed for methodological validation, we urge readers to treat these substantive results cautiously. However, the results point to the ability of our survey module to uncover nuanced fairness positions and to describe fairness preferences in societies more broadly.

5.6 CONCLUSION

This study validates a novel survey tool designed to measure fairness preferences using realistic yet hypothetical scenarios.

We conduct this validation using a two-wave survey covering a representative sample of the US population. Our results demonstrate that fairness preferences are not influenced by the prospect of real-world implementation, even when monetary stakes are high. This conclusion holds true for both the general population and across various demographic subgroups. Moreover, comparing individual responses across the two waves reveals that fairness preferences are stable over time, regardless of whether they originate from hypothetical or incentivized scenarios. We furthermore provide suggestive evidence that the elicited preferences are consistent with established findings on fairness preferences in the US.

Therefore, our validation provides compelling evidence that fairness preferences from hypothetical surveys are not "just cheap talk." Instead, they can yield credible insights into the nature and anatomy of these preferences. We emphasize that these conclusions are context-dependent. Therefore, we currently plan additional validation exercises to show that our findings extend beyond the context of *WEIRD* countries like the US.[12]

[12]Henrich et al. (2010) show that behavior in experiments varies substantially across cultural contexts and that evidence from Western, Educated, Industrial, Rich, and Democratic (WEIRD) countries can hardly be extrapolated to other settings.

Chapter 6

Work-Life Balance (Policies) — Evidence from Germany

6.1 INTRODUCTION

In the past two decades, the realization of work-life balance (WLB)—the striving for a personal balance between market work, leisure activities, and family obligations—has become a widely discussed topic. For instance, web search analyses suggest that the interest in "work-life balance" more than doubled between 2005 and 2020, reflecting its growing relevance in public discourse (see Figure 6.1). At the same time, international institutions and policy initiatives have also prioritized the topic, for instance through the work-life balance directive of the European Union (2019) or the OECD Better Life Index (OECD, 2020).

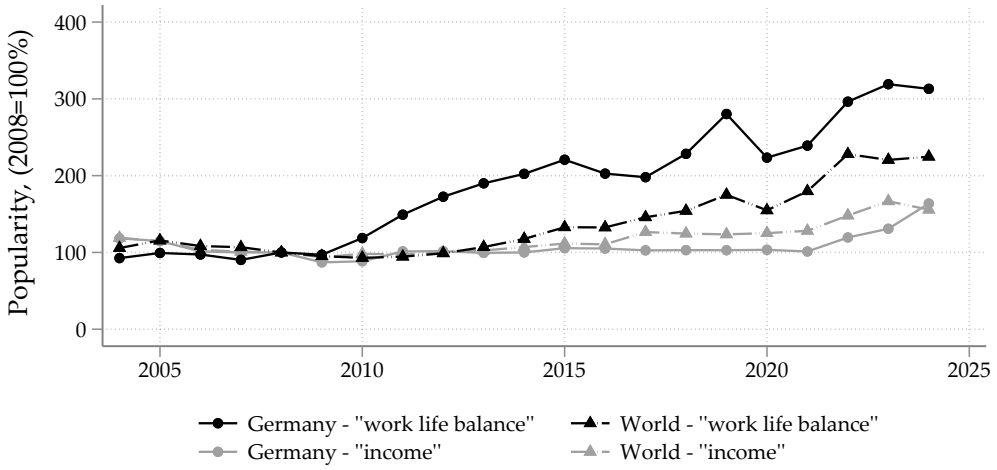
Many of the policies designed to foster WLB make it easier for individuals to reduce their working time, for instance by mandating employers to offer part-time positions or by giving employees the right to reduce their working time, holding their hourly wage constant. While proponents of WLB consider such policies central to improving individual welfare and to achieve more equitable labor market outcomes (OECD, 2020), others are worried that supporting part-time work comes at large fiscal costs that undermine governments' capacity to provide public goods and redistribute economic resources to the poor—especially in the face of labor shortages driven by demographic change.

Motivated by these trade-offs, this paper sheds light on work-life balance (policies) in the German context and is structured along three parts: (i) a descriptive documentation of the state and evolution of WLB; (ii) a causal analysis of a recent WLB policy expansion; and (iii) a stylized framework that makes the costs and benefits of such policies transparent and provides useful sufficient statistics to evaluate the welfare consequences of work-life balance policy reforms.

Work-Life Balance in Germany. In the first part of the analysis, I leverage rich data from the German microcensus and a special version of German social security records that contains administrative information on working hours to present stylized facts about work-life balance in Germany.

The results show that the desire for work-life balance has increased substantially over the last decade. For instance, the share of employees expressing a desire to work less almost doubled between 2010 and 2019. In addition, the desire for work-life balance is highly asymmetric. While high skilled employees express strong desires to work less, low skilled individuals rather strive for an expansion in hours worked, even if they work full-time. This mismatch between desired and actual hours of work accounts for a large share of

Figure 6.1. – Popularity of "Work Life Balance", Google Trends



Note: This figure displays the normalized popularity of the search topic "work-life balance" in Germany and the whole world. Popularity indices on google trends reflect, how often a particular term was searched at a particular point in time in comparison to all searches. To provide an additional level of comparison, the graph also includes the search topic "income". The time series for both terms have been normalized to 100 percent in 2008.
Source: Own calculations based on google trends.

the part-time employment gradient across the skill distribution. Another indication of the growing importance of work-life balance is the shift in the nature of part-time employment, which has—according to stated reasons for part-time work—increasingly moved away from rather involuntary forms of part-time work, e.g., due to a lack of full-time positions or health constraints. However, part-time employment still comes at significant costs at the individual level due to the convexity of earnings in labor hours, i.e., the fact that hourly wages increase with hours worked. In particular, conditional on a fine-grained set of observable characteristics, I find that an increase in working hours by 1 percent is associated with a decrease of hourly wages of around 0.17 percent. The extent of earnings convexity is related to the potential benefit of work-life balance policies, since these policies linearize the earnings schedule by giving employees the right to switch to part-time employment at their current hourly wage.

The Effect of WLB Policies. In a second step, I use administrative linked employer-employee data from social security records to analyze the effect of a work-life balance policy expansion in Germany (*Brückenteilzeitgesetz*). The policy was introduced in 2019 and allows employees to temporarily reduce their working time without requiring the consent of their employer for a period between one and five years. For identification, I exploit the fact that individuals are only eligible for the policy if they work in firms with

specific characteristics that are related to the number employees. To estimate the effects of the policy, I compare the within-individual evolution of the part-time employment probability between employees in eligible and non-eligible firms.

The analysis reveals that the policy had a moderate impact on part-time employment. In particular, the policy increased the part-time employment probability in eligible firms by around 0.4 percentage points relative to individuals in non-eligible firms. This effect is almost exclusively driven by individuals switching from full-time to part-time employment within a given firm. Furthermore, while the policy increased the probability to switch from full-time to part-time employment by around 10 percent, the policy did not have a significant effect on changes from unemployment to part-time employment. Also, in contrast to the objective of the policy, I find no evidence of higher returns from part-time to full-time employment in the aftermath of the policy.

Interpreting Work-Life Balance Policies. In the final part of the paper, I develop a stylized model that interprets WLB policies as a public insurance program against uninsurable employment risks. In the model, individuals face both unemployment and part-time employment risks. While I remain generally agnostic about the sources of these risks, part-time employment risks are consistent with the notion that certain shocks in life make full-time employment impossible or prohibitively costly, for instance, due to time constraints related to child or elderly care provision. By giving individuals the option to switch to part-time employment without having to face a cut in hourly wages, WLB policies provide additional consumption smoothing, but also lead to moral hazard by making part-time employment generally more attractive.

The theoretical analysis provides two insights. First, the local welfare effect of an increase in part-time employment benefits—the WLB policy—depends on the relative magnitude of the insurance value of additional consumption smoothing, and the change in fiscal costs associated with an increase in the part-time employment share. Second, the analysis shows that work-life balance policies and unemployment insurance programs are closely intertwined. In particular, an increase in the part-time employment benefit also makes part-time employment more attractive compared to unemployment, which induces some unemployed individuals to take up part-time work. Under specific circumstances, this effect represents a counteracting beneficial force on fiscal resources that affects the welfare assessment of a WLB policy expansion.

I estimate the sufficient statistics from the model using the empirical estimates from the descriptive and causal analyses in the first two parts of the paper. In particular, I quantify

the benefit of the reform as the consumption smoothing value implied by eliminating the status quo convexity of the earnings schedule. The ingredients on the cost side of the reform rely on the estimated behavioral responses to the WLB policy. According to this quantification, I find that the 2019 WLB policy expansion was locally welfare-improving. While the magnitude of the welfare effect is sensitive to variations in the empirical choices, the sign remains positive for most plausible ranges of parameter values.

Related Literature and Contribution. This paper relates to three strands of the literature. First, the descriptive part on part-time work and part-time hourly wage penalties relates to previous work that estimates the relationship between hours and wages (Bick et al., 2022; Del Rey et al., 2022; Hirsch, 2005; E. Wolf, 2002). The convexity in earnings is also discussed in relation to gender equality. For instance, Goldin (2014) and Goldin and Katz (2011) argue that a large part of the remaining gender wage gap can be explained by the structure of labor markets that facilitate convex returns to hours worked. In contrast to previous work, I estimate hourly wage penalties for Germany based on detailed hours information in administrative social security records that usually do not contain hours information. My descriptive analysis is further not centered around the heterogeneity by gender, but rather focuses on the heterogeneity of part-time work and the desire for work-life balance across the income and skill distribution.

By estimating the effect of a recent part-time entitlement program in Germany, I also relate to literature that focuses on understanding the impact of family-friendly policies in general, and in particular on labor market outcomes. While there exists a large literature discussing the effect of parental leave programs and public childcare availability (Olivetti and Petrongolo, 2017), there is only limited evidence regarding the impact of part-time entitlement programs. There are two notable exceptions. First, Blau and Kahn (2013) show in cross-country regressions that the decline in the labor force participation rates in the United States compared to other advanced nations between 1990 and 2010 can be explained to a large extent by the expansion of paid parental leave, public childcare, and part-time work rights in other advanced economies. However, while part-time entitlement policies are part of the analysis, the authors do not differentiate between different types of policies. Fernández-Kranz and Rodríguez-Planas (2021) analyze the effect of a reform in Spain in the 1990s that gave individuals with children of certain age the right to reduce their working time. They find that the policy led to unintended consequences, since employers are subsequently less likely to hire women in childbearing age. My analysis of the recent WLB policy expansion focuses on estimating the behavioral responses to the reform that are relevant for a general welfare assessment.

The welfare assessment in the final part of the paper is built upon a model of public insurance programs in the presence of uninsurable employment risks, for which private insurance markets do not exist. Thereby, my paper strongly relates to the work by Baily (1978), Chetty (2006b), and Chetty and Finkelstein (2013). The main innovation is that employment risks and policy measures are multi-dimensional, since individuals face both the risk to become unemployed and the risk to become part-time employed. In addition, the linearization of the earnings schedule through WLB policies introduces an additional policy instrument next to unemployment benefits.

Outline. The remainder of the paper is organized as follows. Section 6.2 provides a brief overview of the data sources employed in this paper. Section 6.3 presents stylized facts about part-time work and the desire for work-life balance in the German context. Section 6.4 estimates the behavioral responses of the work-life balance policy expansion in 2019. Section 6.5 introduces a stylized framework that is used to assess the welfare effects of the work-life balance policy expansion. Section 6.6 concludes and discusses avenues for future research.

6.2 DATA

For the empirical analysis, I use three different data sources that leverage the benefits of detailed demographic information of large cross-sectional survey data (Microcensus), administrative data on the employment composition of firms (IAB Establishment Panel), and longitudinal information on individual employment decisions that can be linked to establishment information (Sample of Integrated Employment Biographies).

Microcensus. The German microcensus (*Mikrozensus*) is a representative household survey of the German population that has been conducted since 1957 in regular intervals as part of the official government statistics. It represents a cross-section of around one percent of the German population, i.e, around 810,000 individuals in roughly 370,000 households, and covers a broad set of topics ranging from demographics, educational qualifications, and labor market participation to general life circumstances. The analyses in this paper are based on the annual scientific use files (SUF) from 1991 until 2020, each of which represents a seventy percent sub-sample of the original data in the respective year.[1] To facilitate long-run analyses, I harmonize the annual cross-sectional data set, taking methodological changes in variable definitions over time into account. All data

[1]For details on the microcensus in general and the scientific use files in particular, see the information provided by the Research Data Centres of the Statistical Offices of the Federation and the Federal States.

are weighted using survey weights calibrated to census statistics. I use data from the microcensus to describe the long-run evolution of work-life balance in Section 6.3 and for supplementary insights into the effect of work-life balance policies in Section 6.4. Appendix Table F.6 displays selected summary statistics. The benefit of microcensus data compared to employment data from social security records is that it also covers non-employed individuals and civil servants, and contains detailed demographic information, e.g., on marital status and household composition. In addition, the microcensus includes questions about preferred working hours and about the reasons for part-time work.

IAB Establishment History Panel (BHP). The annual establishment history panel (*Betriebs-Historik-Panel, BHP*) provided by the Institute for Employment Research (*Institut für Arbeitsmarkt- und Berufsforschung, IAB*) is a 50 percent random sample of all establishments in which employees subject to social insurance contributions are registered. The cross-sectional data set is available from 1975 for West Germany and from 1992 for East Germany and aggregates employee-level information from social security notifications as of June 30 of each year. In particular, it contains information on the industry, the employment and wage structure, as well as worker inflows and outflows of the respective establishment.[2] Appendix Table F.7 displays selected summary statistics that already provide some useful information on the prevalence and distribution of part-time work at the establishment level. For instance, the average part-time share in an establishment lies at 68 percent, which is primarily driven by establishments with a low number of employees, since the establishment size distribution is right-skewed. Small establishments are most active in trade while larger establishments are dominated by the manufacturing industry. In this paper, the main purpose of the BHP data is to extract establishment-level information on work force composition and link this information to individual data on employment biographies.

Sample of Integrated Employment Biographies (SIAB). The sample of integrated employment biographies (*Stichprobe der Integrierten Erwerbsbiographien*, SIAB) is a two percent random sample of the integrated employment biographies that are held by the Institute for Employment Research (IAB) in a respective year. It contains detailed information on employment and unemployment spells that are based on administrative records from social security notifications and unemployment benefit receipt. Among others, it captures information on the daily earnings in the registered job, qualification, occupational

[2]For details on BHP data, see the information provided by the Institute for Employment Research, and in particular Ganzer et al. (2022).

information, and an indicator for part-time employment.[3] The SIAB data can be linked to establishment level information from the BHP above, in particular to the information on firm size and work force composition. I follow the data preparation of vom Berge and Schmucker (2023) to transform the spell-level data set into an individual-level data set. I will use this data to estimate the effects of work-life balance policies in Section 6.4. Appendix Table F.8 displays summary statistics of SIAB data for the year 2020. At the individual level, consistent with the variation of the part-time share across the establishment size distribution, the average part-time share lies at around 34 percent.

While employee-level information on working hours is usually not part of SIAB data, hours information can be integrated into the regular SIAB data for five years between 2010 and 2014. The supplementary hours data is based on employer's reporting to the German Social Accident Insurances (*Deutsche Gesetzliche Unfallversicherung, DGUV*) and has been harmonized by vom Berge et al. (2023).[4] The harmonized hours variable accounts for differential reporting procedures by employers at different points in time and broadly reflects the contractual weekly working hours of an employee. Appendix Table F.9 displays selected summary statistics for the data that includes working hours. The average weekly working hours in 2014 is estimated at around 39 (21) hours for full-time (part-time) employed individuals. The five-year SIAB sample including hours information is used to describe the extent of hourly wage penalties by estimating the degree of earnings convexity in Section 6.3.

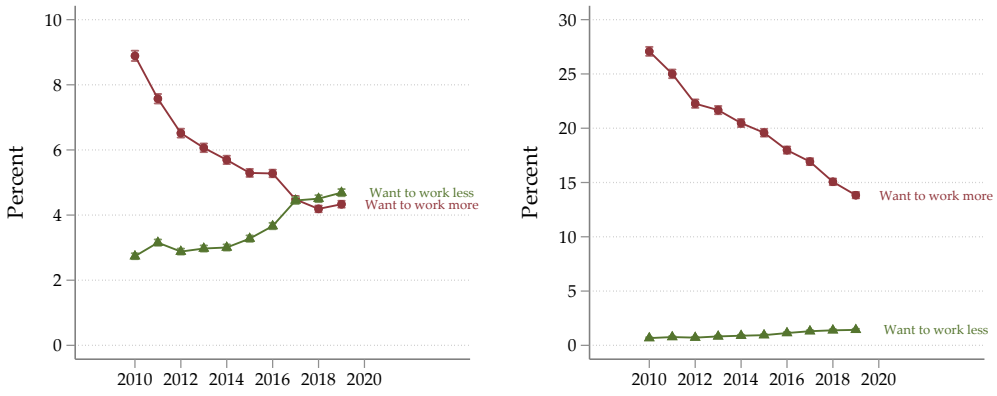
6.3 STYLIZED FACTS

In the following section, I describe four key aspects about the state of work-life balance in Germany, in particular about the nature of part-time and full-time employment. The first three key insights are based on data from the German microcensus and relate to (i) the evolution of (desired) work-life balance over time, (ii) the ability gradient of part-time work, and (iii) the reasons behind part-time employment. The fourth key insight is

[3]For details on SIAB data, see the information provided by the Institute for Employment Research, and in particular Schmucker et al. (2023).

[4]The exclusive availability of individual hours information between 2010 and 2014 can be explained by changes in reporting requirements in the German social insurance system. In particular, since DGUV contributions are exclusively paid for by the employer, employers were for a long time only required to report aggregate establishment level information on the number of hours worked. In 2009, employer's reporting to the German Social Accident Insurance was integrated into the regular employee-level reporting procedure which required employers to report the working hours at the individual level. This integration was partially repealed in 2016 (vom Berge et al., 2023).

Figure 6.2. – Desired Change in Working Hours



(A) Full-Time (B) Part-Time

Note: This figure displays for each year between 2010 and 2019 the share of individuals who state a desire to increase (decrease) their regular weekly working hours considering a correspondingly higher (lower) salary. Results are shown separately for full-time and part-time employees. Part-time (PT) and full-time (FT) employment is based on normal working hours and a self-reported assessment (PT: 24 hours or less, FT: 37 hours and more, 25-36 hours: self-reported part-time status). The figure includes point estimates and associated 95 percent confidence intervals. The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

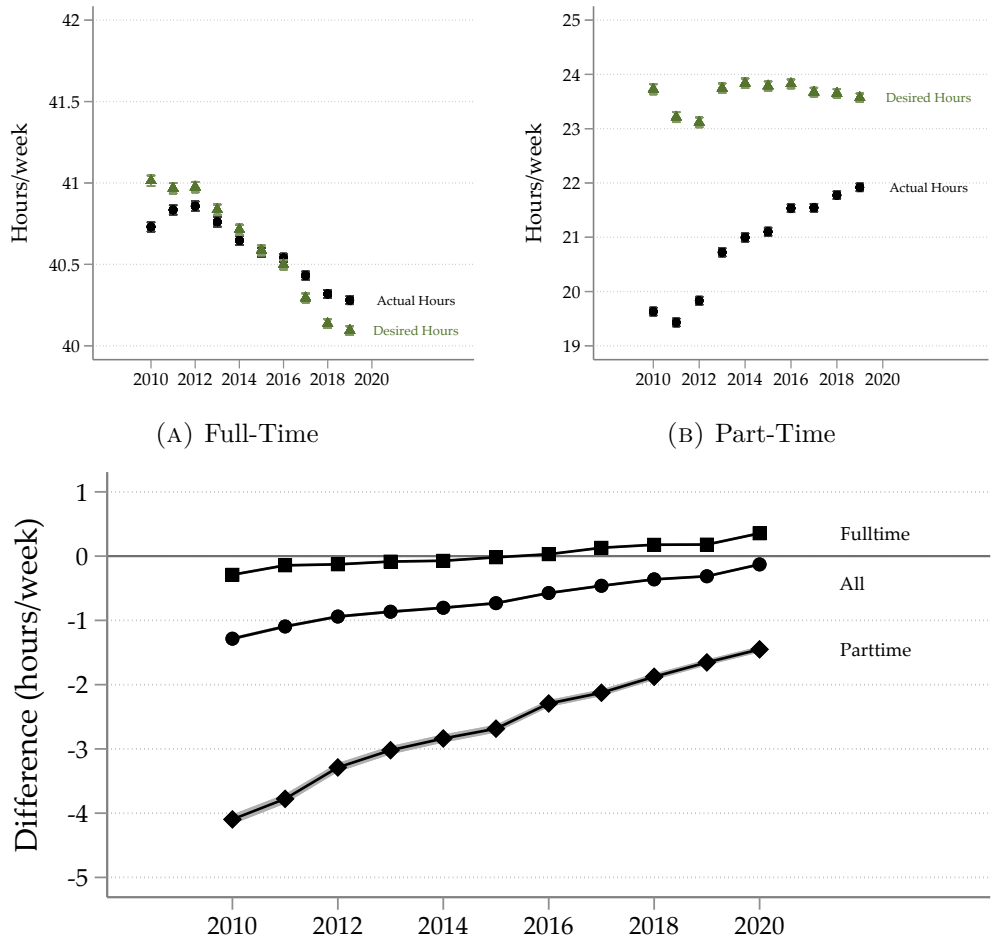
based on SIAB data and describes (iv) to what extent part-time employment is associated with hourly wage penalties. These stylized descriptive facts will motivate the analysis of work-life balance policies in the subsequent sections.

The Desire for Work-Life Balance Has Increased Over Time. In the microcensus, individuals have been asked since 2010 about their actual and desired working times. Importantly, individuals were asked about a desired reduction (expansion) in working times under correspondingly lower (higher) earnings. As shown by Figure 6.2a, the share of full-time employed individuals with a desire to work more has dropped from around nine percent in 2010 to around four percent in 2019. At the same time, the desire to work less has almost doubled over the same time period. Unsurprisingly, part-time workers on average show a higher desire to work more given that they already work lower hours. However, as shown by Figure 6.2b, the evolution over time closely mirrors the one for full-time workers. The pattern is also not driven by compositional changes, since it holds uniformly within demographic subgroups that differ in education, civil status, gender, and age (see Appendix Figures F.1-F.4).

The increasing desire for work-life balance is also revealed in the comparison between actual and desired weekly working hours (see Figure 6.3). In 2010, employed individuals worked on average more than one hour per week less than they would have liked to. Ten

years later, this gap almost completely vanished. While part-time employees still have a desire to expand their working time, full-time employees nowadays work on average more than they would like to. This is particularly remarkable, since the average actual working hours of full-time employed individuals already declined by around half an hour per week from 2010 until 2020. However, the average desired hours declined even more from around 41 hours to around 40 hours per week. This explains why the reduction in working hours was not sufficient to fully realize individuals' desired work-life balance and provides an explanation for why policy makers still perceive work-life balance policies as a necessary tool to assist individuals to realize their desired working times.

Figure 6.3. – Actual and Desired Working Hours



(A) Full-Time (B) Part-Time

(C) Difference

Note: This figure shows the actual and desired weekly working hours for all years between 2010 and 2019 for full-time (part-time employees) in Figure 6.3a (6.3b). The corresponding differences between actual and desired working times are shown in Figure 6.3c. Part-time (PT) and full-time (FT) employment is based on normal working hours and a self-reported assessment (PT: 24 hours or less, FT: 37 hours and more, 25-36 hours: self-reported part-time status). The figure includes point estimates and associated 95 percent confidence intervals. The figure focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

The Desire for Work-Life Balance Is Highly Asymmetric. While the desire for increased work-life balance holds across demographic groups, the level varies strongly across the skill distribution. In particular, the desire for work-life balance increases with the skill level, both when measured in terms of educational achievement or in terms of the net hourly wage of individuals.[5] For instance, among full-time employees with lower education, i.e., individuals with primary or incomplete secondary education, less than two percent prefer to work less considering correspondingly reduced earnings (see Figure 6.4). This share increases to almost ten percent for individuals with post-graduate university education. The ability gradient is also visible for individuals working part-time. For instance, in the top decile of the net hourly wage distribution, close to four percent of individuals demand to work less while the share is close to zero at the bottom of the wage distribution.[6]

The reverse side of the higher desire for lower hours at the top of the skill distribution is a higher degree of under-employment at the bottom. In particular, among part-time employed individuals with lower education, between 20 and 25 percent would like to work more. For individuals with higher education, this share is much lower and lies only at around 10 percent. The magnitude of over- and underemployment is substantial. On average, if individuals demand a reduction (increase) in working hours, they want to work around ten hours less (more) than they currently do (see Appendix Figure F.5).

The descriptive insight implies that the observed variation in working hours across the skill distribution is to some extent an indirect consequence of unrealized desired working times. As shown by Figure 6.5, the part-time share among employed individuals in 2019 is 40 percent at the bottom of the skill distribution and only 15 percent at the top. If individuals would realize their desired working time, e.g., if empowered by work-life balance policies, this gradient would be flattened substantially. However, since WLB policies usually only foster transitions from full-time to part-time employment, such policies might be particularly beneficial for high-skilled individuals, since large shares of low skilled individuals already work part-time and rather demand to work more.

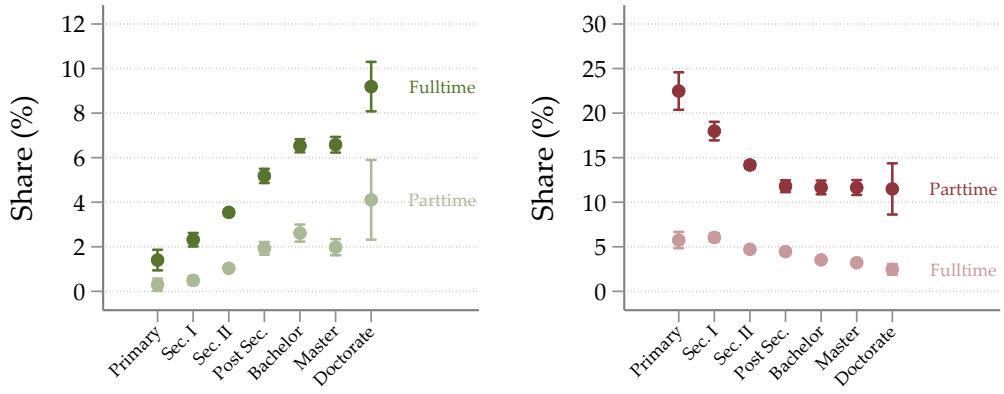
Part-Time Work Is Increasingly Voluntary. Another piece of evidence for the heterogeneous, but increasing attractiveness of work-life balance is the shift in the nature of part-time work from involuntary to voluntary forms of part-time employment. This

[5]While it would be preferable to analyze gross hourly wages, microcensus data only contains net earnings information. However, assuming a monotonic relationship between gross and net earnings, net hourly wage deciles plausibly reflect gross hourly wage deciles.

[6]In the Appendix, I show that this pattern is not driven by the age composition of individuals with different education, since the ability gradient exists within different age groups (see Appendix Figure F.4)

Figure 6.4. – Desire for Work-Life Balance across Ability (2019)

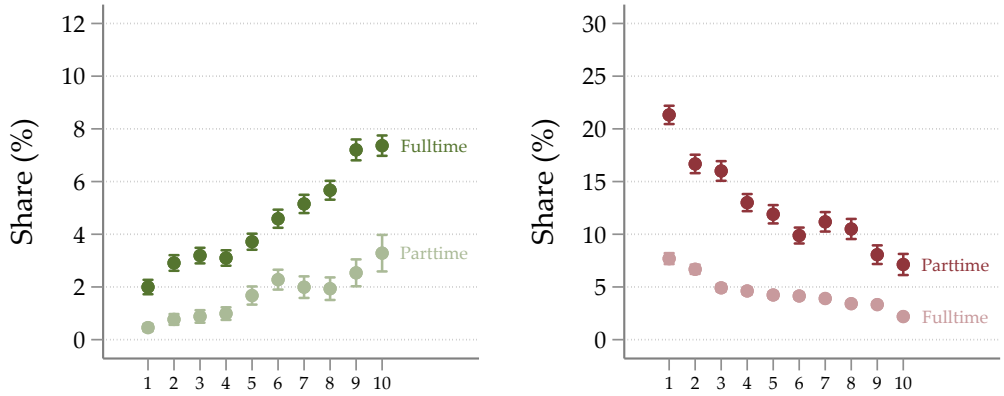
Education



(A) Desire to work less

(B) Desire to work more

Hourly Wage Deciles

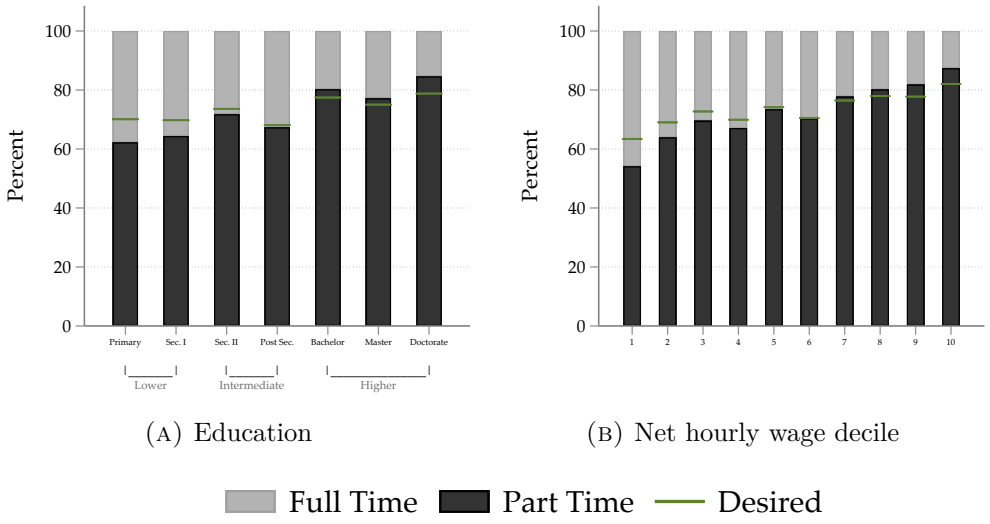


(C) Desire to work less

(D) Desire to work more

Note: This figure displays for 2019, how the share of individuals who state a desire to increase (decrease) their regular weekly working hours considering a correspondingly higher (lower) salary varies with ability and by part-time employment status. The figure shows point estimates and associated 95 percent confidence intervals. Figures 6.4a and 6.4b show the desire for work-life balance across different levels of education. Education information has been harmonized based on the International Standard Classification of Education (ISCED-2011). Figures 6.4c and 6.4d show the desire for work-life balance across net hourly wage deciles. The net hourly wage is estimated based on total individual net earnings information and the normal weekly working hours of individuals. To further account for measurement error in both earnings and hours, the actual net hourly wage is a predicted value based on a set of observable characteristics, i.e. age, education, labor force status, and occupational information based on the International Standard Classification of Occupations (ISCO-08). Part-time (PT) and full-time (FT) employment is based on normal working hours and a self-reported assessment (PT: 24 hours or less, FT: 37 hours and more, 25-36 hours: self-reported part-time status). The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

Figure 6.5. – (Preferred) Part-Time Share across Ability (2019)



(A) Education (B) Net hourly wage decile

Note: This figure shows the part-time share across different skill levels measured by education (Figure 6.5a) and net hourly wage deciles (Figure 6.5b). For each skill level, the graph also shows the part-time share in accordance with individuals desired working time as green horizontal line. The net hourly wage is estimated based on total individual net earnings information and the normal weekly working hours of individuals. To further account for measurement error in both earnings and hours, the actual net hourly wage is a predicted value based on a set of observable characteristics, i.e. age, education, labor force status, and occupational information based on the International Standard Classification of Occupations (ISCO-08). Part-time (PT) and full-time (FT) employment is based on normal working hours and a self-reported assessment (PT: 24 hours or less, FT: 37 hours and more, 25-36 hours: self-reported part-time status). The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

shift can be observed by analyzing the reasons for part-time employment recorded in the microcensus. Involuntary part-time work can defined as those types of part-time employment that are rationalized by external restrictions such as the unavailability of full-time employment options, illnesses and disabilities. In contrast, voluntary part-time work is characterized by the absence of such restrictions. In some cases, e.g., in case of part-time employment due to childcare, the distinction between voluntary and involuntary part-time employment is blurred. The reported reason to work part-time "because of childcare" can then be driven by the unavailability of care options (external constraints) or the genuine desire to spend more time with children (absence of external constraints). For the most recent years, the data allows to explicitly distinguish these cases.

Fifteen years ago, the unavailability of full-time employment options accounted for around one quarter of all part-time employment. As of 2020, this figure of involuntary part-time work now lies below ten percent (see Figure 6.6). While childcare and family obligations are still among the primary reasons for why individuals work part-time, there was a notable increase in the share of individuals that work part-time for "other reasons". These

other reasons can be interpreted as voluntary forms of part-time employment, since they are not associated with any restriction related to illnesses, disabilities, care for elderly or dependent persons, childcare, or other family obligations.

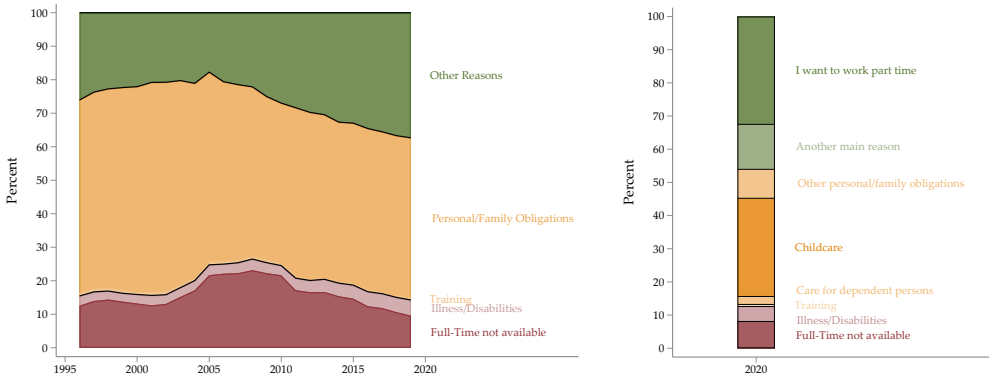
Figure 6.6b illustrates a more fine-grained separation of reasons for the most recent year. As of 2020, around 50 percent of part-time employees do not mention any restriction that is related to their decision to work part-time. While care-related part-time employment accounts for another 40 percent of part-time employment, the vast majority of this part-time employment is not driven by external constraints as highlighted by Figure 6.7. For instance, among those individuals, for which childcare is the main reason of part-time employment, more than 70 percent are not restricted by the unavailability of adequate, affordable or close care facilities. The share is similar for other care-related part-time employment. Taking the (absence of) such restrictions into account further validates the observation that part-time employment as of today can be considered primarily voluntary.

However, similar to the desire for work-life balance, the voluntary degree of part-time employment is not uniformly distributed across the skill distribution. At the top of the skill distribution, 80 percent of part-time employment can be considered voluntary. At the bottom, however, this figure drops to less than 60 percent (see Figure 6.8). The reasons for part-time employment also vary strongly across other dimensions of heterogeneity. For instance, while women report primarily family-related reasons for part-time employment, other voluntary reasons seem to be particularly prevalent among men and older individuals (see Appendix Figures F.7-F.8). These results highlight that work-life balance policies might well assist with the reconciliation of employment and family obligations, but that the policies can also facilitate part-time employment that is not associated with external constraints.

Part-Time Employment Is Linked to Lower Hourly Wages. The increasing desire for work-life balance in the form of voluntarily reduced working hours in exchange for correspondingly reduced earnings raises the question what prevents individuals to change their working time. One potential barrier is the fact that part-time employment does not only come with *proportionally* reduced earnings, but also depresses the hourly wage, both in the cross-section and conditional on observable characteristics (Bick et al., 2022).

To estimate, how hourly wages vary with working hours in Germany, I use administrative data from the Sample of Integrated Employment Biographies (SIAB) between 2010 and 2014. The core variables of interest for the analysis are the employment status, weekly working hours, and weekly earnings. To further limit the impact of misreporting, I exclude

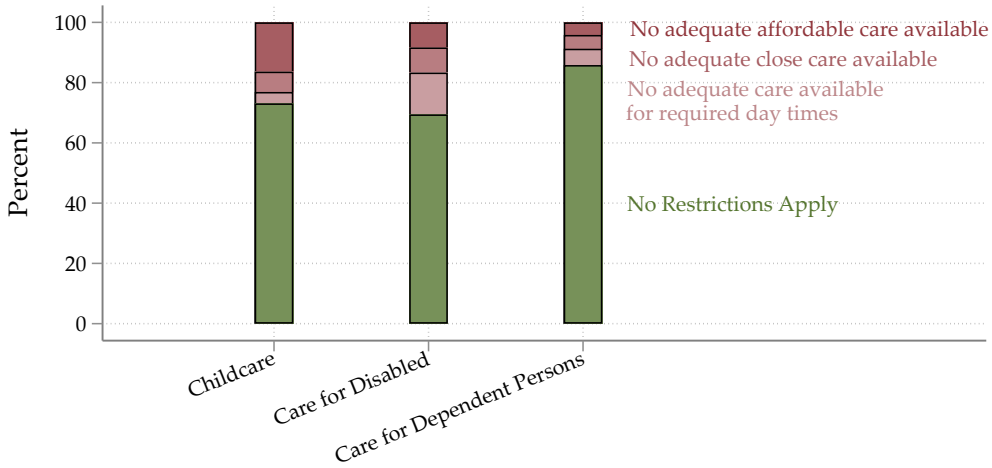
Figure 6.6. – Reasons for Part-Time Employment



(A) 1996 - 2019 (B) 2020

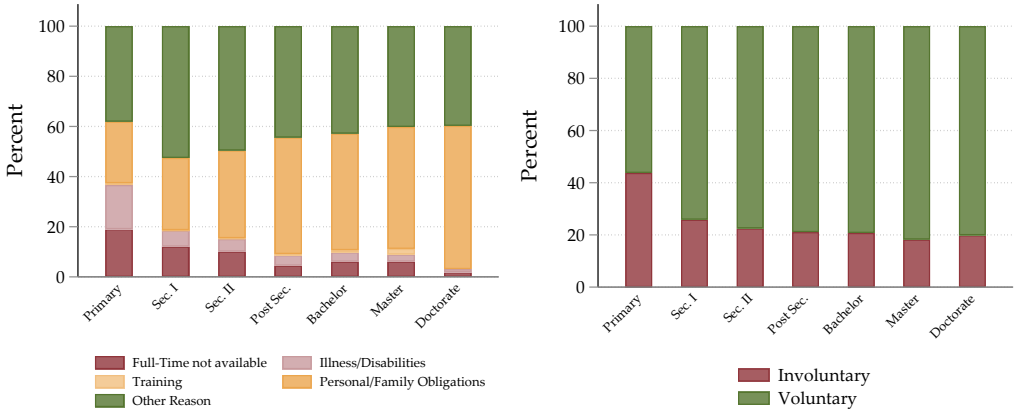
Note: This figure shows how the distribution of reasons for part-time work changed over time for those individuals that are employed part-time in each respective year. Figure 6.6a displays the long run evolution using a harmonized version of the stated reason while Figure 6.6b provides a more detailed breakdown that is only possible for the most recent year. Both figures use a color scheme that classifies reasons for part-time work according to the degree of restrictions faced by individuals, where red (green) colors indicate stronger (weaker) restrictions. The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

Figure 6.7. – Care-Related Part-Time Work, Restrictions (2020)



Note: This figure displays for different types of care-related part-time employment (childcare, care for people with disabilities, and care for dependent persons), whether individuals are restricted by external constraints in their working time decision. In particular, involuntary care-related part-time work (red area) is defined as part-time work driven by external restrictions related to the unavailability of adequate care facilities. It is explicitly differentiated between the unavailability of *affordable* and *close* care facilities, as well as the unavailability of care facilities that offer the required day times. Voluntary care-related part-time work (green area) is defined as care-related part-time work that is not driven by any of these restrictions. The analysis focuses on part-time employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

Figure 6.8. – Reasons for Part-Time Work by Education (2020)



(A) Original (B) Reclassification

Note: This figure displays how the distribution of reasons for part-time work varies across education for those individuals that are employed part-time in 2020. Figure 6.8a displays the detailed breakdown of stated reasons for part-time employment. Figure 6.8b classifies all reasons into voluntary and involuntary components. Childcare, elderly care and family obligations are classified as voluntary if individual stated that neither missing, affordable, nor close care facilities prevent them from working full-time (see details in Figure 6.7). The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

extreme hours values below (above) p1 (p99) from the analysis. The earnings variable of interest is based on the sum of all income from employment notifications, i.e. the sum of daily gross wages subject to social insurance contributions.

Equipped with data on earnings and hours, I assume that the hourly wage follows the functional form

$$w = \theta \frac{l^{\beta}}{1+\beta}, \tag{6.1}$$

where θ is some measure of skill and β governs the degree of earnings convexity, i.e., how strong the hourly wage varies with working hours.[7] Based on this functional form, one could naively estimate the descriptive relation between the hourly wage and weekly working hours as

$$ln(w_{it}) = \alpha + \beta ln(h_{it}) + \epsilon_{it} \tag{6.2}$$

[7]Appendix Figure F.9 illustrates the corresponding difference between a linear and a convex earnings schedule under the respective functional form assumption.

where w_{it} is the hourly wage of individual i at time t and h_{it} are the respective working hours. The coefficient of interest β describes, how the hourly wage varies with working hours using both cross-sectional and inter-temporal variation of working hours and hourly wages. However, this reduced-form estimation suffers from division bias if working hours are measured with noise (Borjas, 1980). In particular, since the hourly wage variable is itself a function of the noisy working hours variable, i.e., since $w = \frac{y}{h}$, the coefficient of interest is biased towards minus one. To address division bias, I instrument the noisy measure of weekly working hours with the part-time employment status p_{it}. Since this alternative measure of working hours is measured independently from working hours, this eliminates the division bias and yields an unbiased estimate of β. The corresponding first and second stage regression equations read

$$ln(h_{it}) = \alpha + \delta p_{it} + \mu_t + \mathbf{X}_{it} + \epsilon^1_{it}, \tag{6.3}$$

$$ln(w_{it}) = \alpha + \beta \widehat{ln(h_{it})} + \mu_t + \mathbf{X}_{it} + \epsilon^2_{it}. \tag{6.4}$$

Across all specifications, I control for time fixed effects denoted by μ_t. In addition, depending on whether one is interested in the part-time wage penalty in the raw cross-section or for individuals of similar characteristics, I also control for a vector of (interacted) observable characteristics $\mathbf{X}_{it}$ containing information about experience, occupation, education, industry, and the state of residence.

Table 6.1 summarizes the results of the analysis. In the raw cross-section, the second stage coefficient of 0.56 indicates that the hourly wage increases by 0.56 percent for a one percent increase in working hours. Controlling for a very fine-grained set of individual characteristics, this measure declines substantially, but still lies at 0.17 for the most rigorous specification that only compares individuals within the same narrow cell defined via congruent experience, occupation, education, industry and state. The measure of 0.17 implies that a part-time employee with a 50 (75) percent contract has approximately an 11 (5) percent lower hourly wage relative to a comparable full-time employee. This suggests that the earnings schedule is convex and that part-time employment is indeed linked to lower hourly wages.[8]

[8] Appendix Table F.10 provides supplementary heterogeneity analyses and shows that men face higher hourly wage penalties, whereas the degree of earnings convexity seems relatively constant over time and across different education levels.

Table 6.1. – Hourly Wages and Working Hours

Type of Controls	δ (s.e.)	β (s.e.)	C.-D. F-Stat.	N_T (N_S)
Year FE (5)	-0.70 (0.001)	0.56 (0.002)	1.32e06	1,777,497 (0)
Year FE + Sep. Controls (641)	-0.66 (0.001)	0.25 (0.001)	885,592	1,777,497 (3)
Year FE + Int. Controls (240,622)	**-0.60 (0.001)**	**0.17 (0.002)**	**326,096**	**1,777,497 (742,578)**
Year FE + Ind. FE (425,944)	-0.45 (0.001)	-0.01 (0.003)	187,923	1,777,497 (159,555)
Year FE + Ind. FE, Between Est. (114,991)	-0.27 (0.001)	0.04 (0.004)	100,430	422,105 (2,153)

Note: This table displays, how the hourly wage rate varies with working hours. The analysis is based on equations (6.3) and (6.4). The table displays five different specifications that differ in the type of controls used in the analysis (see column 1). The first specification only controls for year fixed effects. The second specification includes additional categorical information on experience (9 categories), occupation (339 categories), education (7 categories), industry (272 categories), and state (16 categories). The third specification includes fully interacted controls. The fourth specification controls for year and individual fixed effects. The fifth specification replicates the fourth specification, but only includes changes in part-time employment associated with individuals changing the establishment. For each specification, the table shows the types of controls and number of estimated fixed effect coefficients in column 1. Column 2 (3) displays the first (second) stage coefficient δ (β) and respective standard errors. Column 4 displays the Cragg-Donald Wald F-Statistic providing information about the strength of the instrument in the first stage. Column 5 displays the total number of observations N and the number of singleton observations N_S that are dropped in the respective specification. Highlighted gray cells refer to estimates used in the welfare analysis in Section 6.5.4 (see also Table 6.6).
Source: Own calculations based on SIAB (2010-2014).

Importantly, note that the hourly wage penalty vanishes, if one only uses within-individual changes in part-time employment for the identification of earnings convexity (see fourth specification in Table 6.1). However, while this specification eliminates all time-constant heterogeneity between part-time and full-time employees, it is mainly based on employees that are *allowed* to change their working time within the same establishment, and therefore suffers from selection bias. This is confirmed by an analysis of within-individual changes in part-time employment based on across-establishment switches, where earnings convexity prevails (see last specification in Table 6.1). The divergence between aggregate hourly wage penalties and within-individual hourly wage penalties provides a rationale for why policy makers may want to facilitate employment status switches within firms. These policies force employers to accept individual working hours changes and at the same time prevent a downward adjustment in hourly wages and thereby linearize the convex earnings schedule.

The descriptive section has shown that the past decade in Germany was characterized by a shift towards lower working times and an increasing prominence of voluntary forms of part-time employment. Still, the desire for more work-life balance persists, especially among high-skilled individuals who increasingly demand a reduction in working times.

Also, the empirical analysis reveals that part-time employment is not only associated with proportionally lower earnings, but also comes with lower hourly wages. This hourly wage penalty could be a barrier that prevents individuals to switch from full-time to part-time employment. For instance, individuals might not be willing to accept an hourly wage cut in exchange for lower working hours. Also, since employers might not be allowed to adjust wages downward for those individuals that switch from full-time to part-time employment, employees might not even have the possibility to reduce their working hours in the same firm—even if they would accept lower hourly wages in return.

6.4 POLICY ANALYSIS

To facilitate changes from full-time to part-time employment (and vice-versa), many countries around the world have enacted work-life balance policies that make it easier for individuals to change their working time at a given employer without having to face the negative consequences of reduced hourly wages. After giving a broad overview over such policies, I estimate the behavioral responses to such policies for a recent expansion of work-life balance policies in Germany.

In general WLB policies that provide an entitlement to part-time work can be classified along their *type* and their *length*. First, entitlements to part-time work can be of a general or specific type. While general entitlements allow employees to reduce their working time in any situation, specific entitlements are targeted to cover specific circumstances, e.g., during parental leave, elderly care, partial retirement, or in case of chronic illnesses. Second, entitlements can be of non-temporary or temporary nature. Non-temporary entitlements allow permanent reductions in working time and leave a return to full-time employment at the discretion of the employer. In contrast, temporary entitlements provide the opportunity to reduce working hours for a limited amount of time, allowing the employee to return to their original working hours afterwards. Table 6.2 provides an overview of different policies and their classification in selected countries. In general, specific temporary entitlements are most prevalent. However, as in the case of Germany, specific and general entitlements might also coexist.[9]

[9]As shown in Appendix Table F.1, even though some countries do not have direct regulation that mandates part-time employment rights, such rights can be part of the negotiations between trade unions and employers.

Table 6.2. – Work-Life Balance Policies in Selected Countries

Country	Description
Austria	Specific & temporary entitlements (chronic illness, parenthood).
France	Specific & temporary entitlements (family reasons, parental leave, ill relatives, creation of a company, arduous job).
Germany	General temporary (permanent) entitlement to work part-time since 2019 (2001). Specific temporary entitlements (partial retirement - 1990, parental leave - 2007, home care - 2008/2015).
Netherlands	General entitlement to work part-time.
United Kingdom	Specific (childcare, elderly care) and general entitlements.

Note: This table summarizes WLB policies in selected countries that provide employees with an entitlement to part-time work, i.e., that allow them to reduce their working time without requiring the consent of their employer. The table describes the nature of part-time entitlement policies, in particular whether they are of temporary or non-temporary nature and under which circumstance the entitlement applies.
Source: See detailed overview in Appendix Table F.1.

6.4.1 Institutional Background in Germany

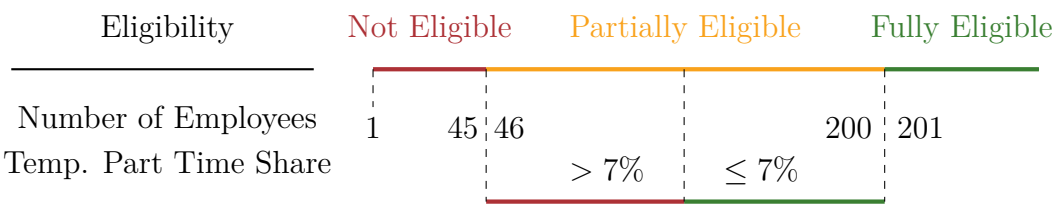
In Germany, work-life balance policies have been expanded since the 2000s. In 2001, Germany introduced a general entitlement to part-time work that allowed individuals to choose a reduction in working hours, but left an eventual future increase in working hours at the discretion of the employer. This general entitlement was supplemented by specific entitlements during parental leave (2007) and during situations of home care provision (2008, 2015).[10]

2019 Policy Reform. In the subsequent analyses, I focus on the most recent policy expansion in 2019. This policy expansion under the *Brückenteilzeitgesetz* gave employees a general and temporary entitlement to reduce their working time. In particular, eligible employees can request to reduce their working time to a desired amount for a specific period between one and five years. After this pre-specified period has passed, employees automatically go back to their original working time. The policy thereby had the explicit rationale to facilitate temporary transitions to part-time employment while preventing a lock-in effect.

Eligibility to the temporary part-time work entitlement is a combination of individual and firm level determinants. First, employees are eligible for the policy if they worked in the respective firm for at least six months. More importantly, they are only eligible if they work in a firm with a certain employment size. Firms with up to 45 employees were not affected by the policy, while firms with more than 200 employees are fully affected. Small firms were excluded from the policy, because it was acknowledged that it might be

[10]For a detailed overview over the regulations in Germany over time, see Appendix Table F.3

Figure 6.9. – Temporary Part-Time Entitlement Policy (2019), Eligibility



Note: This figure visualizes the eligibility for the German part-time entitlement policy expansion in 2019 (*Brückenteilzeitgesetz*) in terms of the number of employees and the part-time share in the firm. For details, see Appendix Table F.4.
Source: Regulations of the *Brückenteilzeitgesetz*.

difficult for them to flexibly accommodate these entitlements without harming business activity. In medium-sized firms with between 46 and 200 employees, special restrictions apply. In particular, the entitlement to temporary part-time is only granted if the share of employees working in temporary part-time is below approximately seven percent.[11] Figure 6.9 visualizes the eligibility along these lines.

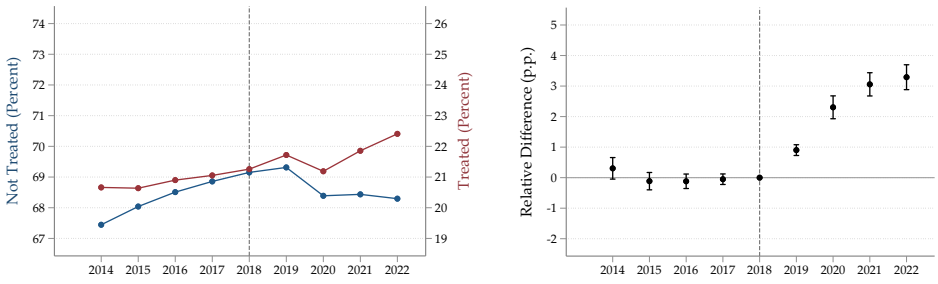
Cross-Sectional Evidence. To take a first descriptive look at the potential effect of the policy, I focus on establishment-level data from the IAB establishment history panel (BHP) and describe how the part-time share in affected firms evolved compared to unaffected firms.[12] Figure 6.10a shows that both treated and untreated firms were on an increasing trend of part-time employment, although at a different level. In treated firms, the average part-time share in 2018 was at around 21 percent. At the same time, the average part-time share in non-treated firms was at around 69 percent. Figure 6.10b shows the corresponding difference in the part-time share between the two groups controlling for year and eligibility-group fixed-effects. Since I am interested in how the aggregate part-time share in the economy evolved differentially, each firm observation is weighted by total firm level employment the pre-reform period.[13] As shown by the figure, a divergence between treated and untreated firms in terms of the part-time share opens up

[11] More precisely, employers in medium-sized firms can decline the request if there is already one employee working in temporary part-time for each 15 employees. For details, see Appendix Table F.4.

[12] It is important to note that the empirical classification entails some difficulties. While the policy operates at the firm level, the BHP is an establishment level data set. I assign treatment status based on the establishment size and composition, since firms are not observable in the data. Any misclassification would yield a lower bound estimate, since establishments would be falsely allocated to the control group in case a small ineligible establishment is part of a firm that is large enough to be actually eligible. In the remainder of the paper, I use the terms firms and establishments interchangeably.

[13] This weighting allows for an individual-level interpretation of treatment effects when using firm level data (see, e.g., Risch, 2024).

Figure 6.10. – Work-Life Balance Policy and Part-Time Share



(A) Unweighted Part-Time Share (B) Weighted Difference

Note: This figure shows the evolution of the raw part-time share in treated and untreated establishments (Figure 6.10a) and the estimated difference when controlling for year and treatment group fixed effects (Figure 6.10b). Treatment status is determined based on the employment size and composition at the establishment level (see Figure 6.9). In both panels, 2018, i.e., the year before the policy came into effect, is marked by a dashed vertical line. Coefficients in the right panel can be interpreted as the percentage point difference of the part-time share between the treatment and control group. Coefficients are plotted against the year before the introduction of the policy. 95 percent confidence intervals are based on robust standard errors.
Source: Own calculations based on BHP (2014-2022).

after the policy was introduced. In particular, the part-time share in treated firms in the medium run is around three percentage points higher compared to the share in unaffected firms.[14] While the aggregate cross-sectional evolution of the part-time share already hints at some effects of the policy, this evidence cannot be interpreted as causal, since the composition of firms and individuals possibly changed independent from the policy over time. To take compositional changes of firms and individuals into account, I thus now turn to an estimation of the causal effect at the individual level.

6.4.2 Identification and Estimation

The effect of the policy can operate along different channels. First, employed individuals in eligible firms may decide to switch from full-time to part-time employment while staying in their respective firm. This is the intended effect of the policy. Second, employed individuals in non-eligible firms may decide to switch from full-time to part-time *and* switch to an eligible firm which is now more attractive due to the policy. Third, previously unemployed individuals may increasingly start to work part-time in eligible firms, if the policy made part-time employment generally more prevalent in eligible firms.

[14]This descriptive pattern can also be observed in other data sources containing information about employment size of firms or establishments, e.g., see analysis based on Microcensus in Appendix Figure F.10.

Identification. To estimate the policy effect across the three margins, I employ an employee level analysis where I compare within-individual changes in the probability to work part-time before and after the policy change between employees that work in firms of different eligibility. I thereby differentiate between *within-establishment changes* from full-time to part-time employment and *between-establishment changes* from full-time to part-time employment, as well as changes from unemployment to part-time employment. The key identifying assumption is one of parallel trends. In the absence of the policy, the employment outcomes of individuals who (start to) work in different types of firms, would have continued in parallel. While this assumption is not directly testable, the dynamic estimation of the conditional outcome differences as an event study can provide some plausibility test. In supplementary analyses, I show that (i) results are robust to the inclusion of differential time trends of firm and individual level characteristics that are correlated with eligibility, (ii) endogenous changes in eligibility are not driving the effect, and (iii) an establishment-level analysis yields similar conclusions as the baseline analysis at the employee-level.

Estimation. I analyze the reform from an employee perspective based on a dynamic differences-in-differences (DiD) specification

$$Y_{it} = \sum_{t \neq 2018} \beta^t d_t \times E_{it} + \gamma_i + \mu_t + \epsilon_{it} \tag{6.5}$$

where Y_{it} are different outcomes of individual i in year t, most importantly (i) a part-time indicator that takes the value of 100 if the individual works part-time, and (ii) a switching indicator that takes the value 100 if the individual switches from full-time to part-time in a given period. I measure the dynamic effect of the treatment by interacting time dummies d_t with the eligibility indicator E based on the employment size and composition in the respective year. Time-constant individual-level differences in the probability to work part-time (to switch) are controlled for by individual fixed effects γ_i while μ_t captures year specific effects common to all employees. $\epsilon_{i,t}$ is the idiosyncratic error term.

I estimate the effect for the time period between 2014 and 2021 to obtain dynamic estimates for four years prior to the policy and three years after the policy.[15] Since I am primarily interested in the labor supply choice of regular employees, I focus on individuals in their prime working age between 25 and 64 who work at least six months in the

[15] Any years prior to 2014 are not used due to a major revision of the part-time indicator in 2012 that impedes the quality of the part-time outcome variable (Fitzenberger and Seidlitz, 2020).

establishment. Trainees, interns, family members in agriculture, retired persons, as well as individuals in voluntary service are excluded. Treatment assignment is based on the establishment size and the establishment level part-time share according to the thresholds displayed in Figure 6.9.[16]

6.4.3 Main Results

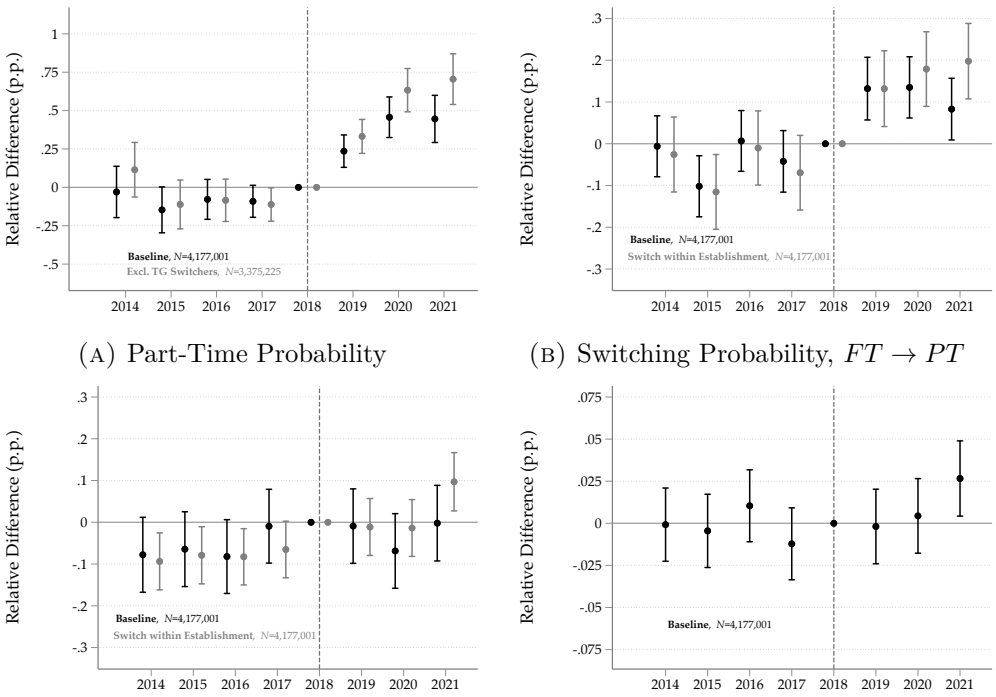
Figure 6.11 shows the main results from the event-study specification for different outcome variables of interest. The corresponding difference-in-difference point estimates are shown in Table 6.3. Before 2019, there was no statistically significant differential trend in the probability to work part-time between individuals in eligible and non-eligible firms. After the policy is introduced, the probability to work part-time increases by around 0.4 percentage points in the medium run (see Figure 6.11a). This increase is driven by the timely increase in the probability to switch from full-time to part-time employment by around 0.2 percentage points (see Figure 6.11b). Compared to an annual baseline switching probability of around 1.4 percent (see Appendix Table F.8), this reflects a relative increase of around 14 percent.

Figure 6.11a also provides a first verification that selection into treatment, i.e., individuals switching to establishments of certain composition or establishments manipulating the eligibility by restricting employment size, does not drive the results. In particular, the effect is even stronger if one focuses exclusively on those employees that do not change the eligibility status over the whole time period between 2014 and 2021 (gray estimates). Figure 6.11b further shows that the effect persists if one focuses exclusively on changes from full-time to part-time employment *within* the same establishment (gray estimates). This is in line with the mechanism of the policy that gave individuals the right to temporarily change the working hours at their current employer.

The results confirm that the policy allowed some individuals to switch into part-time employment that previously weren't able or willing to do so. Since the policy had the explicit objective to provide individuals with an option to only *temporarily* reduce their working time, it is also of interest whether the policy managed to achieve higher return rates, i.e, that more individuals switched from full-time to part-time employment after the policy. However, at least until 2021, there is no evidence of such a pattern, since the within-individual changes in the switching probability from part-time to full-time employment do not differ between individuals in eligible and non-eligible establishments

[16]Since I don't have information on the temporary part-time share at the firm level, I use the general part-time share at the establishment level as a proxy for this unobserved variable.

Figure 6.11. – Effect of Work-Life Balance Policy



(A) Part-Time Probability

(B) Switching Probability, $FT \rightarrow PT$

(C) Switching Probability, $PT \rightarrow FT$

(D) Switching Probability, $U \rightarrow PT$

Note: This figure shows the dynamic difference-in-difference estimation of the policy effect based on equation (6.5). Treatment assignment is based on the establishment size and the part-time share of the contemporaneous establishment of an employee according to the relevant thresholds displayed in Figure 6.9. Figure 6.11a displays results using a part-time indicator (0,100) as outcome variable of interest and distinguishes between estimates for the full baseline sample (black) and the sample of employees that do never change treatment status (gray). Figure 6.11b (6.11c) displays results using a switching indicator (0,100) from full-time (part-time) to part-time (full-time) employment as outcome variable of interest, i.e., $FT \rightarrow PT$ ($PT \rightarrow FT$), and distinguishes between estimates using all switches (black) and estimates using only within-establishment switches (gray). Figure 6.11d displays results using a switching indicator (0,100) from unemployment to part-time employment, i.e. $U \rightarrow PT$. All analyses focus on individuals in their prime working age between 25 and 64 who work at least six months in the establishment. Trainees, interns, family members in agriculture, retired persons, as well as individuals in voluntary service are excluded. Standard errors are clustered at the level of employees.
Source: Own calculations based on SIAB (2014-2021).

(see Figure 6.11c). In addition, the policy also did not incentivize individuals to take up (part-time) employment, since the probability to switch from unemployment to part-time employment does not change differentially between individuals starting to work in eligible and non-eligible firms (see Figure 6.11d).

6.4.4 Robustness and Heterogeneity

Additional Controls. Despite the absence in pre-trends in the event study analysis, one might be concerned that differences in pre-reform characteristics between employees working in eligible and non-eligible firms are driving the effect. With reference to the descriptive statistics in Appendix Tables F.7 and F.8, one could be particularly concerned

Table 6.3. – Effect of Work-Life Balance Policy, Difference-in-Difference Estimates

	(1) PT	(2) $FT \rightarrow PT$	(3) $PT \rightarrow FT$	(4) $U \rightarrow PT$	(5) $ln(y)$
Post × Elig.	0.440	0.210	0.016	0.011	-0.016
	(0.061)	(0.026)	(0.025)	(0.006)	(0.001)
	0.000	0.000	0.514	0.097	0.000
Year FE	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes
Obs. (million)	4.177	4.177	4.177	4.177	4.154

Note: This table shows difference-in-difference estimates for the baseline analysis using the part-time employment indicator (PT, column 1), and the switching indicator from full-time to part-time employment ($FT \rightarrow PT$, column 2), from part-time to full-time employment ($PT \rightarrow FT$, column 3), and from unemployment to part-time employment ($U \rightarrow PT$, column 4). Column 5 displays results for log earnings as outcome variable. All specifications include individual and year fixed effects. For each specification, the table shows the difference-in-difference coefficient, standard errors clustered at the individual level in brackets, as well as associated p-values. Highlighted gray cells refer to estimates used in the welfare analysis in Section 6.5.4 (see also Table 6.6).
Source: Own calculations based on SIAB (2014-2021).

about differential industry and age composition. For instance, it could be the case that industry specific shocks in 2019 affected employees in (non-)eligible establishments differentially, which in turn could explain the diverging evolution of the part-time employment probability starting in 2019. However, as shown in column 3 of Table 6.4, the effect persists when controlling for industry specific time trends, i.e., when I only focus on outcome deviations from the respective industry-specific time trend. Also, differences in the age composition don't change the effect size substantially as shown by column 4, where I control for birth decade specific time trends. This might be explained by the fact that the baseline sample already harmonizes the age of individuals to some degree. Expanding the baseline sample with individuals of age 25-64 to the full sample—illustrated in column 2—indeed reduces the magnitude of the effect, plausibly because individuals at the beginning and end of their working life are less likely to react to the reform, e.g., due to concerns about dynamic career prospects, or because they are already close to retirement.

Selection into Treatment. In the baseline analysis, I use the contemporaneous establishment size and composition for the treatment assignment. This has the advantage to incorporate responses that are driven by changes in eligibility at the establishment level, e.g., due to firm size growth. However, if firms react to the policy by manipulating the eligibility threshold or if individuals switched systematically to treated establishments, the analysis would be confounded. While I already showed that excluding individuals facing varying eligibility status does not alter the results (see Figure 6.11a), one can also address this concern more directly by fixing the treatment assignment in the pre-reform year, or by instrumenting the contemporaneous treatment with the pre-reform assignment (see,

Table 6.4. – Effect of Work-Life Balance Policy, Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Alt. Sample	Add. Controls		Alt. Spec.	
					E_{2018}	IV (E_{2018})
Post × Elig.	0.440	0.175	0.302	0.452	0.520	0.509
	(0.061)	(0.061)	(0.065)	(0.061)	(0.061)	(0.072)
	0.000	0.004	0.000	0.000	0.000	0.000
Year FE	Yes	Yes	-	-	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year	-	-	Yes	-	-	-
Birth × Year	-	-	-	Yes	-	-
Obs. (million)	4.177	4.568	4.177	4.177	3.602	3.602

Note: This table shows difference-in-difference estimates for the baseline outcome variable of interest, i.e., the part-time employment indicator. Column 1 shows the baseline estimate. Column 2 replicates the analysis without any sample restriction on age. Column 3 (4) add additional controls, in particular industry (birth decade) specific time trends. Columns 5 and 6 present results from two alternative specifications. In the first alternative specification displayed in column 5, eligibility is not allowed to vary over the years, but held fix in the year before the policy change (E_{2018}). In the second alternative specification displayed in column 6, eligibility in a specific year is instrumented with the eligibility in the pre-reform year. For each specification, the table shows the difference-in-difference coefficient, standard errors clustered at the individual level in brackets, as well as associated p-values. Highlighted gray cells refer to estimates used in the welfare analysis in Section 6.5.4 (see also Table 6.6).
Source: Own calculations based on SIAB (2014-2021).

e.g., Risch, 2024). Columns 5 and 6 in Table 6.4 show the corresponding results based on these alternative specifications. As expected based on the previous analyses, the effects are of similar magnitude as the baseline estimate.

Establishment-Level Analysis. The baseline analysis uses employees as the unit of observation, since reactions to the policy occur at this level, and because the data structure then allows to control for time-constant unobserved heterogeneity at the individual level. A disadvantage of this approach is that it is not possible to control for time-constant unobserved heterogeneity of certain establishments.[17] In Appendix F.2, I thus estimate the effect of the policy at the establishment level using BHP data. The magnitude of the effect is broadly in line with the baseline results, indicating that the policy increased the part-time share among eligible establishments by 0.5-1 percentage points when focusing on a comparison of the within-establishment changes between eligible and ineligible establishments.

Heterogeneity. As indicated by the previous analyses, the policy had some moderate average effect on the probability to be employed part-time. Motivated by Section 6.3

[17] As discussed in Section 6.2, SIAB data is a two percent random subsample of *individual* employment biographies. Therefore, there are only few individuals that are observed working in the same establishment. This makes it infeasible to control for establishment fixed effects in the baseline analysis.

Table 6.5. – Effect of Work-Life Balance Policy, Heterogeneity

	Sex		Education	
	Men	Women	Lower Educ.	Higher Educ.
Post × Eligible	0.764	0.571	0.407	0.436
	(0.068)	(0.110)	(0.078)	(0.130)
	0.000	0.000	0.000	0.001
Year FE	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Mean (2018)	12.69	56.26	32.62	27.38
Obs. (million)	2.170	2.007	2.023	0.983

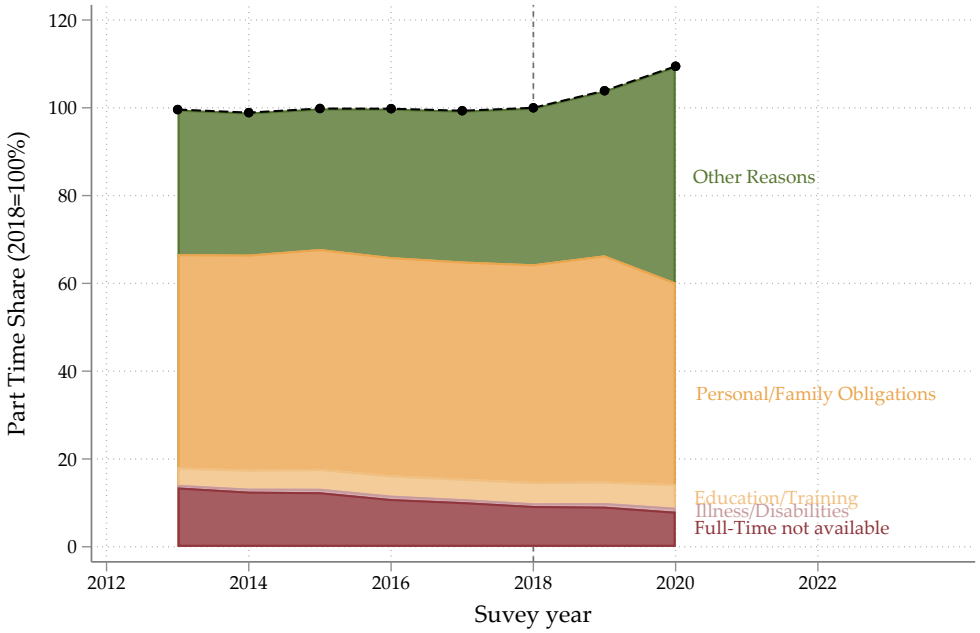
Note: This table shows difference-in-difference estimates for the baseline outcome variable of interest, i.e., the part-time employment indicator, for different sub-samples of the population. The table distinguishes between the effect of the policy among men and women in columns 2 and 3, and among different levels of education in columns 4 and 5. Lower education refers to all individuals with education up to university entrance certificate, but without vocational education. Higher education refers to all individuals with a university entrance certificate and a vocational degree or university education. All specifications include individual and year fixed effects. For each specification, the table shows the difference-in-difference coefficient, standard errors clustered at the individual level in brackets, as well as associated p-values.
Source: Own calculations based on SIAB (2014-2021).

which revealed substantial differences in the desirability for work-life balance, I further estimate the effect separately for men and women, and across different education levels. As shown by Table 6.5, the absolute increase in the part-time employment probability is higher for men than for women, both in absolute terms and relative to the baseline part-time employment share. This is consistent with the fact that men face higher hourly wage penalties compared to women and thus benefit more from the policy (see Appendix Table F.10). Also, the effect is slightly larger in absolute and relative terms for individuals with higher education. This is in line with the descriptive pattern observed in Section 6.3 that revealed a stronger unrealized desire for lower working hours at the top of the skill distribution.

Unfortunately, administrative SIAB data does not allow us to distinguish *why* individuals switched to part-time employment as a response to the policy, e.g., whether it was driven by individuals with a requirement to spend more time with their children, increasing care for elderly relatives, or whether it was rather related to a genuine increase in leisure-related activities. However, by looking at data from the German microcensus, one can provide some suggestive answers to this question. In line with the estimated treatment effect, Figure 6.12 shows that the probability to work part-time increased substantially after the introduction of the policy among individuals in treated establishments.[18] The division into the different reasons for part-time employment highlights that this increase is not associated with childcare, family obligations or involuntary forms of part-time

[18] Appendix Figure F.10 shows that this relative increase is not visible for individuals working in establishments that are not affected by the policy.

Figure 6.12. – Work-Life Balance Policy and the Nature of Part-Time Employment



Note: This figure displays the normalized part-time share for individuals that are employed part-time in establishments with 50 employees and more. Since Microcensus data only contains coarse categories of the number of employees in the local establishment, this employment size cut-off is based on the closest possible cut to the threshold relevant for the policy (see Figure 6.9). The part-time share is normalized to 100 percent in 2018, the year before the policy came into effect. The figure further shows the distribution of stated reasons of part-time employment. The analysis focuses on part-time employed individuals with positive earnings that are not in education and are between 25 and 64 years old. Civil servants are excluded.
Source: Own calculations based on Mikrozensus (1996-2020).

employment. In contrast, the increase of part-time employment seems to be associated with exclusively voluntary forms of part-time employment. In addition, this effect seems particularly pronounced for men (see Appendix Figure F.11), confirming the heterogeneity analysis from the causal analysis based on SIAB data.

6.5 WELFARE EFFECTS

In the previous sections, I showed that part-time employment often comes with hourly wage penalties through convex earnings schedules (Section 6.3) and that work-life balance policies eliminating the convexity—by giving individuals the right to switch to part-time at their current hourly wage—affects individual employment choices (Section 6.4). In this section, I introduce a stylized framework that brings these elements together and allows for a welfare assessment of the work-life balance policy expansion. I first describe the stylized model before discussing the sufficient statistics governing the trade-offs of the

policy expansion. Finally, I bring these sufficient statistics to the data and provide a back-on-the-envelope calculation of the welfare effects of the German WLB policy expansion in 2019. According to this calculation, the WLB policy expansion was welfare-improving for plausible ranges of parameter choices.

6.5.1 Model

I discuss the welfare effects of work-life balance policies from the perspective of a representative agent facing un-insurable employment risks. In particular, similar to the risk to become unemployed, there are exogenous shocks associated with the risk to become part-time employed, e.g., due to exogenous time constraints from child care, the sudden illness of an elderly relative requiring intensive care, or a preference shock that makes full-time employment prohibitively costly. Work-life balance policies are a tool to insure individuals against these adverse shocks that force them to be part-time employed. At the same time, as for the risk to become unemployed, individuals can exert effort to avoid this risk, e.g., by moving closer to childcare facilities, searching an elderly care provider, or by switching to another full-time job that provides increased flexibility. Therefore, as for any insurance policy, a classical trade-off between insurance and moral hazard emerges.

Static Representation. I follow the approach by Baily (1978), Chetty (2006b), and Chetty and Finkelstein (2013) and discuss the dynamic nature of this problem and associated trade-offs in a simplified static model. The main innovation is that agents no longer only face the risk to become unemployed. In contrast, individuals face two types of employment risk. First, individuals can lose their job and become unemployed (U). Second, individuals can receive a shock that makes full-time employment (F) impossible and only allows them to work part-time (P). The time spent in each of the three states is denoted by e_F, e_P and $e_U = 1 - e_T$ respectively, with $e_T = e_F + e_P$ representing the total time spent in either full-time or part-time employment.

Individuals. Spending time in full-time and part-time employment comes with increasing and convex effort costs $\psi_F(e_F)$ and $\psi_P(e_P)$ respectively, i.e., $\psi'_F(e_F) > 0, \psi''_F(e_F) > 0, \psi'_P(e_P) > 0, \psi''_P(e_P) > 0$. Under both full-time and part-time employment, individuals pay taxes τ_U that finance benefits in unemployment b_U. Under full-time employment, individuals in addition pay taxes τ_P that finance benefits under part-time employment. Individuals cannot save and borrow and thus live hand-to-mouth. We write short hand $b = (b_F, b_P)$ and $e = (e_F, e_P)$. The utility, income, and consumption levels of individuals in all three states read

$$\begin{aligned} &u_F(c_F), \;\; c_F = y_F - \tau_U(b_U) - \tau_P(b_P) \\ &u_P(c_P), \;\; c_P = y_P - \tau_U(b_U) + b_P \\ &u_U(c_U), \;\; c_U = y_U + b_U \end{aligned} \tag{6.6}$$

with $u'_F > 0, u''_F < 0, u'_P > 0, u''_P < 0, u'_U > 0, u''_U < 0$.[19]

Government. The government provides two forms of insurance that are financed through taxes. Taxes have an actuarially fair structure, according to which the unemployment benefits are financed through taxes on all employed individuals and part-time employment benefits are financed through taxes on all full-time employed individuals. The government thus faces two constraints given by

$$\begin{aligned} I_1 &= (e_F + e_P)\,\tau_U - (1 - e_F - e_P)\,b_U = 0, \\ I_2 &= e_F \tau_P - e_P b_P = 0. \end{aligned} \tag{6.7}$$

6.5.2 Welfare-Maximization — First Best Situation

In a hypothetical first best situation, the social planner chooses benefit levels and effort directly. Since there is no ex-ante heterogeneity between individuals, utilitarian welfare is simply a weighted sum of utilities in the different employment states net of the effort costs. Thus, the planner's maximization problem subject to the government constraints reads

$$\begin{aligned} \max_{e,b} \;\; & W(b,e) = e_F\, u_F(c_F) + e_P\, u_P(c_P) + (1 - e_F - e_P)\; u_U(c_U) - \psi_F(e_F) - \psi_P(e_P) \\ & \text{s.t. } \tau_U(b_U, e_F, e_P) = \frac{(1 - e_F - e_P)}{e_F + e_P} b_U \quad , \quad \tau_P(b_P, e_F, e_P) = \frac{e_P}{e_F} b_P. \end{aligned} \tag{6.8}$$

Any local optimum must satisfy the following first order conditions

[19] Note that both utility and effort cost functions are allowed to be state-dependent at this stage. For instance, utility functions could be different under full-time and part-time employment, if there are complementarities between leisure and consumption. Effort costs could be different in the two states if, for instance, the opportunity costs of part-time employment have a different structure, e.g., due to the presence of children in the part-time employment state.

$$
\begin{aligned}
&\mathbf{b_U}: \frac{e_F}{(e_F+e_P)}u'_F + \frac{e_P}{(e_F+e_P)}u'_P = u'_U,\\
&\mathbf{b_P}: u'_F = u'_P,\\
&\mathbf{e_F}: \psi'_F(e_F) = \Big(u_F(c_F) - u_U(c_U)\Big) + \frac{b_U}{(e_F+e_P)^2}\left[e_F u'_F(c_F) + e_P u'_P(c_P)\right] + b_P\frac{e_P}{e_F}u'_F(c_F),\\
&\mathbf{e_P}: \psi'_P(e_P) = \Big(u_P(c_P) - u_U(c_U)\Big) + \frac{b_U}{(e_F+e_P)^2}\left[e_F u'_F(c_F) + e_P u'_P(c_P)\right] - b_P u'_F(c_F).
\end{aligned}
\tag{6.9}
$$

Combining and rearranging theses conditions yields

$$
\begin{aligned}
&\mathbf{b}: u'_U = u'_P = u'_F,\\
&\mathbf{e}: \psi'_F(e_F) - \psi'_P(e_P) = \left[u_F(c_F) - u_P(c_P)\right] + u'_F b_P\left(1 + \frac{e_P}{e_F}\right).
\end{aligned}
\tag{6.10}
$$

Intuitively, the planner chooses benefit levels to achieve full insurance, i.e., the marginal utilities are equalized across states.[20] Also, the planner chooses effort levels in both employment states such that the difference in marginal effort costs in the two states equals the social benefit of differential effort in the two states. The social benefit consists of private benefits that are characterized by the utility difference between full-time and part-time employment, and the fiscal externality associated with higher full-time relative to part-time employment.[21]

6.5.3 Welfare-Maximization — Second Best Situation

Under the first best situation, the social planner can mandate effort levels directly. However, in reality, effort levels are choices by individuals that are out of reach from the direct influence of the government. As a consequence, the first best allocation cannot be achieved.

To discuss the welfare effects of increasing part-time insurance, I thus now turn to the second best situation, in which the government only chooses benefit levels and tax rates.

[20] I assume the welfare function is concave at the critical point, meaning that the critical point identified by the first-order conditions corresponds to a local maximum. A formal test, which involves an examination of the Hessian matrix is omitted at this stage.

[21] The private benefit would disappear, if utility functions were state-independent. In this case, the equalization of marginal utilities would imply that the utility values in the three states also coincide.

For this purpose, I first describe choices based on individual utility maximization and derive relevant elasticities from comparative statics at the optimum. Subsequently, I discuss the fiscal effects of changes in the benefit level drawing on these elasticities and on the government budget constraints that are imposed by the actuarially fair insurance. Equipped with these ingredients, I then derive sufficient statistics for the local welfare effects of changes in the part-time employment benefit, i.e., the work-life balance policy.

Individual Optimization. In the second best situation, individuals choose themselves of how much time to spend in each of the three states by maximizing individual utility $V(e_F, e_P)$. The corresponding individual optimization problem reads

$$\begin{aligned} \max_{e_F, e_P} V(e_F, e_T) &= e_F \; u_F\left(c_F\right) + e_P \; u_P\left(c_P\right) + \left(1 - e_F - e_P\right) \; u_U\left(c_U\right) - \psi_F\left(e_F\right) - \psi_P\left(e_P\right), \\ \text{s.t. } c_F &= y_F - \tau_U(b_U) - \tau_P(b_P), \quad c_P = y_P - \tau_U(b_U) + b_P, \quad c_U = y_U + b_U. \end{aligned} \tag{6.11}$$

The individual utility $V(e_F, e_P)$ coincides with the welfare function of equation (6.8), since there is no ex-ante heterogeneity across individuals.[22] The first order conditions (FOCs) of the individual optimization problem read

$$\begin{aligned} u_F - u_U &= \psi_F'\left(e_F\right), \\ u_P - u_U &= \psi_P'\left(e_P\right), \end{aligned} \tag{6.12}$$

which can be combined to

$$\psi_F'\left(e_F\right) - \psi_P'\left(e_P\right) = u_F - u_P. \tag{6.13}$$

In the individual optimum, the private benefit from providing differential effort in the two employment states must equal the difference in marginal effort costs.[23] Consequently, in contrast to the optimality condition under the first best situation in equation (6.10),

[22] While this is a simplification, the focus on a representative agent allows me to assess work-life balance policies in the absence of any re-distributive motive, e.g., based on heterogeneous ability.

[23] In Appendix F.3, I show that the critical point defined by the first order conditions of the individual optimization problem is indeed a local maximum.

individuals do not take into account that spending more time in full-time relative to part-time employment comes with a fiscal externality. The individual optimality conditions can be used to assess behavioral responses to changes in benefit levels.

Individual Optimization — Elasticities. Using the total differential of the first order conditions from the individual maximization problem allows me to separately analyze changes in the four policy variables, i.e., b_U, b_P, τ_U, τ_P on individual choices, i.e., the time spent in each of the states e_F, e_P.[24] Based on these comparative statics, I define the following elasticities that will be relevant for the welfare assessment:

1. Part-Time Share Elasticity w.r.t. b_P

$$\mathcal{E}_{\frac{e_P}{e_T},b_P} = \frac{\partial \frac{e_P}{e_T}}{\partial b_P} \frac{b_P}{\frac{e_P}{e_T}} = \frac{e_F}{e_T} \left[\frac{\partial e_P}{\partial b_P} \frac{b_P}{e_P} - \frac{\partial e_F}{\partial b_P} \frac{b_P}{e_F} \right] \geq 0 \tag{6.14}$$

2. Unemployment Elasticity w.r.t. b_P

$$\mathcal{E}_{1-e_T,b_P} = \frac{\partial 1 - e_T}{\partial b_P} \frac{b_P}{1-e_T} = -\frac{\partial e_F}{\partial b_P} \frac{b_P}{e_F} \frac{e_F}{1-e_T} - \frac{\partial e_P}{\partial b_P} \frac{b_P}{e_P} \frac{e_P}{1-e_T} \rightarrow \text{ambig. sign} \tag{6.15}$$

3. Unemployment Elasticity w.r.t. b_U

$$\mathcal{E}_{1-e_T,b_U} = \frac{\partial 1 - e_T}{\partial b_U} \frac{b_U}{1-e_T} = -\frac{\partial e_F}{\partial b_U} \frac{b_U}{e_F} \frac{e_F}{1-e_T} - \frac{\partial e_P}{\partial b_U} \frac{b_U}{e_P} \frac{e_P}{1-e_T} \geq 0 \tag{6.16}$$

4. Part-Time Share Elasticity w.r.t. b_U

$$\mathcal{E}_{\frac{e_P}{e_T},b_U} = \frac{\partial \frac{e_P}{e_T}}{\partial b_U} \frac{b_U}{\frac{e_P}{e_T}} = \frac{e_F}{e_T} \left[\frac{\partial e_P}{\partial b_U} \frac{b_U}{e_P} - \frac{\partial e_F}{\partial b_U} \frac{b_U}{e_F} \right] \rightarrow \text{ambig. sign} \tag{6.17}$$

The first two elasticities describe behavioral responses to changes in the part-time employment benefit b_P. In particular, an increase in b_P makes part-time employment more attractive and thus unambiguously increases the share of part-time employees (first elasticity). In addition, depending on whether the employment response among full-time or part-time employees is larger, the unemployment share can increase or decrease (second elasticity). The third and fourth elasticity describe the corresponding behavioral responses to changes in the unemployment benefit b_U. In particular, an increase in b_U makes both

[24] The partial derivatives are shown in Appendix F.3.

full-time and part-time employment less attractive, and thereby unambiguously increases the share of unemployed individuals (third elasticity). Whether the part-time employment share increases or decreases depends on how strong the reactions of full-time and part-time employees are in relative terms, i.e., intuitively if more part-time employees or more full-time employees switch to unemployment (fourth elasticity). These four elasticities are required to characterize the behavioral responses affecting the fiscal effects of benefit changes.[25]

Fiscal Effects of Benefit Changes. Using the constraints imposed by the actuarially fair insurance, i.e., equation (6.7), and taking the endogeneity of choices with respect to benefits via the elasticities above into account, one can express the fiscal costs of higher benefits through the required changes in tax rates.[26] Importantly, even though the two insurance policies are actuarially fair in their own domain, the analysis of fiscal effects reveals that there are spill-overs between the two types of insurance policies.

Fiscal Effects — Increasing Part-Time Employment Benefits Increasing consumption in the part-time employment state through changes in b_P requires changes in tax rates τ_P and τ_U to keep the government constraints satisfied. In particular, the required changes in tax rates are

$$\begin{aligned} \frac{\partial \tau_P}{\partial b_P} &= \frac{e_P}{e_F}\left(1 + \frac{e_T}{e_F}\,\mathcal{E}_{\frac{e_P}{e_T},b_P}\right) \geq 0, \\ \frac{\partial \tau_U}{\partial b_P} &= \frac{(1-e_T)}{e_T^2}\frac{b_U}{b_P}\mathcal{E}_{1-e_T,b_P} \rightarrow \text{ambiguous sign.} \end{aligned} \tag{6.18}$$

Intuitively, raising benefits under part-time employment comes with a higher tax rate τ_P, because all infra-marginal part-time employees now receive higher benefits. In addition to this mechanical effect, an increase in b_P also induces some individuals to switch to part-time employment. The strength of this behavioral effect is determined by the elasticity of the part-time share with respect to b_P.

[25] In principle, other combinations of elasticities are also possible. I choose these elasticities, since they are helpful to provide transparent sufficient statistics for the welfare effect of reforms, and since they are closely linked to the empirical estimates from Section 6.4.

[26] The derivations based on the total differential of government constraints are relegated to Appendix F.3.

An increase in b_P has also indirect effects on unemployment insurance. For instance, if some part of the increase in the part-time share is driven by changes in the share of unemployed individuals, this has a fiscal saving effect and thus relaxes the constraint imposed by the actuarially fair unemployment insurance. In this case, τ_U would decline as a response to an increase in b_P. A similar connection between the two insurance policies is also present when I evaluate the fiscal effects of changing unemployment benefits b_U.

Fiscal Effects — Increasing Unemployment Benefits Increasing consumption in unemployment through changes in b_U requires the following changes in tax rates:

$$\begin{aligned} \frac{\partial \tau_U}{\partial b_U} &= \frac{1-e_T}{e_T}\left(1+\frac{\mathcal{E}_{1-e_T,b_U}}{e_T}\right) \geq 0 \\ \frac{\partial \tau_P}{\partial b_U} &= \frac{b_P}{b_U}\,\frac{e_P}{e_F^2}\,e_T\,\mathcal{E}_{\frac{e_P}{e_T},b_U} \rightarrow \text{ambiguous sign} \end{aligned} \tag{6.19}$$

The required change in τ_U contains a mechanical effect and a behavioral effect. First, an increase in unemployment benefits requires more resources for all people that are already unemployed, i.e., the infra-marginal recipients of b_U. In addition, a higher benefit induces some individuals to switch to unemployment which requires an additional increase of the tax rate τ_U that is given by the elasticity of unemployment with respect to the benefit level.

Similar to the spillover from part-time insurance to unemployment insurance above, an increase in the unemployment benefit has also indirect effects on the part-time employment insurance, and therefore the tax rate τ_P. In particular, if the part-time share declines, i.e., if the induced increase in unemployment is mainly driven by part-time employees, this comes with some fiscal saving, since fewer individuals now receive the part-time employment benefit b_P. Correspondingly, the tax rate τ_P declines. In contrast, if the part-time share increases, i.e., if the increase in unemployment is mainly driven by full-time employees, this requires an increase in τ_P.

Welfare Effects of Changes in Benefits. Equipped with the information on how both individuals and government constraints react to changes in the benefit level, I now evaluate local welfare effects of changes in the part-time employment benefit. The second best maximization problem of the social planner reads

$$\begin{aligned}\max_{b_U,b_P} W\left(b,e\right) = {} & e_F\ u_F\left(y_F - \tau_U(e,b) - \tau_P(e,b)\right) \\ & + e_P\ u_P\left(y_P - \tau_U(e,b) + b_p\right) \\ & + \left(1 - e_F - e_P\right)\ u_U\left(y_U + b_U\right) \\ & - \psi\left(e_F\right) - \psi\left(e_P\right) \end{aligned} \tag{6.20}$$

$$\begin{aligned}&\text{s.t. } e_F = e_F^*(b_U,b_P) \quad , \quad e_P = e_P^*(b_U,b_P) \\ &\text{s.t. } \tau_U\left(b_U,e_F,e_P\right) = \frac{\left(1-e_F-e_P\right)}{e_F+e_P} b_U \quad , \quad \tau_P\left(b_P,e_F,e_P\right) = \frac{e_P}{e_F} b_P.\end{aligned}$$

Welfare Effects — Part-Time Employment Benefits Taking into account that individuals are already optimizing, the behavioral response to the change in the benefit level has no direct effect on utility (envelope condition), and only enters via the government budget constraint. A local change in the benefit level of part-time employment thus reads

$$\frac{\partial W}{\partial b_P} = e_P \left(u_P' - \frac{\left(e_F u_F' + e_P u_P'\right)}{e_T} \left[\frac{\left(1-e_T\right)}{e_T} \frac{b_U}{b_P} \frac{\mathcal{E}_{1-e_T,b_P}}{e_P} \right] - u_F' \left[1 + \frac{e_T}{e_F}\ \mathcal{E}_{\frac{e_P}{e_T},b_P} \right] \right).$$

For ease of exposition, one can normalize this welfare change of an increase in b_P by relating it to the welfare change of a budget neutral one dollar increase in the wage bill under full-time employment. Defining the average marginal utility under employment as $\overline{u_T'} = \frac{e_F u_F' + e_P u_P'}{e_T}$, the normalized welfare change reads

$$M_W(b_P) = \frac{\frac{1}{e_P}\frac{\partial W}{\partial b_P}}{\frac{1}{e_F}\frac{\partial W}{\partial y_F}} = \underbrace{\frac{\left(u_P' - u_F'\right)}{u_F'}}_{\text{Insurance Value}} - \underbrace{\frac{e_T}{e_F}\mathcal{E}_{\frac{e_P}{e_T},b_P}}_{\text{Direct Fiscal Costs}} - \underbrace{\frac{\overline{u_T'}}{u_F'}\frac{\left(1-e_T\right)}{e_T}\frac{b_U}{b_P}\frac{\mathcal{E}_{1-e_T,b_P}}{e_P}}_{\text{Spillover to Unemployment Insurance}} . \tag{6.21}$$

The marginal welfare effect of transferring resources from full-time to part-time employment consists of three parts. The first part represents the insurance value that is driven by the difference in marginal utility under full-time and part-time employment. The second part represents the direct fiscal costs of transferring resources from full-time to part-time employment that is driven by the increase in the part-time employment share. The third

part reflects spillovers to unemployment insurance that can counteract or reinforce the fiscal costs depending on how the share of unemployed individuals changes as a response to an increase in b_P.

Welfare Effects — Unemployment Benefits Similar to the increase in b_P, an increase in the benefit under unemployment b_U affects social welfare by

$$\frac{\partial W}{\partial b_U} = (1-e_T)\left\{u'_U - \left(1+\frac{\mathcal{E}_{1-e_T,b_U}}{e_T}\right)\frac{e_F u'_F + e_P u'_P}{e_T} - \frac{e_T}{1-e_T}\,\frac{e_P}{e_F}\,\frac{b_P}{b_U}\,u'_F\,\mathcal{E}_{\frac{e_P}{e_T},b_U}\right\} \tag{6.22}$$

If part-time employment was non-existent, i.e. under $e_F = e_T$ and $e_P = 0$, the equation reduces to the standard formula of Chetty (2008). We can also normalize this welfare gain by comparing it to a one dollar increase in the wage bill both under full-time and part-time employment, i.e.,

$$M_W(b_U) = \frac{\frac{1}{1-e_T}\frac{\partial W}{\partial b_U}}{\frac{1}{e_T}\frac{\partial W}{\partial y_F} + \frac{1}{e_T}\frac{\partial W}{\partial y_P}} = \underbrace{\frac{\left(u'_U - \overline{u'_T}\right)}{\overline{u'_T}}}_{\text{Insurance Value}} - \underbrace{\frac{\mathcal{E}_{1-e_T,b_U}}{e_T}}_{\text{Direct Fiscal Costs}} - \underbrace{\frac{u'_F}{\overline{u'_T}}\,\frac{e_T}{1-e_T}\,\frac{e_P}{e_F}\,\frac{b_P}{b_U}\,\mathcal{E}_{\frac{e_P}{e_T},b_U}}_{\text{Spillover to Part-Time Insurance}}. \tag{6.23}$$

The first part of the marginal welfare effect represents the insurance value of higher unemployment benefits, which is given by the difference between the marginal utility under unemployment and the average marginal utility under employment. It corresponds to the value of consumption smoothing coming with higher unemployment benefits. The second term represents the direct fiscal cost of transferring resources from full-time and part-time employment to unemployment. These fiscal costs are given by the increase of the share of unemployed individuals. The third part represents spillovers to the part-time insurance. This spillover is driven by the fact that changes in the share of unemployed individuals also potentially changes the distribution of employment states. This change can both counteract or reinforce the fiscal effect from an increase in b_U. If the part-time share increases, e.g., because full-time employees switch to unemployment at a higher rate than part-time employees, the spillover to part-time insurance imposes additional fiscal cost. If the increase in unemployment is rather coming from part-time employees, i.e. the part-time share declines, this counteracts some of the fiscal costs, because the government then needs to pay less part-time employment benefits.

Combination and Sufficient Statistics. The mutual dependence of unemployment and part-time employment insurance makes it difficult to assess the local effects of benefit changes of b_P and b_U in isolation. Therefore, since the focus of this paper is the effect of part-time employment benefits, I now assume that the unemployment benefit has already been chosen optimally, i.e., that the first order condition implied by the derivative of the welfare function with respect to b_U in equation (6.22) indeed identifies a local welfare maximum. In this case, one can use the first order condition for the optimal unemployment benefit to characterize the spillovers in the welfare effects of changes in the part-time employment benefit b_P. The marginal welfare gain of part-time employment benefits can then be rewritten as

$$M_W(b_P|b_U^*) = \underbrace{\frac{(u_P' - u_F')}{u_F'}}_{\text{Insurance Value}} - \underbrace{\frac{e_T}{e_F}\mathcal{E}_{\frac{e_P}{e_T},b_P}}_{\text{Direct Fiscal Costs}} \underbrace{- \frac{e_T}{e_F}\left(\frac{\mathcal{E}_{\frac{e_P}{e_T},b_U}}{\frac{e_T\left(u_U' - u_T'\right)}{u_T'} - \mathcal{E}_{1-e_T,b_U}}\right)\mathcal{E}_{1-e_T,b_P}}_{\text{Spillover to Unemployment Insurance}}. \quad (6.24)$$

This formula shows the local welfare effects of an increase in b_P, accounting for an optimally set unemployment benefit b_U.[27] It is general in the sense that it requires limited assumptions, e.g., about the nature of utility functions and the strength of behavioral responses. To bring this formula to the data, however, one requires some additional structural assumptions that will be discussed in the subsequent section.

6.5.4 Evaluating a Work-Life Balance Policy Expansion

To evaluate the welfare effects of a work-life balance policy expansion through the lens of the stylized model, I make some additional structural assumptions that are necessary to characterize the insurance value and the fiscal costs of the reform empirically, informed by the estimates of behavioral responses in Section 6.4.

Utility Function. To estimate the insurance value of the reform, one requires an estimate of the corresponding consumption smoothing value. For this purpose, I assume that utility is state-independent and embeds constant relative risk aversion, i.e.,

$$u_P = u_F = u_U = u = \frac{c^{1-\gamma}}{1-\gamma}, \quad u' = c^{-\gamma}, \quad u'' = -\gamma c^{-\gamma-1}, \quad (6.25)$$

[27] The derivation is relegated to Appendix F.3.

where γ is the coefficient of relative risk aversion.

Earnings Convexity. The income of individuals follows a convex earnings schedule in working hours l that is normalized around the hours under full-time employment, i.e,

$$y(l) = \theta l_F \left(\frac{l}{l_F}\right)^{1+\beta}, \tag{6.26}$$

where l_F represents the working hours under full-time employment and θ represents the hourly wage of full-time employees. This normalization implies that variations in the convexity parameter β do not change the earnings of full-time employees, and is only relevant for deviations from full-time employment. Thus, individuals in part-time employment face an implicit tax through the deviation from the linear earnings schedule, i.e.,

$$y(l) = \underbrace{\theta l}_{\text{Linear Earnings}} - \underbrace{\left(\theta l - \theta l_F \left(\frac{l}{l_F}\right)^{1+\beta}\right)}_{\text{Implicit Tax through Convex Earnings}}. \tag{6.27}$$

Progressive Income Tax. Consumption is given by gross income net of taxes, whereby taxes are approximated by a non-linear form with constant progressivity following Heathcote et al. (2017). The degree of tax progressivity is denoted by the parameter ρ. The total tax $T(y(l))$ is thus the sum of the implicit tax liability through convex earnings and the explicit tax liability through the progressive income tax, i.e.,

$$T(y(l)) = \underbrace{\left[\theta l - \theta l_F \left(\frac{l}{l_F}\right)^{1+\beta}\right]}_{\text{Implicit Tax through Convex Earnings}} + \underbrace{\left[\theta l_F \left(\frac{l}{l_F}\right)^{1+\beta} - \left(\theta l_F \left(\frac{l}{l_F}\right)^{1+\beta}\right)^{1-\rho}\right]}_{\text{Explicit Tax through Progressive Tax System}} \tag{6.28}$$

Based on the assumptions above, and using a first order Taylor approximation around the marginal utility under full-time employment, the insurance value in equation (6.24) can be approximated as

$$
\begin{aligned}
\frac{(u'_P - u'_F)}{u'_F} &= \frac{u'_F + u''_F\,(c_P - c_F) - u'_F}{u'_F} \\
&= \frac{u''_F}{u'_F}\,(c_P - c_F) \\
&= \left(\frac{c_F - c_P}{c_F}\right)\gamma \\
&= \left(1 - \frac{\left(\theta l_F \left(\frac{l_P}{l_F}\right)^{1+\beta}\right)^{1-\rho}}{(\theta l_F)^{1-\rho}}\right)\gamma \\
&= \left(1 - \left(\left(\frac{l_P}{l_F}\right)^{1+\beta}\right)^{1-\rho}\right)\gamma \\
&= \left(1 - k^{(1+\beta)(1-\rho)}\right)\gamma
\end{aligned}
\tag{6.29}
$$

where $k = \frac{l_P}{l_F}$ is the hours share of part-time employees relative to full-time employees.

Empirical Welfare Effects. The parameter values needed to estimate the insurance value and the fiscal costs are based on information from the literature and empirical estimates from the policy expansion in Section 6.4. Table 6.6 provides an overview with detailed explanations. Under the structural assumptions and with these baseline parameter values, one can estimate the reduced equation for the marginal welfare effect which reads

$$
\underbrace{\left(1 - k^{(1+\beta)(1-\rho)}\right)\gamma}_{\text{Insurance Value}} \quad - \quad \underbrace{\frac{1}{1 - s_{PT}}\mathcal{E}_{s_{PT},b_P}}_{\text{Fiscal Costs}} \quad = \quad M_W(b_P). \tag{6.30}
$$

Figure 6.13 displays the baseline results. The insurance value of 0.83 implies that a transfer of one Euro from full-time employment to part-time employment is worth an additional 83 cents through increased consumption smoothing. At the same time, transferring one Euro from full-time to part-time employment comes with a loss of 37 cents due to the behavioral responses of individuals. This yields an aggregate marginal welfare effect of 0.46 cents for each dollar transferred from full-time to part-time employment. Consequently, the 2019 reform was locally welfare-improving under the baseline parameter choices.[28]

[28] In Appendix F.3, I also evaluate the "bang-for-the-buck" of the policy by estimating the marginal value of public funds (MVPF) following Hendren and Sprung-Keyser (2020), and find that the implied MVPF is larger than those of other insurance policies.

Table 6.6. – Welfare Effect of Work-Life Balance Policy, Parameters

Panel A:

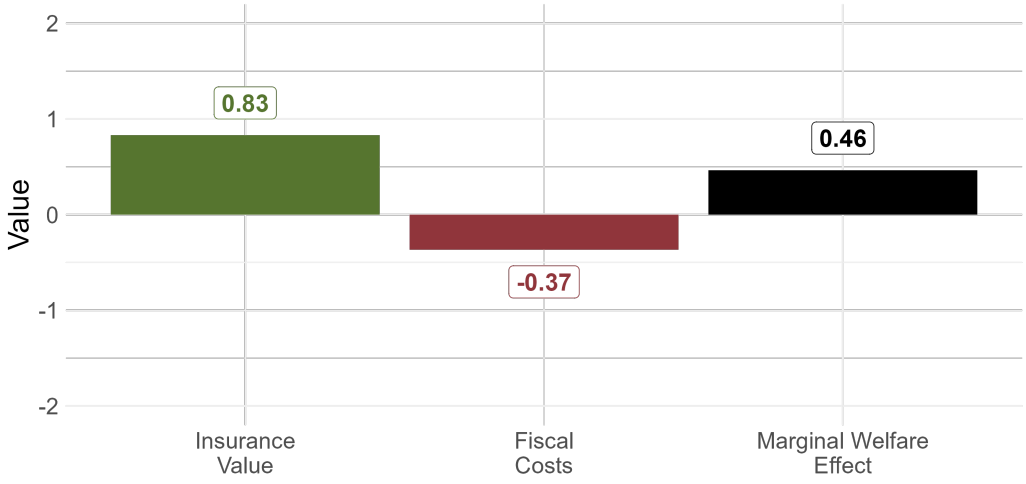
Parameter	Explanation	Value	Source
γ	CRRA Coefficient	1.40	Evans (2005)
ρ	Progressivity Parameter	0.16	Heathcote et al. (2020)

Panel B:

Parameter	Interpretation in Model	Value	Details
β	Earnings Convexity	0.17	Second stage estimate from instrumental variable regression, 2010-2014 (see Table 6.1)
k	Hours Share, Part-Time	0.40	Based on first stage estimate (1-δ) from instrumental variable regression, 2010-2014 (see Table 6.1)
s_{PT}	Part-Time Share	0.34	Average part-time share, 2018 (see Appendix Table F.8)
l_F	Hours of Full-Time Employees	38.88	Average weekly working hours of full-time employees in 2014 (see Table F.9)
θ	Hourly Wage, Full-Time	(13.26) 14.69	Average hourly wage of full-time employees, 2014, original (in brackets) adjusted by wage growth index to 2018 value (see Appendix Table F.9)
$d(\frac{e_P}{e_T})$	Change in Part-Time Share	0.0044	Baseline estimate from policy analysis (see Table 6.4)
$\mathcal{E}_{\frac{e_P}{e_T},b_P}$	Elasticity of the Part-Time Share w.r.t. benefit level	0.24	See Appendix F.3
$\mathcal{E}_{1-e_T,b_P}$	Elasticity of Non-Employment w.r.t. benefit level	≈ 0	See discussion in Section 6.4, and Figure 6.11d

Note: This table displays baseline choices for parameter values that are relevant for the estimation of the sufficient statistics. Panel A summarizes parameter choices that are based on the literature and not estimated in the context of the policy reform. Panel B summarizes parameter choices that are based on the empirical analysis in the context of the work-life balance policy expansion in 2019. For the adjustment of the hourly wage, I use the calendar and seasonally adjusted index of gross earnings provided by the German Federal Statistical Office.

Figure 6.13. – Empirical Welfare Effects of WLB Policy



Note: This figure visualizes the empirical money-metric welfare effects of the German work-life balance policy expansion in 2019. The figure differentiates between the insurance value (green bar) and the direct fiscal costs (red bar) of the policy. The black bar visualizes the aggregate welfare effect. All calculations are based on the baseline parameter values in Table 6.6. For an individual variation of parameters and their influence on the welfare assessment, see the Interactive Work-Life Balance Policy Assessment Tool (I-WLB-PAT).
Source: Own calculations based on estimates from Table 6.6.

Determinants of Marginal Welfare Effect and Robustness. Both components of the marginal welfare effect—insurance value (IV) and fiscal costs (FC)—are sensitive to variations in the different parameters that can be assessed with a focus on the partial derivatives of equation (6.30) given by

$$\frac{\partial IV}{\partial \gamma} = 1 - k^{(1+\beta)(1-\rho)} \geq 0 \qquad \frac{\partial FC}{\partial s_{PT}} = \frac{-1}{(1-s_{PT})^2}\mathcal{E}_{\frac{e_P}{e_T},b_P} \leq 0,$$

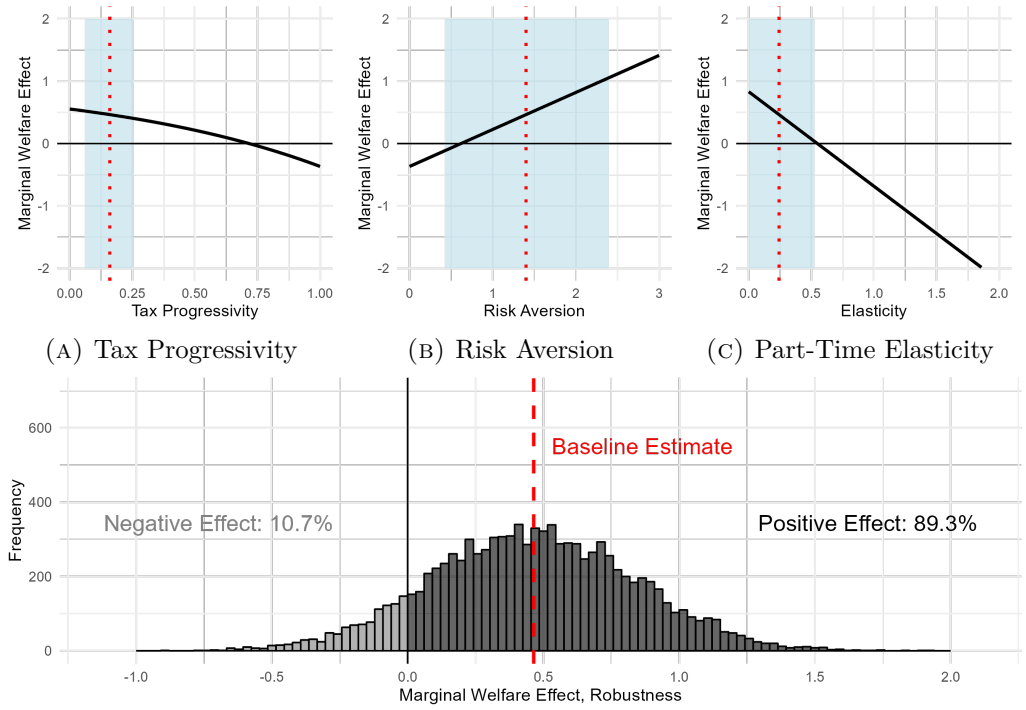
$$\frac{\partial IV}{\partial \rho} = k^{(1+\beta)(1-\rho)} ln(k)(1+\beta)\gamma \leq 0,$$

$$\frac{\partial IV}{\partial k} = -k^{(1+\beta)(1-\rho)-1}\gamma \leq 0 \qquad \frac{\partial FC}{\partial \mathcal{E}_{\frac{e_P}{e_T},b_P}} = \frac{e_T}{e_F} \geq 0,$$

$$\frac{\partial IV}{\partial \beta} = -k^{(1+\beta)(1-\rho)} ln(k)(1-\rho)\gamma \geq 0.$$

The equations highlight that the insurance value increases in the coefficient of relative risk aversion γ, since consumption smoothing across states then comes at a higher monetary value. Also, a higher progressivity of the tax code, i.e. a lower progressivity parameter ρ, reduces the insurance value, since consumption is then already more equal in the two states. Similarly, if the hours difference between full-time employment and part-time employment is smaller, i.e. k is higher, the insurance value decreases as well. In contrast, a more convex earnings schedule (higher β) increases the insurance value, since

Figure 6.14. – Marginal Welfare Effect, Robustness



(A) Tax Progressivity (B) Risk Aversion (C) Part-Time Elasticity

(D) Joint Variation

Note: This figure visualizes, how robust the marginal welfare effect is to the variations of three core parameters, the tax progressivity parameter ρ, the coefficient of relative risk aversion γ, and the part-time employment elasticity $\mathcal{E}_{\frac{e_P}{e_T},b_P}$. Figures 6.14a-6.14c display how isolated variations of these parameter values influence the marginal welfare effect, holding all other values constant at the baseline level. The vertical red line indicates the baseline parameter values shown in Table 6.6. Figure 6.14d displays results from a Monte-Carlo simulation, where all three parameters are modified independently and simultaneously. In particular, I draw at each iteration a value for all three parameters from a normal distribution with mean at the baseline value and respective standard deviations of 0.5, 0.05, and 0.15. Based on these standard deviations, 95% of the parameter values fall in the light blue areas. The Figure displays the distribution resulting from a Monte Carlo simulation with 10,000 draws. The baseline estimate of the welfare effect is shown as a red vertical line. For an individual variation of parameters and their influence on the welfare assessment, see the Interactive Work-Life Balance Policy Assessment Tool (I-WLB-PAT).
Source: Own simulations based on estimates from Table 6.6.

the divergence in consumption between the two states increases correspondingly. On the other hand, the fiscal costs decrease with the part-time share, since the relative change in the part-time share measured by the elasticity implies lower changes in absolute terms. In contrast, a higher elasticity of the part-time share increases fiscal costs.

Figure 6.14 provides information on how sensitive the baseline estimate of the marginal welfare effect is to variations in the tax progressivity parameter ρ, the coefficient of relative risk aversion γ, and the elasticity of the part-time share $\mathcal{E}_{\frac{e_P}{e_T},b_P}$. For changes in the progressivity parameter, the welfare-effect only turns negative for a values of ρ larger than 0.7 (see Figure 6.14a). Such a large progressivity parameter is implausible for Germany, and also falls outside of values estimated in other contexts. For instance, the estimates by Heathcote et al. (2020) show progressivity parameters between 0.09 (United States)

and 0.23 (Denmark). Figure 6.14b reveals that the reform is welfare-improving for CRRA coefficients that are larger than 0.6, holding all other parameters constant. Since there is considerable variation in empirical estimates of CRRA-coefficient (Chetty, 2006a; Evans, 2005; Layard et al., 2008), a value of 0.6 would still be plausible, but rather at the lower end of the spectrum of empirical estimates. Finally, as indicated by Figure 6.14c, isolated variations in the part-time elasticity show that the welfare effect is positive up to an elasticity of around 0.5, a value outside of the statistical confidence bands for the baseline estimate of the German WLB policy expansion (see Appendix Table F.11).

Beyond isolated variations in the parameter values, I also conduct a Monte Carlo simulation that draws random values for the three parameter values independently from normal distributions around the respective baseline estimate. This simulation provides insights about the robustness of the conclusion of a positive welfare effect to joint variations of these parameters. As shown by Figure 6.14d, the simulation reveals that the marginal welfare effect is positive in about 90 percent of all draws. In sum, while the magnitude of the marginal welfare effect is sensitive to variations in parameter values, the sign of the welfare effect is relatively robust.

6.6 CONCLUSION

Motivated by public discussions on the reconciliation of market work, leisure activities, and family obligations, this paper provides empirical and theoretical insights into work-life balance (WLB) policies in the German context. Using rich administrative data from the German microcensus and linked employer-employee records, the empirical analysis documents the state and evolution of WLB, as well as the causal effects of a recent WLB policy reform. The theoretical section introduces a stylized model that interprets WLB policies as an insurance mechanism against employment shocks that hinder full-time work. The empirical estimates are then linked to the model's sufficient statistics to assess the welfare implications of the policy reform

The descriptive analysis reveals that the demand for work-life balance in Germany has increased over the past decade, and is particularly prevalent among individuals with higher education. While some of the desire for increased work-life balance has already manifested in reduced working hours, full-time employees on average still work more than they would like to. One suggested reason for the mismatch between desired and actual hours of work is the convexity of the earnings schedule, which associates shorter working hours

with lower hourly wages. In particular, conditional on a fine-grained set of observables, a half-time employee earns more than 10 percent less per hour than a comparable full-time worker, highlighting the financial costs of reducing working time.

Since the 2000s, policy makers in Germany have made it easier for individuals to reduce their working time without having to face lower hourly wages. To assess the causal effect of such policies on employment decisions, this study evaluates the most recent expansion in 2019, which granted employees the right to temporarily reduce their working hours without employer consent. Exploiting firm size-based quasi-experimental variation in eligibility, I find that the policy led to a moderate differential increase in part-time employment of approximately 0.4 percentage points. This effect is primarily driven by a 14 percent increase in direct transitions from full-time to part-time work, while there is no significant impact on transitions from unemployment to part-time employment. Additionally, contrary to the policy's objectives, there is no evidence that it increased transitions from part-time back to full-time work.

The theoretical analysis illustrates the trade-offs associated with WLB policies. In particular, the proposed model interprets WLB policies as a public insurance program against employment shocks that hinder full-time work, for instance due to childcare or eldercare responsibilities. By allowing individuals to choose part-time employment without having to face cuts in hourly wages, the government linearizes the convex earnings schedule and provides consumption smoothing to individuals that are hit by such shocks. However, this also increases the attractiveness of part-time employment, which introduces moral hazard and potentially comes with a loss of fiscal resources. Informed by the empirical degree of earnings convexity before the policy, and the estimated behavioral responses, the analysis suggests that the 2019 reform was welfare-improving for a plausible range of parameter values.

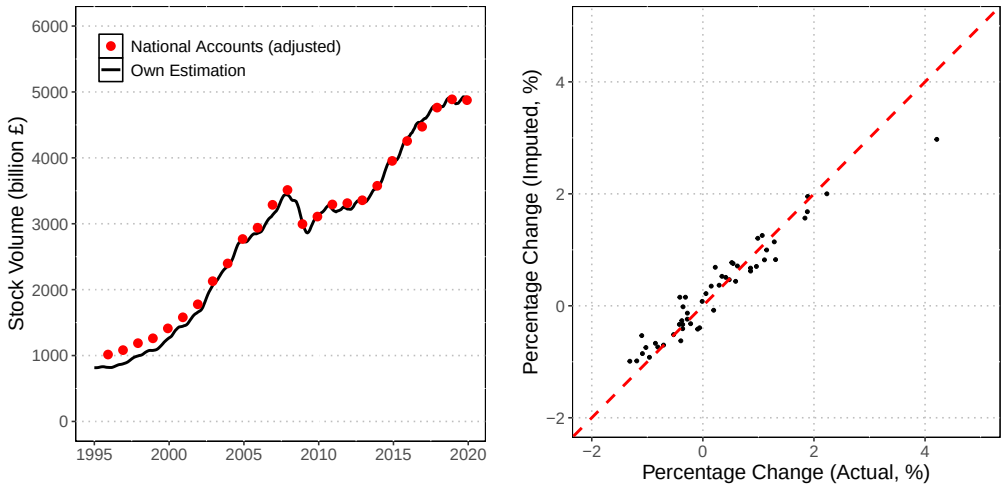
While this study provides important insights into the shifting demand for WLB and the welfare implications of WLB policies, several avenues for future research remain. First, a more precise understanding of the perceived benefits and costs of WLB at the individual level is needed. While large-scale microcensus data enables the analysis of WLB preferences within fine-grained subpopulations, its questions lack the specificity to measure individuals' willingness to pay for reduced working hours. Future research could incorporate stated preference experiments, such as those proposed by Maestas et al. (2023), to elicit individuals' hours choices under different degrees of earnings convexity. Second, the causal analysis in this paper is centered around the reactions to the policy at the level of employees and does not take into account that firms might also change their behavior, e.g., related to hiring decisions and decisions about the earnings structure in

firms. The short-run consequences of the policy might thus be very different from its long-run impact. Dynamic concerns might also influence the behavior at the individual level. For instance, long-term career concerns might hinder individuals to reduce their working hours and thereby limit the effectiveness of work-life balance policies. Finally, the stylized model could be refined by incorporating ex-ante heterogeneity as well as interactions between WLB policies and other tax and transfer instruments. Ex-ante heterogeneity between individuals might be particularly important, since the desire for work-life balance is shown to vary strongly across the skill distribution, and since the part-time employment risk might be very heterogeneous. Moreover, WLB policies share similarities with in-work tax benefits, which also provide consumption smoothing to part-time employees. Furthering our understanding about the complementarities and differences between these policy instruments would thus be an additional valuable path for future research.

Appendix to Chapter 1

A.1 ADDITIONAL RESULTS

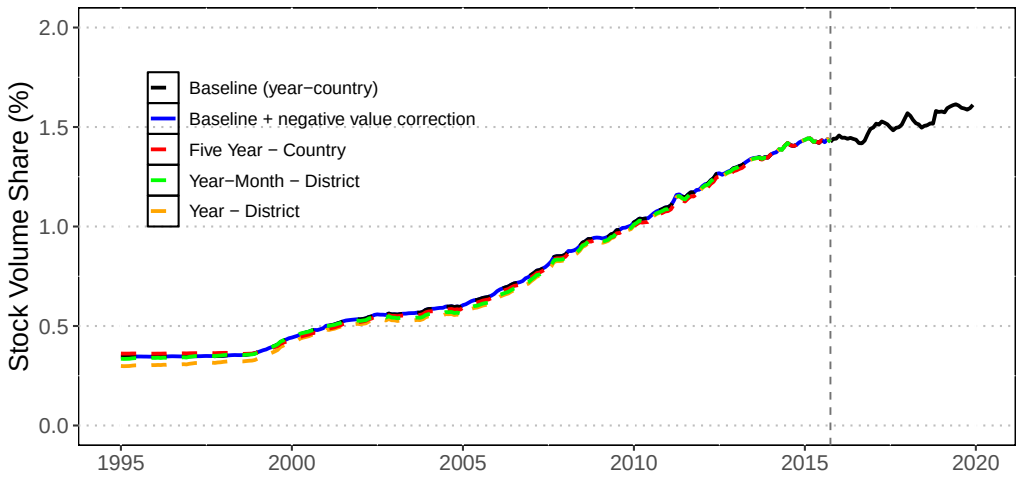
Figure A.1. – Validating Long-Run Time Series



(A) Residential Market Volume (B) Offshore Market Share Growth

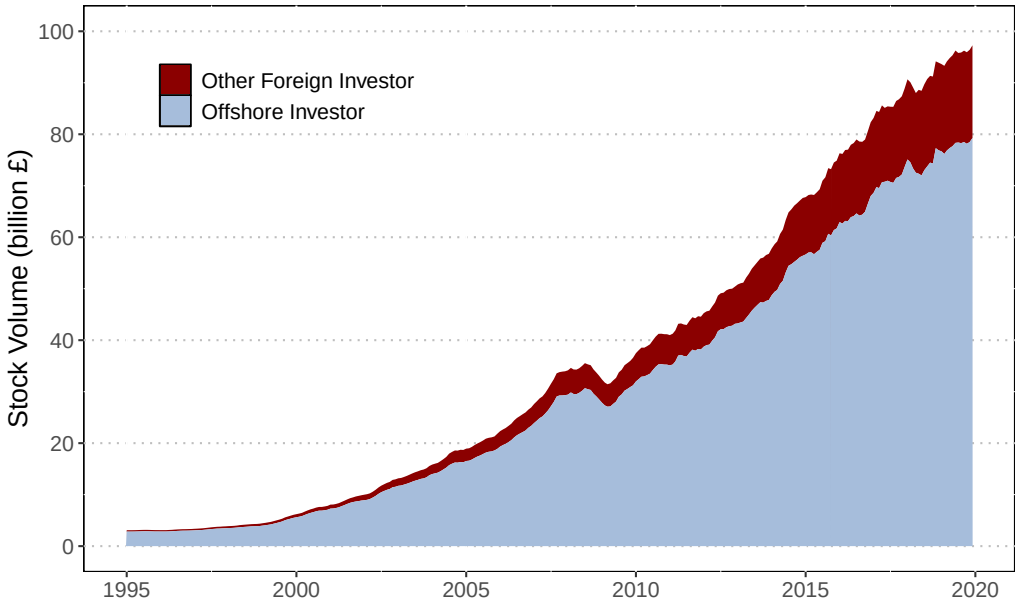
Note: The figure shows the results from two exercises that aim to validate our strategy for estimating long-run time series of real estate values. Figure A.1a compares the total value of residential real estate in England and Wales according to our estimation (black line) with adjusted values from the National Accounts (red points). Figure A.1b compares imputed (y-axis) and actual (x-axis) month-to-month growth rates of the offshore market share for the time between October 2015 and December 2019 when reliable snapshot data is available. For details on the estimation of market shares and the validation exercises, see Appendix A.3.

Figure A.2. – Robustness of Imputation Approach

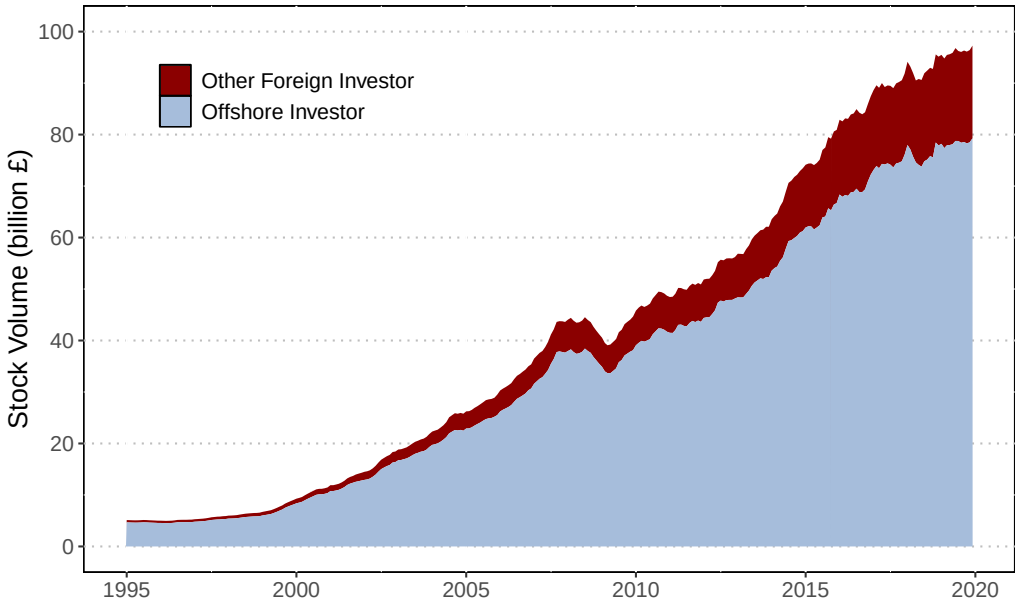


Note: The figure illustrates the robustness of the estimated long-run dynamics in the offshore market share (Figure 1.4) to alternative choices in the imputation procedure. In the baseline, transition probabilities are calculated at the year-country level and negative values at the district level are preserved (black line). Alternative approaches set negative values to zero at each imputation stage (blue line) or calculate transition probabilities at the five-year-country level (red line), month-district level (green line) and year-district level (orange line). For details on the estimation of market shares and the validation exercises, see Appendix A.3.

Figure A.3. – Foreign Corporate Ownership of Residential Real Estate



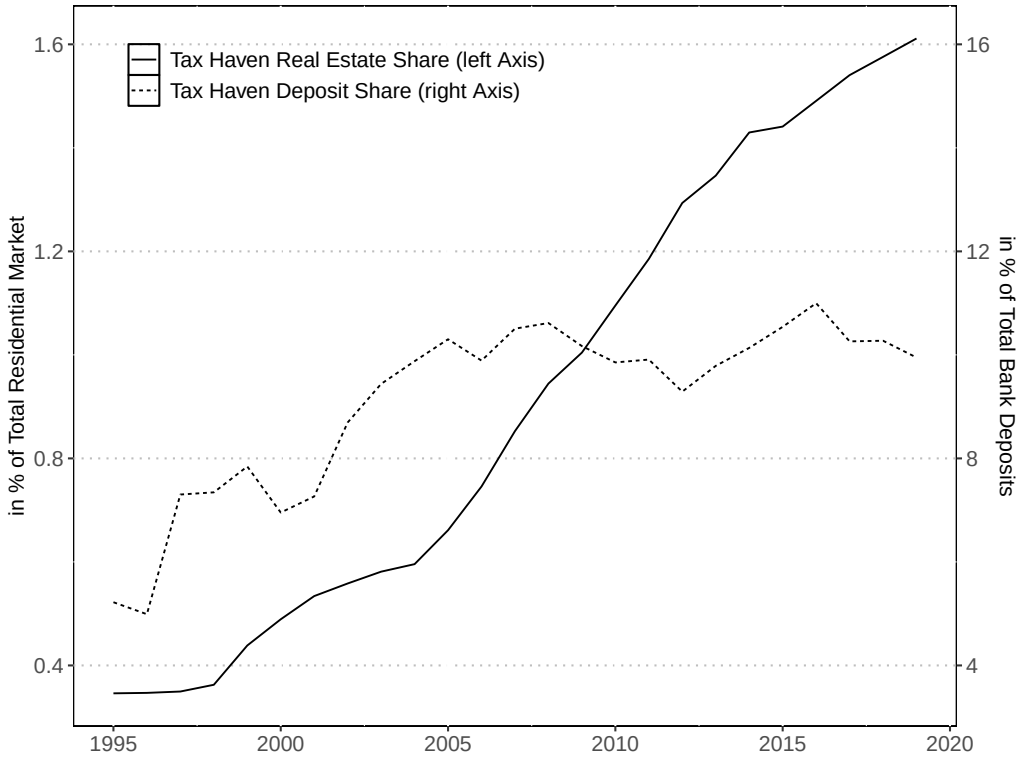
(A) Current Prices



(B) Real Prices

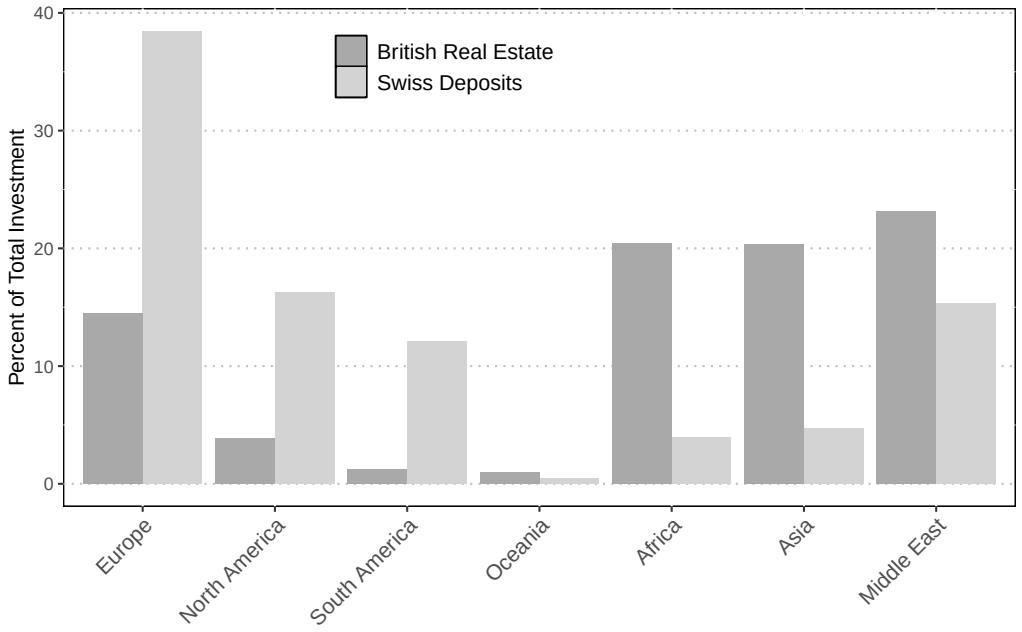
Note: The figure shows the estimated aggregate value of residential real estate in England and Wales held by corporations in offshore tax havens (blue) and other foreign corporations (red). The estimates include direct as well as indirect ownership. The top panel shows values in current prices while the bottom panel shows values in 2019-prices (deflated with CPI). For details on the ownership assignment and the estimation of market shares, see Appendix A.2 and Appendix A.3.

Figure A.4. – Tax Haven Share, Real Estate vs. Bank Deposits



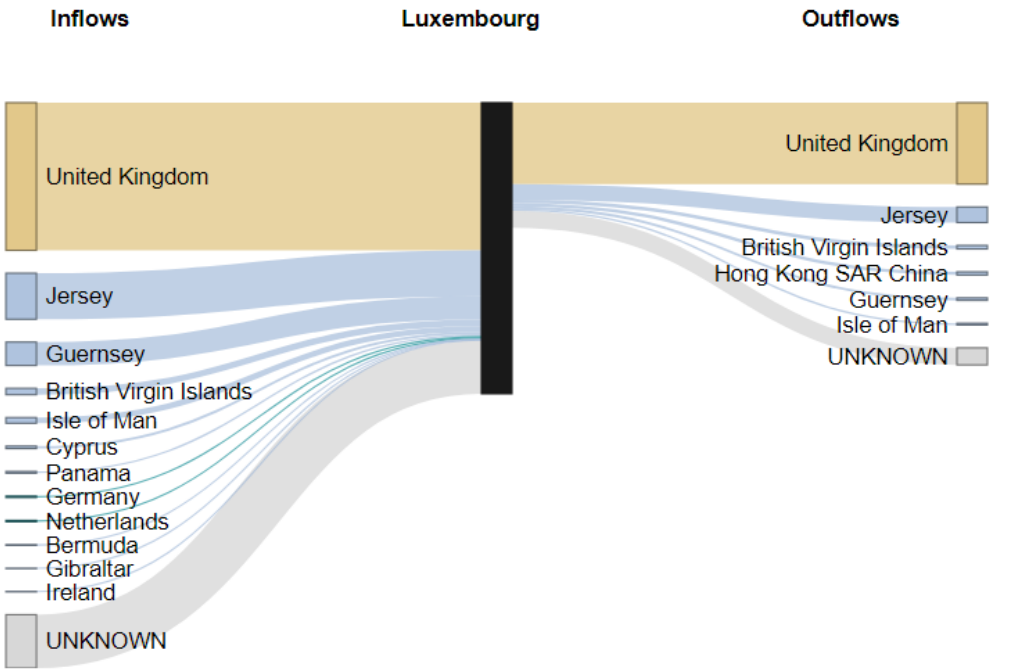
Note: The figure compares the share of aggregate residential real estate value in England and Wales owned by offshore tax havens (solid line, left horizontal axis) to the share of deposits in British banks owned by offshore tax havens (dashed line, right horizontal axis). The tax haven share of deposits is computed based on the Locational Banking Statistics of the Bank for International Settlements.

Figure A.5. – Beneficial Ownership, Real Estate vs. Banking



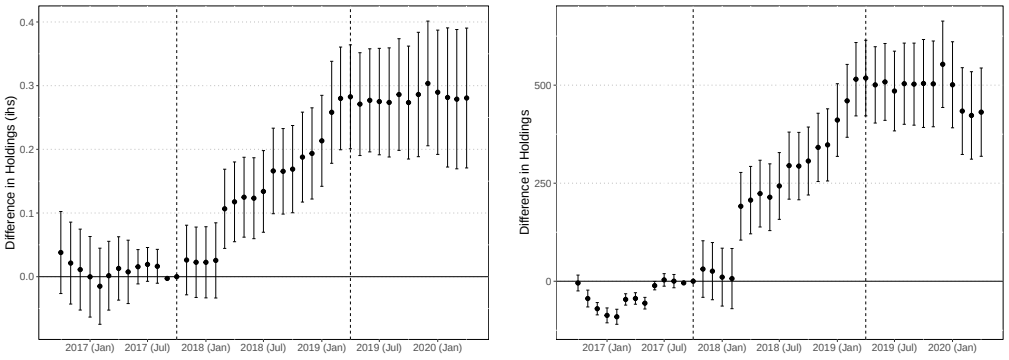
Note: The figure compares the distribution of beneficial ownership across geographical regions for residential real estate in England and Wales (dark gray bars) and for financial assets managed in Switzerland (light gray bars). The beneficial ownership of financial assets is based on the leaked customer records from HSBC Switzerland published by the International Consortium of Investigative Journalists as the *Swiss Leaks*. Both metrics exclude ultimate owners from the United Kingdom.

Figure A.6. – Tax Motive, Flows

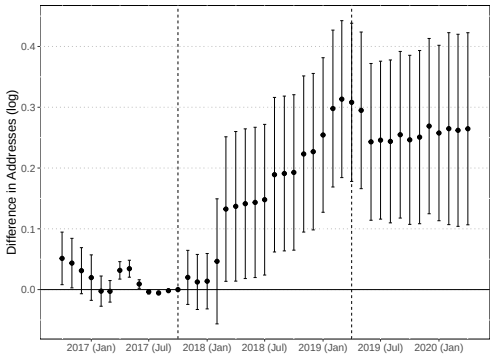


Note: The figure visualizes the flows of property holdings to (from) companies incorporated in Luxembourg between October 2017 and April 2019. Country assignments in flows to (from) Luxembourg are based on the owner that precedes (succeeds) the Luxembourg corporation within at most a time window of 6 months. If no direct predecessor (successor) is visible in the data, we assign this predecessor to the "Unknown" category that consists of private owners, deconstructed properties and newly built properties.

Figure A.7. – The Tax Motive, Robustness



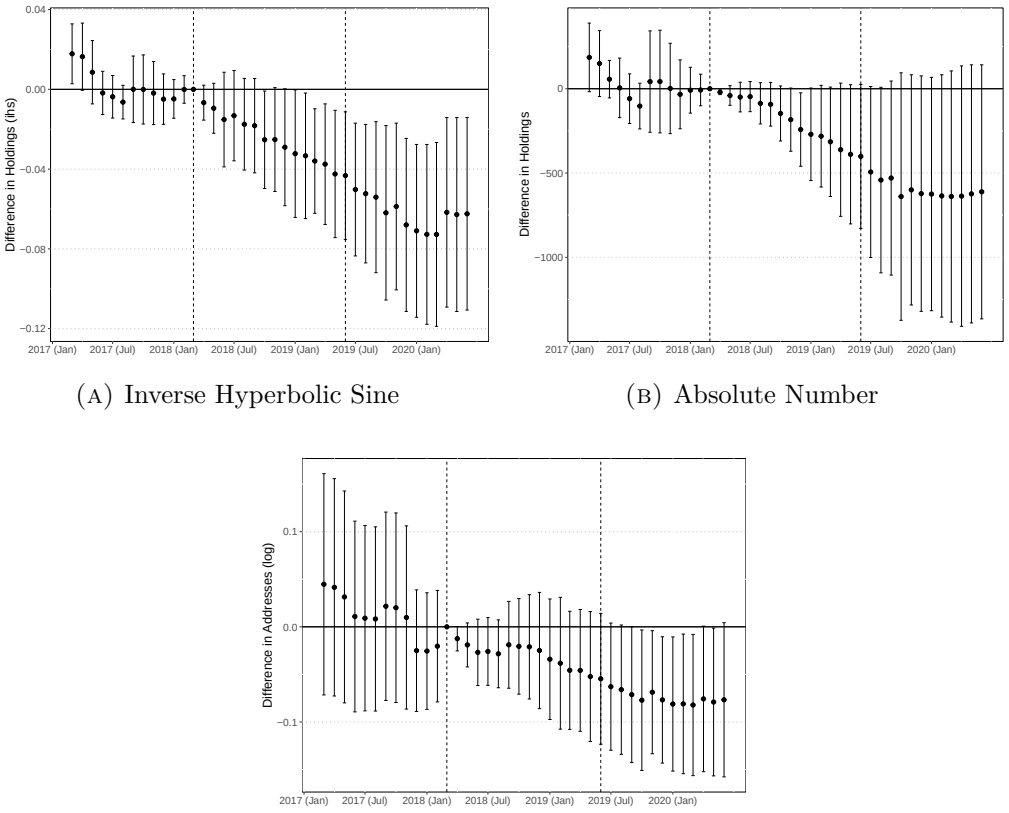
(A) Inverse Hyperbolic Sine (B) Absolute Number



(C) Property Addresses

Note: The figure shows robustness tests of the baseline result on the tax motive. The robustness tests use alternatives to the logarithmic transformation of the outcome variable: the inverse hyperbolic sine (Panel A) and the absolute number without any transformation (Panel B). They also use an alternative unit of observation: the number of property addresses (Panel C).

Figure A.8. – The Secrecy Motive, Robustness



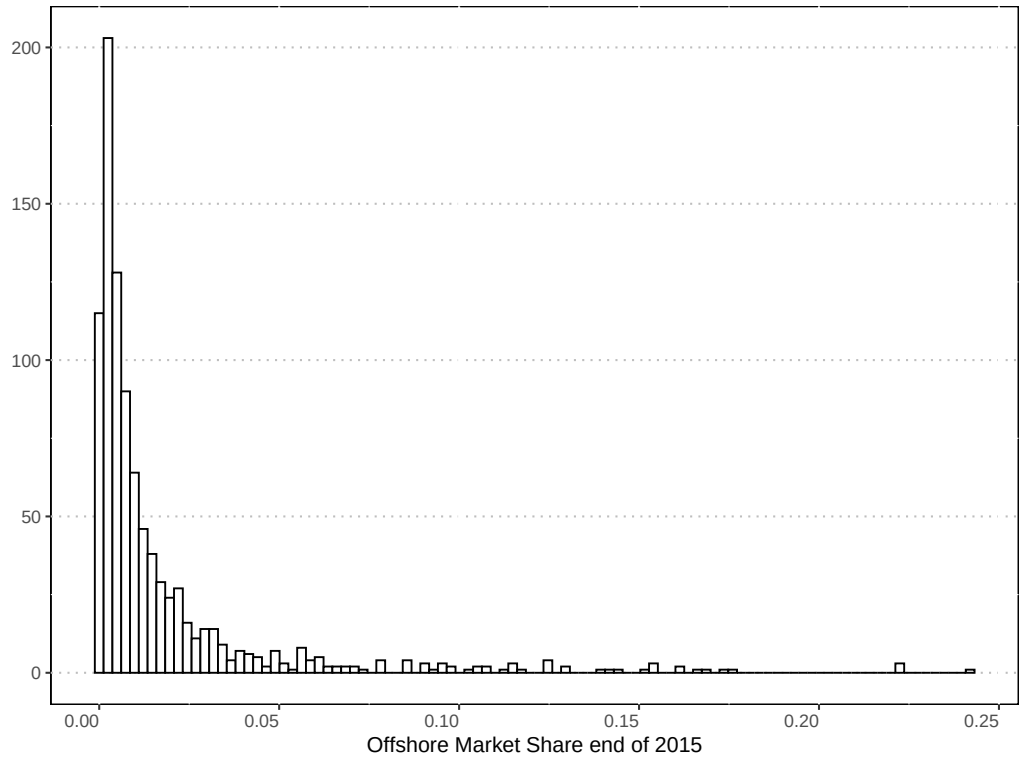
(A) Inverse Hyperbolic Sine

(B) Absolute Number

(C) Property Addresses

Note: The figure shows robustness tests of the baseline result on the secrecy motive. The robustness tests use alternatives to the logarithmic transformation of the outcome variable: the inverse hyperbolic sine (Panel A) and the absolute number without any transformation (Panel B). They also use an alternative unit of observation: the number of property addresses (Panel C).

Figure A.9. – Distribution of Offshore Ownership Shares in London



Note: The figure shows the distribution of the share of residential real estate owned by corporations in offshore tax havens across 983 local areas in London (omitting the 44 areas with no offshore ownership).

A.2 COMBINING DATA SOURCES

This part of the appendix provides details on the combination of price, purchase date, and ownership information to create a comprehensive dataset of British real estate transactions. In Section A.2.1, we describe in detail, how address string information from various sources is harmonized. Section A.2.2 discusses, how ownership information from the land register is combined with information about ownership chains from ORBIS and ultimate ownership details from the ICIJ leak data. Finally, we provide insights on the conditions that are utilized to match price and ownership information (Section A.2.3).

A.2.1 Preparing British Addresses

Understanding the relative importance of foreign held residential real estate requires information that is not readily available from British sources. The Price Paid data records

sales price information for the entire residential market, the different ownership datasets contain the immediate ownership information for the entire corporate market. Our goal is to match these sources to obtain combined price-ownership information for the residential market that is owned by corporate investors. Unfortunately, the UK registers do not publish identifiers linking the Price Paid data and the ownership data. Even if these were published, they would most likely link the data specific identifiers which are of limited use. The main identifier in the Price Paid data is a "transaction id" that can include many properties sold in one transaction. The main identifier in the ownership data is the "land title registration number" that can include dozens of addresses that have the same owner and are registered as one land title. To be able to combine the data, we move all datasets to the level of a postal address. An example for a simple data entry is:

"FLAT 5, 213 SUSSEX GARDENS, LONDON (W2 2RJ)"

This entry can easily be moved to the postal address level:

Postcode	**Street**	**House No.**	**Unit**	**Unit No.**	**Locality**	**No.**
W2 2RJ	SUSSEX GARDENS	213	FLAT	5		

An example for a complicated data entry is:

"FLATS 1-27 WALLACE COURT, 54 TIZZARD GROVE, LONDON (SE3 9EE), FLATS 103-128 WALLACE COURT, 44 TIZZARD GROVE, LONDON (SE3 9EQ)AND FLATS 129-157 WALLACE COURT,52 TIZZARD GROVE, LONDON (SE3 9FE),1-48 GRAYSTON HOUSE, 21 ASTELL ROAD LONDON (SE3 9FN), 49-110 GRAYSTON HOUSE, 1 OTTLEY DRIVE, LONDON (SE3 9FP), 1-62 MALTBY HOUSE, 2 OTTLEY DRIVE, LONDON (SE3 9FJ), 63-105 MALTBY HOUSE, 18 TUDWAY ROAD, LONDON (SE3 9FL), 5-12 OTTLEY DRIVE, LONDON (SE3 9FT), 2-16 (EVEN) TUDWAY ROAD, LONDON (SE3 9FR)"

This entry is separated by our code into multiple postal addresses:

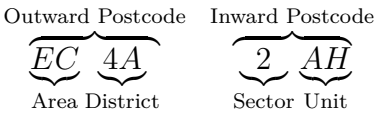
Postcode	**Street**	**House No.**	**Unit**	**Unit No.**	**Locality**	**No.**
SE3 9EE	TIZZARD GROVE	54	FLAT	1	WALLACE COURT	
SE3 9EE	TIZZARD GROVE	54	FLAT	2	WALLACE COURT	
...	...	...	...	...	...	...
SE3 9EE	TIZZARD GROVE	54	FLAT	27	WALLACE COURT	
SE3 9EQ	TIZZARD GROVE	44	FLAT	103	WALLACE COURT	
...	...	...	...	...	...	...
SE3 9FP	OTTLEY DRIVE	1			GRAYSTON HOUSE	94
...	...	...	...	...	...	...

This process is carried out for the following datasets:

- the Price Paid database that collects all residential transactions
- the CCOD database that contains all domestic corporate investment
- the historical CCOD database
- the OCOD database that contains all foreign corporate investment
- the historical OCOD database
- the freedom of information request data containing foreign transactions since 1990
- the Private Eye data containing a 2014 snapshot of foreign ownership

Detecting Postcodes. Our code starts with basic string cleaning exercises such as: Moving everything to uppercase letters, adjusting typos such as "Esat" instead of "East", eliminating apostrophes, aligning cardinal points ("Northerly = North", or "North-West = North West"), aligning punctuation, removing trailing or leading white spaces as well as unnecessary punctuation.

We then start the address preparation by separating data entries along the information on postcodes. We exploit the fact that postcode assignment in the UK follows regularized patterns and contains different layers which are shown in the following exemplary postcode:



Given that the different layers allow for a limited number of patterns, we can use regular expressions to identify the postcode. Since postcodes usually appear at the end of an address entry, any auxiliary information that appears after the postcode is dropped.[1] In

[1]For instance, this auxiliary information includes strings like "and adjacent parking lot" that are not relevant for the preparation of the address.

very rare cases, there is only one postcode for a land title that includes several streets and it is only specified for the first street. If we find any additional street information after a postcode, we assign the same postcode.[2]

Detecting Streets. Having separated addresses by postcodes, we next turn to street names. Since a data entry can contain multiple streets under the same postcode, we first use regular expressions to identify, how many different streets are included in the entry of the land registry.[3]

Depending on the number of matches, we then repeatedly extract streets from the data entry following a two step procedure. The first step makes use of the list of street names provided by the national mapping agency of the UK.[4] In particular, we extract streets in the data entry by looking through all street names associated with the respective outward postcode. The identification of street names via location-specific lists has the advantage that it reduces false positives associated with terms that can be a street name in one place, but are a locality or property name in others. In a second step, if we have not identified a street using the lists above, we extract remaining street names based on street regular expressions using general terms commonly associated with streets.

For each extracted address, we also obtain a leading string that contains house numbers, unit information (e.g. "FLAT 10", "APARTMENT 5"), locational information (e.g. "LAND AT", "BUILDINGS LYING TO THE SOUTH OF"), and locality information (e.g. "WALLACE COURT"). In the next sections, we discuss how we separate these elements.

Separating House Numbers, Units, and Localities. In the data, house number information is placed directly in front of the street while unit and locality information is located in front of the house number (e.g. "FLAT 7, DORCHESTER COURT, 77-81

[2]In particular, we look for the appearance of any of the following street regular expressions: STREET, ROAD, ARCADE, VALE, LANE, AVENUE, CLOSE, WHARF, ACRE, MEWS, WALLS, CHASE, SQUARE, CRESCENT, PLACE, ROW, TERRACE, WAY, QUAY, BOULEVARD, GATE, CLOSE, DRIVE, GROVE, GARDENS, GARDEN, CIRCUS, WALK, EMBANKMENT, HEIGHTS, GARTH, PARKWAY, COURT, YARD, VIEW, HILL, PARADE, APPROACH, RISE, PATH, GLADE, and TERRACE.

[3]We use the same list of street pointers used in the postcode preparation above.

[4]These data are available from the Ordnance Survey.

SLOANE STREET, LONDON (SW1X 9SE)"). This order allows us to identify the house number as a regular expression that directly precedes the street name and contains only non-alphabetical strings.[5]

Once the house number is extracted, we detect units and their corresponding number(s) by searching the remaining entry for elements of an extensive list of unit types.[6] The non-alphabetical regular expressions following the extracted unit types are recorded as the respective unit number(s). Again, unit numbers are allowed to follow the same patterns as house numbers.

The land registers are full of special property names such as "Wallace Court" or "Grayston House". We opt for treating these localities different from the units described above, though the often fulfill the same role in identifying addresses. The most important difference is that unlike flats or apartments, these localities often appear without a number associated with them. We detect localities using regular expressions which is simpler then detecting units after the latter have been taken out of the running string.

Treating House, Unit and Locality Numbers. Having identified the strings that include information about the house, unit and locality numbers, we transform them into a harmonized numeric appearance. For instance, we create integer sequences of simple numeric sequences ("1-3" becomes "1,2,3"). Thereby, we pay special attention to ranges of numbers that also include letters (e.g. "3A-4C") and make sure that all combinations (e.g. "3A, 3B, 3C, 4A, 4B, 4C") are created for the final data.

Whenever one of the above steps leads to more than one postcode, street, house number, unit number, or locality number, we reshape the dataset into long format. This process results in a dataset of postal addresses that are uniquely identified. Whenever data quality concerns prohibit precise recognition, the entire land register entry is dropped from the dataset.

The classification of address information is a challenge that is also discussed in the field of computer science. Bourne et al. (2018) show how a general model of natural language processing (spaCy) has to be trained on the OCOD data to structure the address string snippets from the UK land register. The goal of their analysis is to provide a good

[5]We allow for one-digit alphabetical strings attached to numbers and for particular alphabetical strings such as "ODD" or "EVEN". Thereby, our code is also able to identify more complicated house number combinations like "1-11 (ODDS)", "16-20 (EVEN)", or "14, 17A-18C".

[6]Among others, unit types include "APARTMENT","FLAT","UNIT", "BUILDINGS", "BUNGALOWS", "SUITE", "ROOM", "FLOOR".

aggregate classification of market shares where we are interested in offshore investment specifically and at the property level. We therefore do not infer prices based on area level aggregates.

By manually inspecting the different data sources (FOI, PPD, OCOD, CCOD), our approach allows a characterization that is tailored to capture the heterogeneity in the address recordings in our various data sources. In addition, we aim to provide a transparent classification while machine learning algorithms are often a black box.

A.2.2 Preparing Ownership Information

HM Land Registry differentiates its registers for commercial real estate between real estate titles held by domestic companies (CCOD) and titles held by foreign companies (OCOD).[7]

The Commercial and Corporate Ownership Data (CCOD) contains all freehold and leasehold registered titles to land or property in England and Wales which are held by non-private entities incorporated in the UK. Examples of non-private entities include public limited companies, unlimited companies and limited liability partnerships.[8] While land registration in the UK exists for more than 150 years, is only publicly available and updated on a monthly basis from November 2017 onward. In addition, using the legal procedure of a Freedom of Information (FOI) request, we obtained quarterly data for the time period between March 2014 and September 2017. The latest version of the data contains more than three million property titles including both property-related information (e.g. geographical information, tenure) and owner-related information (e.g. name, company registration number, proprietor type).[9]

The Overseas Company Ownership Data (OCOD) is the corresponding administrative register for property titles registered to companies outside of the UK. OCOD was made publicly available in 2017 as part of the government's strategy to enhance the understanding of the UK housing.[10] Apart from information recorded in the CCOD, the OCOD

[7]Detailed information on CCOD and OCOD is provided by HM Land Registry.

[8]The data does exclude Charitable Incorporated Organizations. Land register data for private individuals is also maintained by HM Land Registry, but not available for the public.

[9]For some properties, the land register also contains information about the price paid. However, since the recordings of the prices is incomplete, we ignore the price information from the land register, since we have comprehensive price information in the transaction data (Price Paid data).

[10]This process had been pushed by a data publication of the current affairs magazine *Private Eye* in 2014. It contains property and ownership information of around 100,000 property titles in England and Wales acquired by foreign companies between 1999 and 2014. For details on the data and related news coverage, see information by the Private Eye Magazine.

contains the country of incorporation of the proprietor(s). Although the country of incorporation has been routinely recorded in the register since 1999, OCOD is also only available from November 2017. In the process of our FOI request, we could extend the coverage obtaining quarterly OCOD data which dates back to October 2015. As of June 2020, OCOD counts around 100,000 property titles which we separate into the addresses described above and in the main text.

Preparing Companies in the British Registers. In order to link ownership information across different data sources, we also harmonize the details about the corporate owners in three steps. First, we align the string entries by treating leading and trailing white spaces, caps, or special characters. Second, we extract company types from the strings based on an extensive list of incorporation types and separate them from the information about the company name.[11] Finally, we harmonize company types and reclassify them in a unified way given that a limited liability company for instance could be recorded as "Limited", "Ltd.", or "Limited Liability Company". For companies incorporated abroad, we use a fourth step to extract and harmonize information about the country of incorporation. This harmonization is not only necessary due to frequent misspellings of country names, but also because the records sometimes present regional or local information instead of the country information (e.g. "Tortola" instead of "British Virgin Islands", or "Abu Dhabi" instead of "United Arab Emirates").

Adding Orbis Ownership Information. While the British land register data separated between real estate held by foreign and domestic companies, a domestic company can directly be owned from abroad. This would disguise foreign ownership as domestic ownership. In order to address this issue, we employ corporate ownership information from the Orbis dataset by Bureau van Dijk (BvD). This proprietary dataset includes data on 400 million companies and their ownership structure. Starting from the British company, we first harmonize company incorporation types to create a 1:1 match with the Orbis dataset and assign a BvD ID to the British companies that appear in the British registers. In the simplest cases, Orbis already has information on the global ultimate owner of a company, defined as holding more than 50% of the voting shares (GUO 50 in Orbis) of a company. In that case, we can simply re-assign the country of origin from "United Kingdom" to that owner for the company in question. If this is not the case, we exploit information on shareholders. If a company in Britain for example has one shareholder in France who is not the ultimate owner, we move to that shareholder and check if it has an ultimate owner. The country that has the majority ultimate owner

[11] Examples of company types include "LTD", "SARL", and "PLC".

share is then designated as the country of origin of an investment. Whenever we cannot identify global ultimate owners, we refrain from reclassifying the country of origin. This is a conservative approach once the shareholder structure has moved offshore but avoids assigning the shareholder country when no global ultimate owner is identified. Therefore, in our dataset, this data effort re-codes the country of origin from "United Kingdom" to the respective country of the GUO 50 we find in Orbis.

Matching ICIJ Ownership Information. Shell companies in tax havens are routinely used to hide ownership. A regular stream of large data leaks has been collected into a consistent database by the International Consortium of Investigative Journalists (ICIJ). The last installment of this dataset was published in May 2022, adding further data from the *Pandora Papers* leak. The dataset used here includes all available information from the *Offshore Leaks* (2013), the *Panama Papers* (2016), the *Bahamas Leaks* (2016), the *Paradise Papers* (2017), and the *Pandora Papers* (2021). Together with the files on Mossack Fonseca in the Paradise Papers, parts of the corporate registers of six well known offshore jurisdictions were leaked with much less media attention but more data.[12] In total, the data contains information on more than 810,000 offshore entities and the respective managerial structures. Most importantly, the data identifies the persons connected to an offshore identity ('officers') and reveals their respective role in the company (e.g. beneficial owner, nominee director, nominee shareholder). After several updates over the last years, this database is now available online as a network database.

Preparing Offshore Entities In order to match the information in the leaked databases to the British land registers, we first prepare the company name and company type information using the approach described above for British companies. The list of company types, including misspelled versions, is longer here (298 company types, only including those that exist in the offshore leaks database) because the data spans multiple corporate registers. For example, limited liability companies could be a Malaysian SDN BLD, a Luxembourgian SARL, or a British LTD. Having aligned company names, we match both data sources using name, type, and country of incorporation. This conservative approach can lead to false positive matches only if two firms with the identical name and the same incorporation type exist in the corporate register of one country. This is legally prohibited in the corporate register of the UK and we have not found such a duplicate in the tax havens for which we have data.

[12]These six jurisdictions are: Aruba, the Cook Islands, the Bahamas, Barbados, Malta, Nevis [part of St. Kitts and Nevis], and Samoa.

Since the country information is therefore important for the match with the leak data and the registers, we go through the country information in the offshore leaks database in detail. In some cases, the description of the jurisdiction is sub-national. We update the sub-national jurisdictions (such as Delaware, Texas, As Al Khaimah, or Dubai) to the country level (USA and United Arab Emirates for these cases). Instead of creating a synthetic Netherlands Antilles, we assign historical Netherlands Antilles cases to Curacao. Other updates concern spelling mistakes usually due to the mistreatment of special characters at some point of the data construction which we re-assign using regular expressions.

Preparing Offshore Officers Officers are the shareholders, directors, and beneficial owners that the offshore leak database identifies. These officer nodes can be linked to entity nodes with undirected edges supplied in the offshore leaks database. This is a many to many match as one officer can be connected to more than one entity and one entity can have more than one officer.

We add one important step to this link: The 'officer' assignment in the offshore leak database includes legal entities if they are entered as shareholders, beneficial owners, or directors. In order to distinguish those from legal entities, we employ our company type recognition code and classify any officer name in which we find a company type (such as "Ltd.") as a legal entity. Whenever we do not, we classify it as a natural person. We confirm the quality of this approach with manual inspection.

Matching Linked Entity-Officer Data with British Register Data The resulting combination of offshore shell companies with their beneficiaries can then be matched to our previously prepared address data. In order to do so, we rely on three prepared pieces of information: The name of a company, its incorporation type, and the country of incorporation. Our matching first uses the original capitalized names and types and the country of incorporation (28% of all matches). For the not yet matched data, we then proceed to match using the prepared company types and the country of incorporation (44% of all matches). Finally, we discard the company type and only match using the name stump and the country of incorporation (28% of all matches). We use all these uncertainty levels in the results of the paper and we have not detected non-randomness of the resulting country distributions across these three uncertainty levels.

Note that we do not use the ICIJ leak data to update our main dataset. A property will still be counted as offshore held for our main statistics even if we could identify the beneficial owner in the ICIJ leak. We analyze the ICIJ match in the respective section in detail.

A.2.3 Matching Price and Ownership Data

Having access to prepared postal address information for the transaction price dataset (Prices Paid data), as well as prepared addresses and company information for the corporate ownership datasets (CCOD, historical CCOD, OCOD, historical OCOD, our freedom of information request, and the Private Eye data) we then proceed to match the two. The central outcome of our address preparation is a unique property identifier. With this identifier, combining ownership and price data is straightforward.

Timing of Transactions and Owner Registrations. Beyond address information, the second ingredient of our match is information on the timing of real estate transactions. The price paid data registers the time that the transaction was recorded while the "date proprietor added" in the ownership data registers the date of entry of the new owner in the land register. Usually, an ownership change is recorded with a lag of 3-4 months in the land register in our data. Correspondence with the British land register confirmed this expected time lag. We double the tolerance and match transactions and ownership changes in a window of -1:8 months. When we observe an ownership change that is not matched to a transaction but we do observe the address in the price paid data at another time, we update ownership at the ownership change. This essentially introduces an assumption that a property that is residential at some point in time can always be treated as part of the residential market. This ensures that we do not miss ownership changes that were not due to a sale.

If a property was sold to a foreign company in 2010 and not sold since then, this sale will show up in the freedom of information data as a transaction. It will also show up in the historical OCOD stock data since the first entry still captures the current owner. It will also show up a few months later in the OCOD dataset. We introduce a data hierarchy going from the last available dataset to the first (OCOD $\rightarrow$ OCOD hist $\rightarrow$ FOI $\rightarrow$ Private Eye; CCOD $\rightarrow$ CCOD hist) and take all other auxiliary information from the data source of the highest point in the hierarchy where the match appears. Here, the Private Eye data is dominated by OCOD and its historical version (and is not registered separately in table 1.1). This is expected and a useful sanity check as the Private Eye data was originally

based on a freedom of information request on foreign property purchases: The same data that was used to construct OCOD in the first place. Our own FOI request does provide purchases of foreign companies that took place before the current owner at the start of the real time database was entered into the register, however. Whatever did not change owner between the start of real time OCOD data and the Private Eye data is registered in the first iteration of OCOD. Whatever did change ownership, is registered in the data we obtained through the FOI request on ownership changes since 1990.

Any deviation in address data quality will prohibit a match. False positives are therefore very unlikely. Our match should therefore be interpreted as a lower bound estimate.

A.2.4 Corporate Ownership Data including Non-Residential Properties

In the analysis of causes of offshore ownership in section 1.4, we make the analysis as broad as possible by focusing on the stock of both non-residential and residential property holdings. Since we are interested in the monthly stock of property titles held by companies registered in different jurisdictions, we use all ownership data after October 2015 for which we have regular snapshot information available. Since the frequency of the snapshots varies over time, we update the stock data using the date proprietor added information if the date proprietor added information is less than four months (six months) away from the snapshot date for snapshots after December 2017 (before December 2017). For instance, a property title held that appears in the snapshot of November 2017, but has a date proprietor added of October 2017 is also assigned to the stock of October 2017 (for which a regular snapshot is not available). The main specification focuses on property titles. As discussed above, since a property title can include multiple unique addresses, we also provide results for unique addresses.

A.3 LONG-RUN DYNAMICS OF THE OFFSHORE MARKET SHARE

To estimate the stock volume of foreign held residential real estate, we combine transaction information on all *residential* real estate from the Price Paid Data (PPD) with ownership information on *all corporate held* real estate from the land register (LR) and ORBIS. The confidence of this stock volume estimation varies over time. While after October 2015, regular snapshots of the land register are available, we only have limited information on

purchases by foreign companies between 1995 and October 2015. We thus discuss the estimation strategy for these two periods, i.e. 1995 to October 2015, and after October 2015 separately.

A.3.1 Estimation after October 2015

Given that snapshots of the land register are available after October 2015 on a regular basis, we can estimate the foreign held stock volume based on the information in each respective snapshot. For this purpose, we first merge the land register snapshot to PPD and supplement it with ownership information from ORBIS.[13]

For the estimation of the aggregate offshore market share at different points in time, we improve upon this data combination by making two adjustments. First, we also include transactions from the land register data that do not appear in the PPD, but whose transacted properties appeared at any point in the PPD. We thereby assume that a property is always either residential or non-residential but does not switch between these categories. Second, we impute the real estate stock value between the land register snapshots by using the transaction dates instead of the dates of the land register snapshot. Purchase prices are adjusted using district level house price indices (HPI). By using the transaction date information separately from the actual snapshot time, we correct for lagged appearances of properties in the land register.[14]

A.3.2 Estimation before October 2015

Before October 2015, we no longer have regular LR snapshots at our disposal. We therefore use all available purchase information before October 2015 to estimate the stock volume and shares of foreign held residential real estate between January 1995 and October 2015.[15]

Stocks. For every year-month t, we are interested in the stock value of real estate that is held by private owners (S_t^P), domestic corporate owners (S_t^D), foreign tax haven corporate

[13]For details on the combination of land register data, Price Paid Data, and ORBIS, see Appendix A.2.

[14]For example, a property might be registered in LR with a date proprietor added (DPA) of August 2018, but only appears in the snapshot of December 2018. Disentangling snapshot time from transaction time might thus increase the accuracy of our stock volume measure.

[15]As above, PPD information is supplemented with land register information on properties that appear at any time in PPD and are thus considered residential.

owners (S_t^{TH}), foreign non haven corporate owners (S_t^{NH}), and foreign corporate owners without country information (S_t^F). Stock estimates at time t refer to the end-of-month stock.

Flows. While we do not observe stocks directly, we observe at every year-month t the flow value F_t^{o-d} from origin o to the destination d. Consistent with the stock owner types, origin and destination can be P, D, TH, NH, F. In addition, the origin of a flow can also be New. Flows from New to one of the other owner type segments are identified from PPD information on whether a property is newly built or an established residential building.

Linking Stocks and Flows. The following equation shows exemplary for the stock of real estate held by tax havens, how stocks and flows can be linked. Note that we link stocks and flows separately for every district i.

$$\begin{aligned} S_{t,i}^{TH} = \; & \underbrace{(1+g_{t-1,i})}_{\text{HPI growth adjustment}} \quad \underbrace{S_{t-1,i}^{TH}}_{\text{Stock end of previous period}} \\ & \underbrace{-F_{t,i}^{TH-P} - F_{t,i}^{TH-D} - F_{t,i}^{TH-NH} - F_{t,i}^{TH-F}}_{\text{Flows out of tax havens into other owner type segments}} \\ & \underbrace{+F_{t,i}^{P-TH} + F_{t,i}^{D-TH} + F_{t,i}^{NH-TH} + F_{t,i}^{F-TH}}_{\text{Flows into tax havens from other owner type segments}} \\ & \underbrace{+F_{t,i}^{New-TH}}_{\text{New Properties}} \end{aligned} \tag{A.1}$$

Importantly, we observe foreign ownership of purchases only since 1995. As a consequence, we do not observe all flows with their origin o. Flows F thus consist of an observed and an unobserved component denoted by $\hat{F}$ and $\tilde{F}$ respectively. For instance, the flow from tax haven to private reads

$$F_{t,i}^{TH-P} = \underbrace{\hat{F}_{t,i}^{TH-P}}_{\text{observed}} + \underbrace{\tilde{F}_{t,i}^{TH-P}}_{\text{unobserved}} . \tag{A.2}$$

The unobserved part of the flow can be estimated, since we know if a property has existed already before the purchase. For instance, denote $F_{t,i}^{?-P}$ the inflow of property value into the private sector that has no previous owner and does not consist of newly built properties.

In order to assign the origin of these flows into destination d, we assume that the origins are proportional to the origins of flows into destination d for which we know the owner. For the example above, it follows that

$$\begin{aligned} \tilde{F}_{t,i}^{TH-P} &= \underbrace{\hat{F}_{t,i}^{?-P}}_{\text{flows to P, unobserved origin}} \times \; r_{y(t)}^{TH-P} \\ &= F_{t,i}^{?-P} \times \frac{\hat{F}_{y(t)}^{TH-P}}{\hat{F}_{y(t)}^{TH-P} + \hat{F}_{y(t)}^{NH-P} + \hat{F}_{y(t)}^{F-P} + \hat{F}_{y(t)}^{P-P}} \end{aligned} \tag{A.3}$$

Note that the ratio $r_{y(t)}^{TH-P}$ is calculated at the country level and at the year level y(t) to have a reasonably large sample size for the estimation of the shares.[16] Importantly, for the years after our first snapshot (2014, Private Eye data), any purchase of an already existing property without owner prior owner information must have had a private owner before this purchase. If the previous owner would have been corporate, we would have observed it in the stock data.

Combination and Inversion. We estimate the stocks at every point in time by inverting equation (A.1) and solving for the stocks going backward in time. This is possible, because we have a correct estimate of the stock of real estate with corresponding owner types in October 2015.[17] Combining the inverted equation (A.1) with equations (A.2) and (A.3), we obtain

[16]For the evolution of the aggregate market share, these choices do not change the results substantially as shown by Figure A.2 in which variations to the geographic level and period of transition probabilities are shown. Also note that we refrain from conditioning the transition probabilities on property characteristics, since our data is relatively scarce in this dimension.

[17]It is possible that the imputation yields negative stock values at the district level if we overestimate the inflows of properties based on the transition probabilities at a given point in time. However, as shown in Figure A.2, putting a zero lower bound threshold to the stock value at the district level in each backward imputation step does not influence the aggregate evolution of the market share.

$$
\begin{aligned}
S^{TH}_{t-1,i} = \frac{1}{(1+g_{t-1,i})} \Bigg[& S^{TH}_{t,i} \\
& + \hat{F}^{TH-P}_{t,i} + \hat{F}^{TH-D}_{t,i} + \hat{F}^{TH-NH}_{t,i} + \hat{F}^{TH-F}_{t,i} \\
& + \tilde{F}^{TH-P}_{t,i} + \tilde{F}^{TH-D}_{t,i} + \tilde{F}^{TH-NH}_{t,i} + \tilde{F}^{TH-F}_{t,i} \\
& - \hat{F}^{P-TH}_{t,i} - \hat{F}^{D-TH}_{t,i} - \hat{F}^{NH-TH}_{t,i} - \hat{F}^{F-TH}_{t,i} \\
& - \tilde{F}^{P-TH}_{t,i} - \tilde{F}^{D-TH}_{t,i} - \tilde{F}^{NH-TH}_{t,i} - \tilde{F}^{F-TH}_{t,i} \\
& - F^{New-TH}_{t,i} \Bigg].
\end{aligned}
\tag{A.4}
$$

Aggregation. In a final step, we aggregate the district level stock estimates and calculate the stock shares by ownership with respect to the total residential market denoted by s. For instance, the aggregate offshore tax haven share is calculated as

$$
s^{TH}_t = \frac{\sum_i S^{TH}_{t,i}}{\sum_i S^{TH}_{t,i} + \sum_i S^{NH}_{t,i} + \sum_i S^{F}_{t,i} + \sum_i S^{D}_{t,i} + \sum_i S^{P}_{t,i}}. \tag{A.5}
$$

A.3.3 Validation

To validate our results, we employ two validation exercises. First, we benchmark our estimate of the total volume of residential real estate against information from the national accounts. Second, we validate our imputation strategy by applying it to the period after October 2015 for which reliable estimates of the offshore market share growth rates exist.

Residential Market Volume. The UK national accounts[18] provide a yearly estimate of the capital stock value for dwellings (residential properties), other buildings and structures, and land since 1995 at the national level. We make two assumptions to make the time series of the national accounts comparable to our estimate. First, we allocate all land and land improvements to the residential market, but exclude other buildings and structures. Second, since the national accounts capture the entire UK, we scale the time series by the land share of England and Wales. Figure A.1a compares the values from the adjusted national accounts with our own estimation at the monthly level.

[18] https://www.ons.gov.uk/economy/nationalaccounts/uksectoraccounts/datasets/thenationalbalancesheetestimates

Offshore Market Share Growth. To validate our imputation strategy, we focus on the two most recent years in our data (October 2015 to December 2019). In this time period, we have credible longitudinal information on the offshore market share growth rates to which we can compare results from our imputation. We take the December 2019 value of real estate as given. We then deliberately ignore information about previous owners at the property level. Instead, we follow the imputation strategy above by classifying previous owners based on the year-specific transition probabilities calculated at the aggregate level. Figure A.1b compares the resulting month-to-month growth rates for the tax haven share in residential real estate to its actual values.

A.4 TAXATION OF CAPITAL GAINS

Capital gains tax (CGT) in the UK was introduced in 1965 and applies to profits resulting from selling, swapping or transferring an asset which increased in value. With respect to real estate holdings, the capital gains tax applies to all real estate property that is (i) not the primary residence of the taxpayer, or (ii) the primary residence of the taxpayer, but is let, used for business, or is larger than 5,000 square meters. While capital gains valuation are usually based the difference between the purchase and sales price, market valuations can be applied if assets are transferred at prices below their market value. Capital gains tax applies to net gains across all taxable assets, i.e. losses from one asset can be deducted from the gains of another asset. Both the tax rates and the tax-free allowances for the capital gains tax vary with the type of investor and the income level of the taxpayer.[19]

Historically, foreign investors—both individuals and corporate—were mostly exempt from capital gains tax.[20] In April 2015, CGT rules were extended to foreign individuals, close companies (companies controlled by five individuals or less), trusts and personal representatives disposing UK residential property. However, both the disposal of UK residential property by diversely held foreign companies and the disposal of non-residential property remained exempt.

The 2017 Policy Change. In November 2017, the government announced to remove these remaining advantages for non-resident investors:

[19]For detailed information on the capital gains tax, see https://www.gov.uk/browse/tax/capital-gains.

[20]From 2013 to 2015, CGT did apply to foreigners subject to the Annual Tax on Enveloped Dwellings (ATED).

> *To align the UK with other countries and remove an advantage which non-residents have over UK residents, all gains on non-resident disposals of UK property will be brought within the scope of UK tax. This will apply to gains accrued on or after April 2019.*
>
> Autumn Budget 2017, p. 35

The details of this policy change were further specified during a consultation process between November 2017 and July 2018. The final draft legislation included provisions to extend CGT coverage to both direct disposals of non-residential properties and indirect disposals of residential property through "property-rich" corporations. In particular, indirect disposals through shares became subject to CGT if the company derives at least 75 percent of its gross asset value from UK real estate and if the person making the disposal held at anytime at least a 25 percent investment in the company during the last two years.[21] Table A.1 summarizes the tax rates applicable to foreigners on gains accrued after April 2019.

The Luxembourg Exemption. At the time of announcement, the new regulation interacted with existing double tax treaties, most notably with the double tax treaty between the UK and Luxembourg. While taxing rights regarding the direct disposal of property are usually not affected by double tax treaties, the UK was restricted in taxing the gains made by foreign investors via shares in UK property rich entities. Although the

[21] While the UK government explained that both the asset value test and the ownership test aggregate disposals of multiple entities and ownership shares held by certain persons through a series of entities, it is unclear how this provision is implemented, if beneficial owners are concealed through offshore structures. Additional exemptions to this rule include UK property used by the UK branch or agency of a foreign company, property held by overseas pension schemes, sovereign wealth funds and charities. For details, see information on the consultation process and the regulatory framework on the capital gains tax for non-residents.

Table A.1. – Capital Gains Tax Rates for Foreigners (April 2019)

Investor Type	Non-Residential Property	Residential Property
Individuals	10% / 20%	18% / 28%
Companies	19%	19%
Trustees	20%	28%

Note: The table displays capital gains tax rates applicable to non-resident individuals, companies, and trustees applicable to capital gains as of April 2019. Non-residential properties include commercial property, for example shops or offices, agricultural land, forests, and any other land or property which is not used as a residence. Residential properties include buildings used or suitable for use as a dwelling as well as the underlying garden or grounds associated with such buildings. Marginal tax rates for individuals differ, since individuals pay lower rates within the basic income tax band and a higher rate on any amount above the basic income band.
Source: HM Revenue & Customs.

UK government was aware of this loophole and enacted general anti-forestalling rules, it was and is still unclear, how the exact coverage and the enforcement of these rules will look like.[22].

As an additional measure against "treaty shopping", the UK implemented the OECD multilateral convention to implement tax treaty related measures to prevent base erosion and profit shifting in October 2018. This multilateral amendment included a particular clause on indirect disposals of "property-rich" entities:

> *For purposes of a Covered Tax Agreement, gains derived by a resident of a Contracting Jurisdiction from the alienation of shares (...) may be taxed in the other Contracting Jurisdiction if (...) these shares or comparable interests derived more than 50 per cent of their value directly or indirectly from immovable property.*
>
> OECD Multilateral Amendment, Article 9, Paragraph 4

This multilateral amendment allowed the UK to regain the rights to capital gains through the indirect disposal of immovable property. However, Luxembourg ratified the amendment only in April 2019 and the corresponding changes on capital gains tax formally apply only to capital gains after April 2020.[23] In mid 2022, the UK and Luxembourg signed a new double taxation treaty that directly integrates this treatment of the indirect disposals of "property-rich" entities:

> *Gains derived by a resident of a Contracting State from the alienation of shares or comparable interests, such as interests in a partnership or trust, deriving more than 50 per cent of their value directly or indirectly from immovable property (...) situated in the other Contracting State may be taxed in that other State.*
>
> 2022 UK-Luxembourg Double Taxation Convention, Article 13, Paragraph 2

This tax treaty came into force at the end of 2023.

[22]See for instance a comment by the law firm Goodwin Procter, but also a more recent assessment by the law firm DLA Piper stating that "UK anti-forestalling rule (...), depending on the facts, could be circumvented."

[23]See information by HM Revenue & Customs on Tax Treaties between the UK and Luxembourg.

A.5 PUBLIC OWNERSHIP REGISTERS

At least since 2016, the installation of ownership registers and the exchange of the corresponding beneficial ownership information has become a core priority of governments around the world to combat illicit financial flows including terrorist financing, money laundering and tax evasion.[24]

While ownership registers have been in place in some of the Overseas Territories (OT) and the Crown Dependencies (CD) for a long time, the access to these registers (if available) was historically very restricted. Even the UK government could not access the recorded information until a series of bilateral agreements was negotiated between the UK government and the OT and CD authorities in 2016.[25] These agreements were supposed to give law enforcement authorities fast and secure access to beneficial ownership information on companies incorporated in OT and CD countries. Importantly, however, OT and CD did not plan to make the ownership information *publicly* available at that time.

The 2018 Policy Announcement. The announcement of public corporate registers has its roots in an amendment to the *Sanctions and Anti-Money Laundering Act* (SAMLA) that entered the parliamentary process already at the beginning of 2018.[26] While the government first opposed the amendment, it became apparent in April 2018 that it had enough supporters to defeat the government.[27] The amendment was approved in the House of Commons on May 1st 2018.[28]

According to the amendment, the UK Secretary of State shall provide assistance to the Overseas Territories in setting up a publicly accessible register of the beneficial ownership of companies registered in each government's jurisdiction. In the case that OT governments fail to install public registers until the end of 2020, the UK government is required by law to formally mandate such a public register:

> *The Secretary of State must, no later than 31 December 2020, prepare a draft Order in Council requiring the government of any British Overseas Territory*

[24] See for instance the joint statement by multiple countries on the systematic sharing of beneficial ownership information and the implementation Financial Action Task Force standards.

[25] See the collection on the bilateral agreements between the UK and different OT and CD jurisdictions.

[26] See the proposal of Amendment 73 to SAMLA at the report stage of the bill.

[27] See for instance a Guardian newspaper article on the SAMLA amendment.

[28] See details on the legislative implementation procedure of SAMLA.

> *that has not introduced a publicly accessible register of the beneficial ownership of companies within its jurisdiction to do so.*
>
> Sanctions and Anti-Money Laundering Act 2018, Section 51, Subsection 4

While the law itself does not specify a particular date at which the registers are expected to be accessible, a Joint Ministerial Council explained in December 2018 that the mandate would request the registers to be operational at the end of 2023.

The Exclusion of Crown Dependencies. While OT were supposed to be directly mandated to set up public beneficial ownership registers in the future, CD were excluded from SAMLA, because their constitutional status implies a higher degree of autonomy. In particular, Guernsey, Jersey and the Isle of Man elect their own legislatures and run their own domestic legislation. Although a formal mandate is not possible, CD have announced in June 2019 to set up public registers voluntarily:

> *Within 12 months of [the publication of the Implementation Review of the 5th AMLD in January 2022], we will each bring forward to our own parliament legislative proposals to establish public access to beneficial ownership data of companies held on a central register.*
>
> Joint commitment by Guernsey, Jersey and the Isle of Man

At the end of 2020, when the OT mandate was supposed to be issued, the government published the draft Order in Council.[29]. It is noteworthy, however, that the UK government did not see the need to officially bring this mandate into force given the "firm commitments from all of the inhabited Overseas Territories to adopt publicly accessible registers".[30]

[29] See draft Order in Council.

[30] See Ministerial Statement on publicly accessible registers in the UK Overseas Territories.

Appendix to Chapter 2

B.1 APPENDIX FOR THE THEORY PART

B.1.1 Proofs of Theorem 1 and Corollary 1

We focus on the statements in Theorem 1 that involve conditional revenue functions for dual earner couples. The extension to single-earner couples follows from the same arguments and is therefore omitted. Recall that a reform in direction $h \in \mathcal{H}$ is Pareto-improving if

$$R_\tau(0,h) \; - \; \max_{y \in \mathcal{Y}} \; h(y) \quad \geq \quad 0 \; . \tag{B.1}$$

A reform in direction $-h \in \mathcal{H}$ is Pareto-improving if

$$R_\tau(0,h) \; - \; \min_{y \in \mathcal{Y}} \; h(y) \quad \leq \quad 0 \; . \tag{B.2}$$

Let $(i,j) \in \{(1,2),(2,1)\}$. Consider a simple reform h so that

$$h(y) \quad = \quad h_i(y_i) \; \mathbf{1}(y_j \in [y_{js}, y_{js} + \ell_j]) \; , \quad \text{where}$$

$$h_i(y_i) \quad = \quad \begin{cases} 0, & \text{if} \quad y_i \leq y_{is} \\ y_i - y_{is} & \text{if} \quad y_{is} \leq y \leq y_{is} + \ell_i \\ \ell_i, & \text{if} \quad y_{is} \geq y_{is} + \ell_1 \end{cases}$$

It follows from (B.1) and (B.2) that a reform in direction h is Pareto-improving if

$$R_\tau(0,h) \quad > \quad \ell_i \; . \tag{B.3}$$

A reform in direction $-h$ is Pareto-improving if

$$R_\tau(0,h) \quad < \quad 0 \; . \tag{B.4}$$

Lemma 1. *If, for some B_j, $\mathcal{R}_i(y_i \mid B_j) > 1$, there exist y_{is} and $\ell_i > 0$ so that (B.3) holds. If $\mathcal{R}_i(y_i \mid B_j) < 0$, there exist y_{is} and $\ell_i > 0$ so that (B.4) holds.*

Proof. For ℓ_i close to zero, we have that[1] $R_\tau(0,h) \simeq \ell_i \mathcal{R}_i(y_i \mid B_j)$. Therefore, for ℓ_i close to zero, (B.3) holds if $\ell_i \mathcal{R}_i(y_i \mid B_j) > \ell_i$, or, equivalently, if $\mathcal{R}_i(y_i \mid B_j) > 1$. Analogously, for ℓ_i close to zero, (B.4) holds if $\ell_i \mathcal{R}_i(y_i \mid B_j) < 0$, or, equivalently, if $\mathcal{R}_i(y_i \mid B_j) < 0$. □

Lemma 2. *If, for some B_j, $\underline{y}_i < \overline{y}_i$ and $\mathcal{R}_i(\overline{y}_i \mid B_j) > \mathcal{R}_i(\underline{y}_i \mid B_j)$, there is a Reform-direction h so that (B.1) holds.*

Proof. We define

$$h = \underline{\tau}_i \, \underline{h}_i(y_i) \, \mathbf{1}(y_j \in B_j) + \overline{\tau}_i \, \overline{h}_i(y_i) \, \mathbf{1}(y_j \in B_j) \, , \quad \text{where}$$

$$\underline{h}_i(y_i) = \begin{cases} 0, & \text{if } \; y_i \le \underline{y}_i \\ y_i - \underline{y}_i & \text{if } \; \underline{y}_i \le y \le \underline{y}_i + \underline{\ell}_i \\ \underline{\ell}_i, & \text{if } \; y_{is} \ge \underline{y}_i + \ell_i \end{cases}$$

and

$$\overline{h}_i(y_i) = \begin{cases} 0, & \text{if } \; y_i \le \overline{y}_i - \overline{\ell}_i \\ y_i - \overline{y}_i & \text{if } \; \overline{y}_i - \overline{\ell}_i \le y \le \overline{y}_i \\ \ell_i, & \text{if } \; y_{is} \ge \underline{y}_i \, . \end{cases}$$

We also choose $\underline{\tau}_i < 0$, $\overline{\tau}_i > 0$ and $\underline{\tau}_i + \overline{\tau}_i = 0$. We finally let $\underline{\ell}_i = \overline{\ell}_i =: \ell_i$. This construction ensures that, conditional on $y_j \in B_j$,

$$h(y) \begin{cases} = 0, & \text{if } \; y_1 \le \underline{y}_1 \\ < 0, & \text{if } \; \underline{y}_1 < y < \overline{y}_1 \\ = 0, & \text{if } \; y_1 \ge \overline{y}_1 \, . \end{cases}$$

so that $\max_y h(y) = 0$. Thus, to verify that (B.1) holds, it remains to be shown that $R_\tau(0,h) > 0$. To see this, note that by the linearity of the Gateaux differential, for ℓ_i close to zero, we have

$$\begin{aligned} R_\tau(0,h) &= \underline{\tau}_i \, \ell_i \, \mathcal{R}_i(\underline{y}_i \mid B_j) + \overline{\tau}_i \, \ell_i \, \mathcal{R}_i(\overline{y}_i \mid B_j) \\ &> \underline{\tau}_i \, \ell_i \, \mathcal{R}_i(\underline{y}_i \mid B_j) + \overline{\tau}_i \, \ell_i \, \mathcal{R}_i(\underline{y}_i \mid B_j) \\ &= 0 \, . \end{aligned}$$

[1] With $\ell_i = 0$, we have that $T^0(y) = T^1(y)$, for all y and hence $y^*(\theta,\tau,h) = y^0(\theta)$. There is neither a change of the tax system nor of behavior, hence there is no change of tax revenue so that $R_\tau(\tau,h) = R_i(y_{is},0,y_{js},\ell_j) = 0$, for all $\tau \ge 0$. In particular, $R_i(0,y_{is},y_{js},\ell_j) = 0$. A first order Taylor approximation, moreover yields

$$R_\tau(0,h) \simeq R_i(y_{is},0,y_{js},\ell_j) + \ell_i \mathcal{R}_i(y_i \mid B_j) = \ell_i \mathcal{R}_i(y_i \mid B_j) \, .$$

□

Lemmas 1 and 2 imply Theorem 1. If a given tax system is Pareto-efficient, then there must not exist a Pareto-improving reform direction. The following Lemma which establishes Corollary 1 exploits this property to give necessary conditions for the Pareto-efficiency of a tax system.

Lemma 3. *If there is no Pareto-improving reform direction, then the following holds, for all* $(i,j) \in \{(1,2),(2,1)\}$ *and for all* $y \in \mathcal{Y}$*: The function* $y_i \mapsto \mathcal{R}_i(y_i \mid y_j)$ *is non-increasing.*

$$\mathcal{R}_i(y_i \mid y_j) \quad \geq \quad 0\,. \tag{B.5}$$

There is a number $\bar{\ell}_j$*, so that, for all* $\ell_j < \bar{\ell}_j$*,*

$$\ell_j\, \mathcal{R}_i(y_i \mid y_j) \quad \leq \quad 1\,. \tag{B.6}$$

Proof. All statements follow from the observation that, for ℓ_j close to zero,

$$\mathcal{R}_i(y_i \mid B_j[y_{js}, \ell_j]) \quad \simeq \quad \ell_j\, \mathcal{R}_i(y_i \mid y_j)\,.$$

Therefore, if (B.6) was violated, there would exists ℓ_j so that $\mathcal{R}_i(y_i \mid B_j[y_{js}, \ell_j]) > 1$. thereby contradicting the non-existence of a Pareto-improving reform direction, see Proposition 1. If (B.5) was violated, there would exist ℓ_j so that $\mathcal{R}_i(y_i \mid B_j[y_{js}, \ell_j]) < 0$, again contradicting the non-existence of a Pareto-improving reform direction, see Proposition 1. If the function $y_i \mapsto \mathcal{R}_i(y_i \mid y_j)$ was increasing, this would imply the existence of ℓ_j and $\underline{y}_i < \overline{y}_i$ with $\mathcal{R}(\overline{y}_i \mid B_j[y_{js}, \ell_j]) > \mathcal{R}_i(\underline{y}_i \mid B_j[y_{js}, \ell_j])$, once more contradicting the non-existence of a Pareto-improving reform direction, see Proposition 2. □

B.1.2 Proof of Theorem 2

Outline. Let $h \in \mathcal{H}$ be a continuous function. We approximate h by a combination of simple tax reforms h_m and show that

$$R_\tau(0,h) \;-\; \max_{y \in \mathcal{Y}}\; h_m(y) \quad \leq \quad 0\,. \tag{B.7}$$

We then extend the argument to the function h.

Constructing h_m.

Consider a partition of $\mathcal{Y}_1$ in m adjacent brackets of equal length ℓ_1, indexed by $k_1 \in K_1 = \{1, 2, \ldots, m\}$, with bracket $B^1_{k_1}$ starting at $(k_1 - 1)\ell_1$ and ending at $k_1\ell_1$. A partition of $\mathcal{Y}_2$ in m adjacent intervals of equal length ℓ_2 is defined analogously, with brackets indexed by $k_2 \in K_2 = \{1, 2, \ldots, m\}$.

A simple tax reform that changes marginal tax rates in the interior of $\mathcal{Y}_1 \times \mathcal{Y}_2$ can now be identified by $k \in K = K_1 \times K_2$ and by the changes in marginal tax rates τ^1_k and τ^2_k applying, respectively, to the earnings of spouse 1 and spouse 2. For such a reform we write

$$h_k(y_1, y_2) \quad = \quad \tau^1_k \; h^1_{k_1+1}(y_1) \; \mathbf{1}(y_2 \in B^2_{k_2+1}) \; + \; \tau^2_k \; h^2_{k_2+1}(y_2) \; \mathbf{1}(y_1 \in B^1_{k_1+1}) \; ,$$

where

$$h^1_{k_1+1}(y_1) = \begin{cases} 0, & \text{if} \quad y_1 \leq k_1 \, \ell_1 \; , \\ y_1 - k_1 \, \ell_1, & \text{if} \quad k_1 \, \ell_1 \leq y_1 \leq (k_1 + 1)\ell_1 \; , \\ \ell_1, & \text{if} \quad y_1 \geq (k_1 + 1)\ell_1 \; , \end{cases}$$

and analogously for $h^2_{k_2+1}$. We write

$$\begin{aligned} h^{int}(y_1, y_2) \quad &:= \quad \textstyle\sum_{k \in K} h^k(y_1, y_2) \\ &= \quad \textstyle\sum_{k \in K} \tau^1_k \; \bar{h}^k_1(y_1, y_2) \; + \; \sum_{k \in K} \tau^1_k \; \bar{h}^k_2(y_1, y_2) \end{aligned}$$

for a reform that combines all these simple tax reforms, where

$$\bar{h}^1_k(y_1, y_2) \quad := \quad h^1_{k_1+1}(y_1) \; \mathbf{1}(y_2 \in B^2_{k_2+1}) \quad \text{and}$$

$$\bar{h}^2_k(y_1, y_2) \quad := \quad h^2_{k_2+1}(y_2) \; \mathbf{1}(y_1 \in B^1_{k_1+1}) \; .$$

Simple tax reforms that change marginal tax rates on the boundary of $\mathcal{Y}_1 \times \mathcal{Y}_2$ are identified by the index k_1 (for the segment with $y_2 = 0$) or by the index k_2 (for the segment with $y_1 = 0$). The change in the marginal tax rate applying to the spouse with positive earnings is denoted either by $\tau^1_{k_1}$ or $\tau^2_{k_2}$. We write

$$h^{b1}(y_1, y_2) \quad := \quad \sum_{k_1 \in K_1} \tau^1_{k_1} \bar{h}^{b1}_{k_1}(y_1, y_2)$$

with

$$\bar{h}^{b1}_{k_1}(y_1, y_2) \quad := \quad h^1_{k_1+1}(y_1) \mathbf{1}(y_2 = 0)$$

for a combination of simple tax reforms that affect the marginal tax rates of spouse 1, conditional on spouse 1 being the couple's single earner. We define the functions $\bar{h}^{b2}_{k_2}$ analogously. We now combine all these simple reforms into a reform direction $h_m : (y_1, y_2) \to h_m(y_1, y_2)$ with

$$h_m(y_1, y_2) \quad = \quad h^{int}(y_1, y_2) + h^{b1}(y_1, y_2) + h^{b2}(y_1, y_2)$$

Note that the reform direction h_m depends on how fine the partition of $\mathcal{Y}_1 \times \mathcal{Y}_2$ into m^2 squares is. Finally, to complete the construction, for $k = (k_1, k_2)$, we define the changes of marginal tax rates stipulated by h_m with respect to the function h that h_m approximates,

$$\frac{h(y_1^{k_1+1}, y_2^{k_2}) - h(y_1^{k_1}, y_2^{k_2})}{\ell_1} \quad =: \quad \tau_k^1$$

and

$$\frac{h(y_1^{k_1}, y_2^{k_2+1}) - h(y_1^{k_1}, y_2^{k_2})}{\ell_2} \quad =: \quad \tau_k^2 \,,$$

where $y_1^{k_1} := k_1 \ \ell_1$ and $y_2^{k_2} := k_2 \ \ell_2$. For later reference, note that this implies, in particular, that

$$\lim_{\ell_1 \to 0} \frac{h(y_1^{k_1+1}, y_2^{k_2}) - h(y_1^{k_1}, y_2^{k_2})}{\ell_1} \quad = \quad h_1(y_1^{k_1}, y_2^{k2}) \quad = \quad \tau_k^1$$

and

$$\lim_{\ell_2 \to 0} \frac{h(y_1^{k_1}, y_2^{k_2+1}) - h(y_1^{k_1}, y_2^{k_2})}{\ell_2} \quad = \quad h_2(y_1^{k_1}, y_2^{k2}) \quad = \quad \tau_k^2 \,.$$

The Relation between h and h_m. Consider the partition of the set $\mathcal{Y}_1 \times \mathcal{Y}_2$ into m^2 squares as described above. Thus, square $k = (k_1, k_2)$ has endpoints $(y_1^{k_1}, y_2^{k_2})$, $(y_1^{k_1+1}, y_2^{k_2})$, $(y_1^{k_1}, y_2^{k_2+1})$, and $(y_1^{k_1+1}, y_2^{k_2+1})$ where $y_1^{k_1} = k_1\ell_1$, $y_1^{k_1+1} = (k_1+1)\ell_1$, $y_2^{k_2} = k_2\ell_2$, $y_2^{k_2+1} = (k_2+1)\ell_2$. We now argue that if $h(0,0) = h_m(0,0)$, then the functions h and h_m coincide at the corners of any square. To see this, let $h(y_1^{k_1}, y_2^{k_2}) = h_m(y_1^{k_1}, y_2^{k_2})$. We now show that this implies that also

$$h_m(y_1^{k_1+1}, y_2^{k_2+1}) = h(y_1^{k_1+1}, y_2^{k_2+1})$$

Note that

$$\begin{aligned} h_m(y_1^{k_1+1}\ell_1, y_2^{k_2+1}) &= h_m(y_1^{k_1}, y_2^{k_2}) \;+\; h_m(y_1^{k_1+1}, y_2^{k_2+1}) - h_m(y_1^{k_1+1}, y_2^{k_2}) \\ &\qquad + \; h_m(y_1^{k_1+1}, y_2^{k_2}) - h_m(y_1^{k_1}, y_2^{k_2}) \\ &= h(y_1^{k_1}, y_2^{k_2}) \;+\; \tau_2^{k_1+1,k_2}\,\ell_2 \;+\; \tau_1^{k}\,\ell_1 \\ &= h(y_1^{k_1}, y_2^{k_2}) \;+\; h(y_1^{k_1+1}, y_2^{k_2+1}) - h(y_1^{k_1+1}, y_2^{k_2}) \\ &\qquad + \; h(y_1^{k_1+1}, y_2^{k_2}) - h(y_1^{k_1}, y_2^{k_2}) \\ &= h(y_1^{k_1+1}, y_2^{k_2+1})\;. \end{aligned}$$

Revenue Implications of h_m.

We seek to characterize the Gateaux differential of tax revenue in direction h_m, $R_\tau(0, h_m)$. Using that

$$\begin{aligned} h_m(y_1, y_2) &= h^{int}(y_1, y_2) + h^{b1}(y_1, y_2) + h^{b2}(y_1, y_2) \\ &= \textstyle\sum_{k\in K} \tau_k^1\; \bar{h}_k^1(y_1, y_2) \;+\; \sum_{k\in K} \tau_k^2\; \bar{h}_k^2(y_1, y_2) \\ &\qquad + \textstyle\sum_{k_1\in K_1} \tau_{k_1}^1 \bar{h}_{k_1}^{b1}(y_1, y_2) \quad + \sum_{k_2\in K_2} \tau_2^{k_2} \bar{h}_{k_2}^{b2}(y_1, y_2) \end{aligned}$$

and the linearity of the Gateaux differential, we obtain

$$\begin{aligned} R_\tau(0, h_m) &= \textstyle\sum_{k\in K} \tau_k^1\; R_\tau(0, \bar{h}_1^k) + \sum_{k\in K} \tau_k^2\; R_\tau(0, \bar{h}_2^k) \\ &\qquad + \textstyle\sum_{k_1\in K_1} \tau_{k_1}^1 R_\tau(0, \bar{h}_{k_1}^{b1}) \quad + \sum_{k_2\in K_2} \tau_{k_2}^2 R_\tau(0, \bar{h}_{k_2}^{b2})\;. \end{aligned}$$

Large m. We have partitioned $\mathcal{Y}_1 \times \mathcal{Y}_2$ into m^2 squares. For $\ell_1 = \frac{\bar{y}_1}{m}$ and $\ell_2 = \frac{\bar{y}_2}{m}$ close to zero it follows from first order Taylor approximations that

$$R_\tau(0, \bar{h}_k^1) \quad \simeq \quad \ell_1\; \ell_2\; \mathcal{R}_1(y_1^{k_1} \mid y_2^{k_2})\;,$$

where $y_1^{k_1} = \ell_1\; k_1$ and $y_2^{k_2} = \ell_2\; k_2$. The approximation is perfect in the limit as $m \to \infty$. Analogously, we have that

$$R_\tau(0, \bar{h}_k^2) \quad \simeq \quad \ell_1\; \ell_2\; \mathcal{R}_2(y_2^{k_2} \mid y_1^{k_1})\;,$$

$$R_\tau(0, \bar{h}_{k_1}^{b1}) \quad \simeq \quad \ell_1\; \mathcal{R}_1^s(y_1^{k_1})\;, \quad \text{and}$$

$$R_\tau(0, \bar{h}_{k_2}^{b2}) \quad \simeq \quad \ell_2\; \mathcal{R}_2^s(y_2^{k_2})\;.$$

Thus, for $\ell_1 = \frac{\bar{y}_1}{m}$ and $\ell_2 = \frac{\bar{y}_2}{m}$ close to zero and by the linearity of the Gateaux differential, we obtain

$$\begin{array}{rcl} R_\tau(0, h_m) & \simeq & \sum_{k\in K} \tau_k^1 \; \ell_1 \; \ell_2 \; \mathcal{R}_1(y_1^{k_1} \mid y_2^{k_2}) \quad + \sum_{k\in K} \tau_k^2 \; \ell_1 \; \ell_2 \; \mathcal{R}_2(y_2^{k_2} \mid y_1^{k_1}) \\ & & + \sum_{k_1\in K_1} \tau_{k_1}^1 \ell_1 \; \mathcal{R}_1^s(y_1^{k_1}) \quad + \sum_{k_2\in K_2} \tau_{k_2}^2 \ell_2 \; \mathcal{R}_2^s(y_2^{k_2}) \; . \end{array} \tag{B.8}$$

Proof that h_m is not a Pareto-Improving Direction

Rewriting h_m. For later reference note that

$$\sum_{k\in K} \tau_k^1 \; \bar{h}_k^1(y_1, y_2) \;\; = \;\; \sum_{k_2=0}^{K_2} \sum_{k_1=0}^{K_1} \tau_{k_1,k_2}^1 \; \bar{h}_{k_1,k_2}^1(y_1, y_2)$$

$$\sum_{k\in K} \tau_2^k \; \bar{h}_k(y_1, y_2)^2 \;\; = \;\; \sum_{k_1=0}^{K_1} \sum_{k_2=0}^{K_2} \tau_{k_1,k_2}^2 \; \bar{h}_{k_1,k_2}^2(y_1, y_2)$$

so that we can write

$$\begin{array}{rcl} h_m(y_1, y_2) & = & \sum\limits_{k_2\in K_2} \sum\limits_{k_1\in K_1} \tau_{k_1,k_2}^1 \; \bar{h}_{k_1,k_2}^1(y_1, y_2) \quad + \sum\limits_{k_1\in K_1} \sum\limits_{k_2\in K_2} \tau_{k_1,k_2}^2 \; \bar{h}_{k_1,k_2}^2(y_1, y_2) \\ & & + \sum\limits_{k_1\in K_1} \tau_{k_1}^1 \; \bar{h}_{k_1}^{b1}(y_1, y_2) \quad + \sum\limits_{k_2\in K_2} \tau_{k_2}^2 \; \bar{h}_{k_2}^{b2}(y_1, y_2) \; . \end{array} \tag{B.9}$$

Focusing on Corners. We seek to show that

$$R_\tau(0, h_m) \; - \max_{y\in\mathcal{Y}} \; h_m(y) \quad \leq \quad 0 \; . \tag{B.10}$$

In doing so we can, without loss of generality, limit attention to the values that the function h_m achieves at the corners of the squares in Figure B.1: The function h_m is piecewise linear and therefore achieves its maximum at a corner. Thus, to establish that (B.10) holds, we show that for all $k_1 \in K_1$ and all $k_2 \in K_2$

$$R_\tau(0, h_m) \; - h_m(y_1^{k_1}, y_2^{k_2}) \quad \leq \quad 0 \; . \tag{B.11}$$

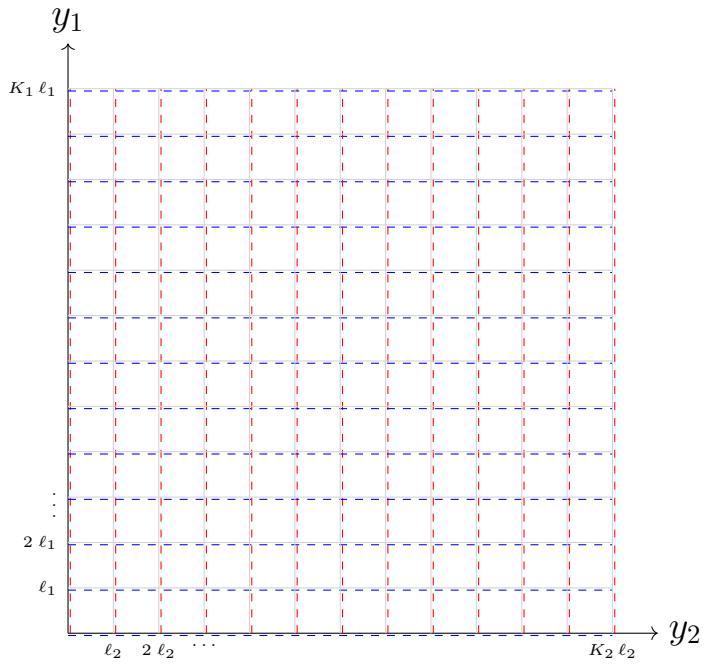
Using (B.8) and (B.9) inequality (B.11) can also be written as

$$A \; + \; B \; + \; C \; + \; D \quad \leq \quad 0 \, , \tag{B.12}$$

where

$$\begin{array}{rcl} A & = & \sum\limits_{k_2\in K_2} \sum\limits_{k_1\in K_1} \tau_{k_1,k_2}^1 \; \ell_1 \; \ell_2 \; \mathcal{R}_1(y_1^{k_1} \mid y_2^{k_2}) \\ & & - \sum\limits_{k_2\in K_2} \sum\limits_{k_1\in K_1} \tau_{k_1,k_2}^1 \; \bar{h}_{k_1,k_2}^1(y_1^{k_1}, y_2^{k_2}) \; , \end{array}$$

Figure B.1. – Summing over all k



Note: A point y_k on the grid has coordinates $y_{k1} = k_1\,\ell_1$ and $y_{k2} = k_2\,\ell_2$. In summing over all k, we can either go from one red line to the next and sum over all points, or go from one blue line to the next and sum over all points. The summation in (B.9) does the former for simple reforms that affect the marginal tax rates for spouse 1 and the latter for simple reforms that affect the marginal tax rates for spouse 2.

$$
\begin{aligned}
B &= \sum_{k_1 \in K_1} \sum_{k_2 \in K_2} \tau^2_{k_1,k_2}\, \ell_1\, \ell_2\, \mathcal{R}_2(y_2^{k_2} \mid y_1^{k_1}) \\
&\quad - \sum_{k_1 \in K_1} \sum_{k_2 \in K_2} \tau^2_{k_1,k_2}\, \bar{h}^2_{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \,, \\
C &= \textstyle\sum_{k_1 \in K_1} \tau^1_{k_1}\, \ell_1\, \mathcal{R}^s_1(y_1^{k_1}) \quad - \sum_{k_1 \in K_1} \tau^1_{k_1}\, h^1_{k_1+1}(y_1^{k_1}) \mathbf{1}(y_2 = 0) \,,
\end{aligned}
$$

and

$$
D = \textstyle\sum_{k_2 \in K_2} \tau^2_{k_2} \ell_2\, \mathcal{R}^s_2(y_2^{k_2}) \quad - \sum_{k_2 \in K_2} \tau^2_{k_2}\, h^2_{k_2+1}(y_2^{k_2}) \mathbf{1}(y_1 = 0) \,.
$$

The inequality in (B.11) follows form the observations that, under the conditions listed in Theorem 1, $A \leq 0$, $B \leq 0$, $C \leq 0$, and $D \leq 0$. In the following we show that $A \leq 0$. $B, C, D \leq 0$ can be shown by a straightforward adaptation of the arguments used to establish that $A \leq 0$.

Lemma 4. *Under the conditions listed in Corollary 1,*

$$
\sum_{k_2=0}^{K_2} \sum_{k_1=0}^{K_1} \tau_1^{k_1,k_2}\, \ell_1\, \ell_2\, \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \;\leq\; \sum_{k_2=0}^{K_2} \sum_{k_1=0}^{K_1} \tau_1^{k_1,k_2}\, h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \,.
$$

Proof. Fix some $k_2 \in \{0, 1, \ldots, K_2\}$. It suffices to show that for any such k_2,

$$\sum_{k_1=0}^{K_1} \tau^1_{k_1,k_2}\ \ell_1\ \ell_2\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad \sum_{k_1=0}^{K_1} \tau^1_{k_1,k_2}\ h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2})\ . \tag{B.13}$$

Step 1. Define the auxiliary function H_1

$$H_1(j_1) \quad := \quad \sum_{k_1=0}^{j_1} \tau^1_{k_1,k_2}\ h^1_{k_1,k_2}(y_1^{k_1}, y_2^{k_2})$$

and let $j_1^* \in \operatorname{argmax}\ H_1(j_1)$. Note that

$$\sum_{k_1=j_1^*}^{j^*} \tau^1_{k_1,k_2}\ h^1_{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \quad = \quad \tau^1_{j_1^*,k_2}\ \ell_1$$

and that

$$\tau^1_{j_1^*,k_2}\ell_1\ \ell_2\ \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \quad \leq \quad \tau^1_{j_1^*,k_2}\ \ell_1$$

since, by assumption, $\ell_2\ \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \leq 1$. Thus we have that

$$\sum_{k_1=j_1^*}^{j^*} \tau^1_{k_1,k_2}\ell_1\ \ell_2\ \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \quad \leq \quad \sum_{k_1=j_1^*}^{j^*} \tau^1_{k_1,k_2}\ h^1_{k_1,k_2}(y_1^{k_1}, y_2^{k_2})\ . \tag{B.14}$$

Step 2. We now argue that also

$$\sum_{k_1=j_1^*-1}^{j^*} \tau^1_{k_1,k_2}\ell_1\ \ell_2\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad \sum_{k_1=j_1^*-1}^{j_1^*} \tau_1^{k_1,k_2}\ h^1_{k_1,k_2}(y_1^{k_1}, y_2^{k_2})\ . \tag{B.15}$$

Note that

$$\sum_{k_1=j_1^*-1}^{j_1^*} \tau^1_{k_1,k_2}\ h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \quad = \quad \sum_{k_1=j_1^*-1}^{j_1^*} \tau_1^{k_1,k_2}\ \ell_1 \quad \geq \quad 0$$

otherwise j_1^* would not be a maximizer of H_1. Also, by assumption,

$$\mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \quad \leq \quad \mathcal{R}(y_1^{j_1^*-1} \mid y_2^{k_2})\ ,$$

so that

$$\sum_{k_1=j_1^*-1}^{j^*} \tau^1_{k_1,k_2}\ell_1\ \ell_2\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2})\ell_2 \sum_{k_1=j_1^*-1}^{j_1^*} \tau^1_{k_1,k_2}\ \ell_1\ .$$

Since, by assumption, $\ell_2\,\mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \leq 1$, (B.15) follows.

Step 3. A repeated application of this argument yields

$$\sum_{k_1=0}^{j^*} \tau_{k_1,k_2}^1 \ell_1\,\ell_2\,\mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad \sum_{k_1=0}^{j_1^*} \tau_{k_1,k_2}^1\, h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2})\,. \tag{B.16}$$

Step 4. We now claim that, for any $z \geq j_1^* + 1$,

$$\sum_{k_1=j_1^*+1}^{z} \tau_{k_1,k_2}^1 \ell_1\,\ell_2\,\mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0\,. \tag{B.17}$$

Note that

$$\sum_{k_1=j_1^*+1}^{z} \tau_{k_1,k_2}^1\, h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \quad = \quad \sum_{k_1=j_1^*+1}^{z} \tau_{k_1,k_2}^1\,\ell_1 \quad \leq \quad 0 \tag{B.18}$$

otherwise j_1^* would not be a maximizer of H_1. In particular, this requires that $\tau_1^{j_1^*+1,k_2} \leq 0$ so that

$$\sum_{k_1=j_1^*+1}^{j_1^*+1} \tau_1^{k_1,k_2} \ell_1\,\ell_2\,\mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0\,,$$

using that, by assumption, $\mathcal{R}(y_1^{j_1^*+1} \mid y_2^{k_2}) \geq 0$. Now suppose that also $\tau_1^{j_1^*+2,k_2} \leq 0$, then, using that $\mathcal{R}(y_1^{j_1^*+2} \mid y_2^{k_2}) \geq 0$, it follows that also

$$\sum_{k_1=j_1^*+1}^{j_1^*+2} \tau_1^{k_1,k_2} \ell_1\,\ell_2\,\mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0$$

If, by contrast, $\tau_{j_1^*+2,k_2}^1 \geq 0$, then, using $\mathcal{R}(y_1^{j_1^*+2} \mid y_2^{k_2}) \leq \mathcal{R}(y_1^{j_1^*+1} \mid y_2^{k_2})$

$$\sum_{k_1=j_1^*+1}^{j_1^*+2} \tau_1^{k_1,k_2} \ell_1\,\ell_2\,\mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \;\leq\; \mathcal{R}(y_1^{j_1^*+1} \mid y_2^{k_2}) \ell_2 \sum_{k_1=j_1^*+1}^{j_1^*+2} \tau_{k_1,k_2}^1\,\ell_1\,,$$

where the right-hand side expression is negative by (B.18). Again, it follows that

$$\sum_{k_1=j_1^*+1}^{j_1^*+2} \tau_{k_1,k_2}^1\,\ell_1\,\ell_2\,\mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0\,.$$

A repeated application of this argument establishes (B.17). In particular, this implies that

$$\sum_{k_1=j_1^*+1}^{K_1} \tau_{k_1,k_2}^1 \ell_1\,\ell_2\,\mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0\,. \tag{B.19}$$

Step 5. The inequality in (B.13) now follows from combining (B.16) and (B.19). □

Proof that h is not a Pareto-Improving Direction

Suppose that, in contrast to what we seek to show, $R_\tau(0,h) - \max_{y\in\mathcal{Y}}\ h(y) > 0$. Since $R_\tau(0,h) = \lim_{m\to\infty}\ R_\tau(0,h_m)$ and since for every $\epsilon > 0$, there is $\hat{m}(\epsilon)$ so that $m > \hat{m}(\epsilon)$ implies that, for all $y \in \mathcal{Y}$, $\mid h_m(y) - h(y) \mid \leq \epsilon$, this implies that, there exists $m > \hat{m}(\epsilon)$ so that $R_\tau(0,h_m) - \max_{y\in\mathcal{Y}}\ h_m(y) > 0$. This contradicts (B.10). Thus, the assumption that h is a Pareto-improving direction has led to a contradiction and must be false.

B.1.3 Proof of Theorem 3

We only show that (2.4) holds. The proof that also (2.5) holds can be obtained along the same lines. Let h be a function that is continuous on the bounded domain $\mathcal{Y}$ and which is hence differentiable almost everywhere. We approximate reform direction h by a collection of simple tax reforms as described in the proof of Theorem 2, see section B.1.2. As a first step, we seek to characterize the Gateaux differential of tax revenue in direction h_m, $R_\tau(0,h_m)$. Using that

$$\begin{aligned} h_m(y_1,y_2) &= h^{int}(y_1,y_2) + h^{b1}(y_1,y_2) + h^{b2}(y_1,y_2) \\ &= \textstyle\sum_{k\in K} \tau_1^k\ h_1^k(y_1,y_2)\ +\ \sum_{k\in K} \tau_2^k\ h_2^k(y_1,y_2) \\ &\quad + \textstyle\sum_{k_1\in K_1} \tau_1^{k_1} h^{bk_1}(y_1)\mathbf{1}(y_2=0) \\ &\quad + \textstyle\sum_{k_2\in K_2} \tau_2^{k_2} h^{bk_2}(y_2)\mathbf{1}(y_1=0) \end{aligned}$$

and the linearity of the Gateaux differential, we obtain

$$\begin{aligned} R_\tau(0,h_m) &= \textstyle\sum_{k\in K} \tau_1^k\ R_\tau(0,h_1^k) + \sum_{k\in K} \tau_2^k\ R_\tau(0,h_2^k) \\ &\quad + \textstyle\sum_{k_1\in K_1} \tau_1^{k_1} R_\tau(0,h^{bk_1}\mathbf{1}(y_2=0)) \\ &\quad + \textstyle\sum_{k_2\in K_2} \tau_2^{k_2} R_\tau(0,h^{bk_2}\mathbf{1}(y_1=0))\ . \end{aligned}$$

Taking Limits. We have partitioned $\mathcal{Y}_1 \times \mathcal{Y}_2$ into m^2 squares. We now take limits and let $m \to \infty$. Note that for $\ell_1 = \frac{\bar{y}_1}{m}$ and $\ell_2 = \frac{\bar{y}_2}{m}$ close to zero it follows from first order Taylor approximations that

$$R_\tau(0,h_1^k) \quad \simeq \quad \ell_1\ \ell_2\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2})\ ,$$

where $y_1^{k_1} = \ell_1\ k_1$ and $y_2^{k_2} = \ell_2\ k_2$. The approximation is perfect in the limit as $m \to \infty$. Analogously, we have that

$$R_\tau(0, h_2^k) \quad \simeq \quad \ell_1\ \ell_2\ \mathcal{R}(y_2^{k_2} \mid y_1^{k_1})\ ,$$

$$R_\tau(0, h^{bk_1}\mathbf{1}(y_2 = 0)) \quad \simeq \quad \ell_1\ \mathcal{R}(y_1^{k_1} \mid 0)\ ,$$

and

$$R_\tau(0, h^{bk_2}\mathbf{1}(y_1 = 0)) \quad \simeq \quad \ell_2\ \mathcal{R}(y_2^{k_2} \mid 0)\ .$$

Thus, for $\ell_1 = \frac{\bar{y}_1}{m}$ and $\ell_2 = \frac{\bar{y}_2}{m}$ close to zero and by the linearity of the Gateaux differential, we obtain

$$\begin{aligned} R_\tau(0, h_m) \quad \simeq \quad & \textstyle\sum_{k\in K} \tau_1^k\ \ell_1\ \ell_2\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) + \sum_{k\in K} \tau_2^k\ \ell_1\ \ell_2\ \mathcal{R}(y_2^{k_2} \mid y_1^{k_1}) \\ & + \textstyle\sum_{k_1\in K_1} \tau_1^{k_1}\ell_1\ \mathcal{R}(y_1^{k_1} \mid 0) \\ & + \textstyle\sum_{k_2\in K_2} \tau_2^{k_2}\ell_2\ \mathcal{R}(y_2^{k_2} \mid 0)\ . \end{aligned}$$

and

$$\begin{aligned} & \lim_{m\to\infty}\ R_\tau(0, h_m) = \\ & \lim_{m\to\infty}\ \textstyle\sum_{k\in K} \tau_1^k\ (y_1^{k_1+1} - y_1^{k_1})\ (y_2^{k_2+1} - y_2^{k_2})\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \\ & + \lim_{m\to\infty}\ \textstyle\sum_{k\in K} \tau_2^k\ \ (y_1^{k_1+1} - y_1^{k_1})\ (y_2^{k_2+1} - y_2^{k_2})\ \mathcal{R}(y_2^{k_2} \mid y_1^{k_1}) \\ & + \lim_{m\to\infty}\ \textstyle\sum_{k_1\in K_1} \tau_1^{k_1}\ (y_1^{k_1+1} - y_1^{k_1})\ \mathcal{R}(y_1^{k_1} \mid 0) \\ & + \lim_{m\to\infty}\ \textstyle\sum_{k_2\in K_2} \tau_2^{k_2}\ (y_2^{k_2+1} - y_2^{k_2})\ \mathcal{R}(y_2^{k_2} \mid 0)\ . \end{aligned}$$

All the expressions on the right-hand side are Riemann sums that converge to the corresponding Riemann integral as $m \to \infty$. Therefore,

$$\begin{aligned} R_\tau(0, h) \quad = \quad & \lim_{m\to\infty}\ R_\tau(0, h_m) \\ = \quad & \textstyle\int_{\mathcal{Y}_1\times\mathcal{Y}_1} h_1(y_1, y_2)\ \mathcal{R}(y_1 \mid y_2)\ dy_2\ dy_1 \\ & + \textstyle\int_{\mathcal{Y}_1\times\mathcal{Y}_1} h_2(y_1, y_2)\ \mathcal{R}(y_2 \mid y_1)\ dy_2\ dy_1 \\ & + \textstyle\int_{\mathcal{Y}_1} h_1(y_1, 0)\ \mathcal{R}(y_1 \mid 0)\ dy_1 \\ & + \textstyle\int_{\mathcal{Y}_2} h_2(0, y_2)\ \mathcal{R}(y_2 \mid 0)\ dy_2\ . \end{aligned}$$

B.1.4 Proof of Corollary 4

Starting from the characterization of $W_\tau(0,h)$ in Theorem 3, we can substitute

$$h_1(y_1,0) + \int_0^{y_2} h_{12}(y_1,s_2)\, ds_2$$

for $h_1(y_1,y_2)$ and

$$h_2(0,y_2) + \int_0^{y_1} h_{21}(s_1,y_2)\, ds_1$$

for $h_2(y_1,y_2)$. This yields

$$\begin{aligned} W_\tau(0,h) \;=\; & \int_{\mathcal{Y}} \Big(\int_0^{y_2} h_{12}(y_1,s_2)\, ds_2 \Big) \mathcal{W}_1(y_1 \mid y_2) dy \\ & + \int_{\mathcal{Y}} \Big(\int_0^{y_1} h_{21}(s_1,y_2)\, ds_1 \Big) \mathcal{W}_2(y_2 \mid y_1) dy \\ & + \int_{\mathcal{Y}_1} h_1(y_1,0) \left(\mathcal{W}_1^s(y_1) + \int_{\mathcal{Y}_2} \mathcal{W}_1(y_1 \mid y_2) dy_2 \right) dy_1 \\ & + \int_{\mathcal{Y}_2} h_2(0,y_2) \left(\mathcal{W}_2^s(y_2) + \int_{\mathcal{Y}_1} \mathcal{W}_2(y_2 \mid y_1) dy_1 \right) dy_2 \,, \end{aligned}$$

One can use integration by parts to verify that

$$\int_{\mathcal{Y}_2} \Big(\int_0^{y_2} h_{12}(y_1,s_2)\, ds_2 \Big) \mathcal{W}_1(y_1 \mid y_2) dy_2 \;=\; \int_{\mathcal{Y}_2} h_{12}(y_1,y_2) \Big(\int_{y_2}^{\bar{y}_2} \mathcal{W}_1(y_1 \mid s_2) ds_2 \Big) dy_2$$

and

$$\int_{\mathcal{Y}_1} \Big(\int_0^{y_1} h_{21}(s_1,y_2)\, ds_1 \Big) \mathcal{W}_2(y_2 \mid y_1) dy_1 \;=\; \int_{\mathcal{Y}_1} h_{21}(y_1,y_2) \Big(\int_{y_1}^{\bar{y}_1} \mathcal{W}_2(y_2 \mid s_1) ds_1 \Big) dy_1$$

Thus,

$$\begin{aligned} W_\tau(0,h) \;=\; & \int_{\mathcal{Y}} h_{12}(y_1,y_2) \Big(\int_{y_2}^{\bar{y}_2} \mathcal{W}_1(y_1 \mid s_2) ds_2 \Big) dy \\ & + \int_{\mathcal{Y}} h_{21}(y_1,y_2) \Big(\int_{y_1}^{\bar{y}_1} \mathcal{W}_2(y_2 \mid s_1) ds_1 \Big) dy \\ & + \int_{\mathcal{Y}_1} h_1(y_1,0) \left(\mathcal{W}_1^s(y_1) + \int_{\mathcal{Y}_2} \mathcal{W}_1(y_1 \mid y_2) dy_2 \right) dy_1 \\ & + \int_{\mathcal{Y}_2} h_2(0,y_2) \left(\mathcal{W}_2^s(y_2) + \int_{\mathcal{Y}_1} \mathcal{W}_2(y_2 \mid y_1) dy_1 \right) dy_2 \,, \end{aligned}$$

Using that $h_{12}(y) = h_{21}(y)$, for all $y \in \mathcal{Y}$ yields the expression in Corollary 4.

B.1.5 Behavioral Responses and Tax Revenue

A Couple's Choice Problem. Given a tax reform (τ,h), we denote by

$$y^{*d}(\omega,\tau,h) \quad \in \quad \text{argmax}_y \; C^1(y) - k_1(y_1,\omega_1) \;-\; k_2(y_2,\omega_2) \,, \tag{B.20}$$

the couples' optimal earnings choices conditional on being a dual-earner couple. We denote by $v^d(\omega, \tau, h)$ the indirect utility associated with a solution to the problem in (B.20). Analogously, we write

$$y_i^{*s}(\omega_i, \tau, h) \quad \in \quad \text{argmax}_y \, C^1(y_i, 0) - k_i(y_i, \omega_i) \tag{B.21}$$

for the optimal choice conditional on the spouse with index $i \in \{1, 2\}$ being the single earner and $v_i^s(\omega_i, \tau, h)$ for the indirect utility associated with a solution to the problem in (B.21). A couple will choose to be a dual earner couple when the resulting payoff of $v^d(\omega, \tau, h) - \varphi_1 - \varphi_2$. exceeds the payoff from its alternative options. Alternatives are that only spouse 1 or only spouse 2 exerts productive effort with payoffs that are, respectively, given by $v_1^s(\omega_1, \tau, h) - \varphi_1$ and $v_2^s(\omega_2, \tau, h) - \varphi_2$. Thus, for dual earner couples,

$$v^d(\omega, \tau, h) - v_1^s(\omega_1, \tau, h) \; := \; \hat{\varphi}_2(\omega, \tau, h) \quad \geq \quad \varphi_2 \, , \quad \text{and}$$

$$v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h) \; := \; \hat{\varphi}_1(\omega, \tau, h) \quad \geq \quad \varphi_1 \, .$$

It also has to be true that $v^d(\omega, \tau, h) \geq \varphi_1 + \varphi_2$. Otherwise, the couple would be better off with no earnings at all. A couple will choose spouse i as its single-earner when this yields a payoff that dominates what is achievable as a dual earner couple or with spouse $-i$ as the single earner. The latter implies

$$v_i^s(\omega_i, \tau, h) - \varphi_i \quad \geq \quad v_{-i}^s(\omega_{-i}, \tau, h) - \varphi_{-i} \, .$$

The payoff with i as the single earner must also dominate having no earnings at all,

$$v_i^s(\omega_i, \tau, h) \; := \; \hat{\hat{\varphi}}_i(\omega_i, \tau, h) \quad \geq \quad \varphi_i \, .$$

Finally, for couples with no earnings, it must be true that, for both i, $\varphi_i \geq \hat{\hat{\varphi}}(\omega_i, \tau, h)$ and $v^d(\omega, \tau, h) \leq \varphi_1 + \varphi_2$. The following Proposition states implications of these conditions under a set of additional assumptions, see Figure B.2 for an illustration.

Proposition 5. *Suppose that the following premises are satisfied:*

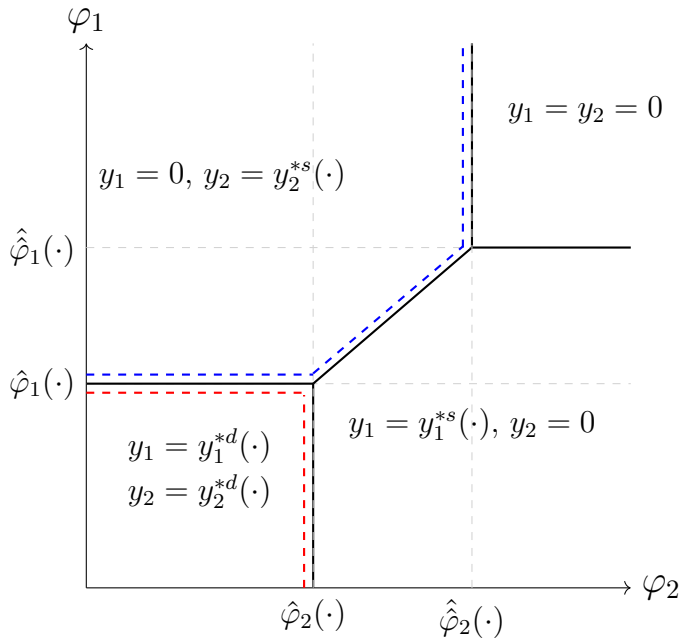
i) For all i, $y_i^{*s}(\omega_i, \tau, h) > y_i^{*d}(\omega, \tau, h)$.

ii) For all i, $\lim_{\omega_i \to 0} v_i^s(\omega_i, \tau, h) = 0$ and $\lim_{\omega_{-i} \to 0} v^d(\omega, \tau, h) = v_i^s(\omega_i, \tau, h)$.

Then, for any $\omega >> 0$, and for any $i = 1, 2$,

$$\hat{\hat{\varphi}}(\omega_i, \tau, h) \quad > \quad \hat{\varphi}_i(\omega, \tau, h) \, . \tag{B.22}$$

Figure B.2. – A Couple's Discrete Choice Problem (I)



Note: The figure shows how the earnings choices of the spouses in a couple with productivities $\omega = (\omega_1, \omega_2)$ depend on their fixed costs $\varphi = (\varphi_1, \varphi_2)$. For fixed costs in the lower left rectangle both spouses generate income. For fixed costs in the upper right area, the couple has no earnings at all. For fixed costs in the upper left area, spouse 1 is the single earner. For fixed costs in the lower right area, spouse 2 is the single earner.

Hence, $y^*(\omega, \varphi, \tau, h) =$

$$\begin{cases} y^{d*}(\cdot), & \text{if } \varphi_1 < \hat{\varphi}_1(\cdot) \text{ and } \varphi_2 < \hat{\varphi}_2(\cdot)\ , \\ (y_1^{*s}(\cdot), 0), & \text{if } \varphi_1 < \min\{\Delta v^s(\cdot) + \varphi_2, \hat{\hat{\varphi}}_1(\cdot)\} \text{ and } \varphi_2 > \hat{\varphi}_2(\cdot)\ , \\ (0, y_2^{*s}(\cdot)), & \text{if } \varphi_1 > \hat{\varphi}_1(\cdot) \text{ and } \varphi_2 < \min\{-\Delta v^s(\cdot) + \varphi_1, \hat{\hat{\varphi}}_2(\cdot)\}\ , \\ (0, 0), & \text{if } \varphi_1 > \hat{\hat{\varphi}}_1(\cdot) \text{ and } \varphi_2 > \hat{\hat{\varphi}}_2(\cdot)\ . \end{cases} \tag{B.23}$$

where $\Delta v^s(\omega, \tau, h) := v_1^s(\omega_1, \tau, h) - v_2^s(\omega_2, \tau, h)$.

A formal proof of Proposition 5 can be found at the end of this Section. The first premise of Proposition 5 is fulfilled under a progressive tax system that gives rise to positive jointness. Under such a system, the marginal tax rate on the earnings of one spouse is increasing in the earnings of the other spouse. Thus, when spouse i is a single earner her marginal tax rate is lower as compared to any other constellation in which she is part of a dual earner couple. Therefore, y_i^* is decreasing in y_{-i}. The second premise appears natural on the assumption that y_i^* goes to zero as ω_i goes to zero and that, in this case, also i's

contribution to the couple's payoff goes to zero.[2] Under these assumptions, the cutoff types $\hat{\hat{\varphi}}(\omega_i, \tau, h)$ and $\hat{\varphi}_i(\omega, \tau, h)$ are ordered as follows: Spouse i has positive earnings – whatever the earnings choice for spouse $-i$ – when $\varphi < \hat{\varphi}$. Likewise, spouse i has no earnings when $\varphi > \hat{\hat{\varphi}}$, again, irrespective of whether $y_{-i} = 0$ or $y_{-i} > 0$. For intermediate fixed costs, however, whether or not spouse i has positive earnings depends on whether the other spouse has $y_{-i} = 0$ or $y_{-i} > 0$.

In the sequel, we will characterize behavioral responses to taxation and, ultimately, derive sufficient statistics formulas for the conditional revenue functions. In doing so we will proceed under the assumption that, in the status quo, i.e. for $\tau = 0$, $\hat{\hat{\varphi}}(\omega_i, \tau, h) > \hat{\varphi}_i(\omega, \tau, h)$ holds for all i and all ω. Imposing this assumption eases the exposition. Otherwise we would have to get into a distinction of various subcases.

Behavioral Responses to a Simple Tax Reform: Dual Earner Couples. Consider a simple tax reform (τ, h) so that, conditional on $y_1 \in B_1(y_{1s}, \ell_1)$, the couples' tax burden increases by $\tau h_2(y_2)$, where

$$h_2(y_2) = \begin{cases} y_2 - y_{2s}, & \text{if } y_2 \in B_2(y_{2s}, \ell_2) \ , \\ \ell_2, & \text{if } y_2 \geq y_{2s} + \ell_2 \ , \\ 0, & \text{otherwise} \ . \end{cases}$$

Let $\mathcal{Y}_{h_2} = \{y \mid h_2(y) \neq 0\}$ be the "treated" range of incomes and denote by

$$\Omega^d(\tau, h_2) = \{\omega \mid y^{*d}(\omega, \tau, h_2) \in \mathcal{Y}_{h_2}\}$$

the corresponding set of productive abilities.[3] Obviously, there will be behavioral responses to infinitesimal changes of τ only for couples with $\omega \in \Omega(\tau, h_2)$. For all other couples, $\mathcal{Y}(h_2)$ is a dominated region and what happens there has no consequence for their choices.

Now consider a couple with $\omega \in \Omega(\tau, h_2)$. The couples' behavioral response at the intensive margin is given by

$$y^{*d}_{\tau}(\omega, \tau, h_2) = (y^{*d}_{1,\tau}(\omega, \tau, h), y^{*d}_{2,\tau}(\omega, \tau, h)) \ ,$$

[2]It follows form the Spence-Mirrlees single-crossing condition that y^*_i is a decreasing function of ω_i. To avoid lengthy formal arguments, we refrain from stating further assumptions on the effort cost function and the tax function that imply premise ii).

[3]We allow for the possibility that optimal choices are not unique. The statement $y^{*d}(\omega, \tau, h_2) \in \mathcal{Y}_{h_2}$ is a shorthand for: There is an element y of $\operatorname{argmax}_y C^1(y) - k_1(y_1, \omega_1) - k_2(y_2, \omega_2)$ that belongs to $\mathcal{Y}_{h_2}$.

i.e. by the derivative of its optimal earnings choice conditional on being a dual-earner couple.[4] Since indirect utility $v^d(\omega, \tau, h)$ goes down, some couples give up on having two earners. This is the behavioral response at the extensive margin. It is illustrated by the dashed red lines in Figure B.2. They indicate that both $\hat{\varphi}_1(\omega, \tau, h)$ and $\hat{\varphi}_2(\omega, \tau, h)$ shrink in response to the reform. Couples with fixed costs between the black and red dashed lines move out of the set of dual-earner couples and into the set of single-earner couples.

Single Earner Couples. Consider a simple tax reform so that, conditional on $y_2 = 0$, the couples' tax burden increases by $\tau h_1(y_1)$, where

$$h_1(y_1) \;=\; \begin{cases} y_1 - y_{1s}, & \text{if} \quad y_1 \in B_1(y_{1s}, \ell_1) \;, \\ \ell_1, & \text{if} \quad y_1 \geq y_{1s} + \ell_1 \;, \\ 0, & \text{otherwise} \;. \end{cases}$$

Again, let $\mathcal{Y}_{h_1} = \{y \mid h_1(y) \neq 0\}$ be the "treated" range of incomes and denote by $\Omega_i^s(\tau, h_i) = \{\omega \mid y_i^{*s}(\omega, \tau, h) \in \mathcal{Y}_{h_1}\}$ the corresponding set of productive abilities. The behavioral response at the intensive margin is given by $y_{1,\tau}^{*s}(\omega_i, \tau, h_1)$. i.e. by the derivative of the optimal earnings choice conditional on spouse 1 being the couple's single earner. When indirect utility $v_1^s(\omega_i, \tau, h)$ goes down some couples give up on having spouse 1 as the single earner. Depending on their fixed costs, some move into being a dual earner couple, others assign the role of the single earner now to spouse $-i$ and some give up on having earnings at all. These behavioral responses at the extensive margin are illustrated by the dashed blue lines in Figure B.2.

Implications for Tax Revenue. We now clarify the repercussion that behavioral responses have for tax revenue. As will become clear, even though the earnings function that has been characterized in Proposition 5 exhibits discontinuous jumps at the cutoff levels for fixed costs, tax revenue is a differentiable function of the reform intensity τ if the distribution of fixed costs is "well-behaved." Consider a tax reform h and denote by $r(\omega, \tau, h)$ the change in tax revenue that is due to the set of couples who all have productive abilities ω, but may differ in their fixed costs of generating income. Thus,

$$\begin{aligned} r(\omega, \tau, h) + r^0(\omega) \;=\; & \sigma^d(\omega, \tau, h)\; T^1(y^{d*}(\omega, \tau, h), \tau, h) \\ & + \sigma_1^s(\omega, \tau, h)\; T^1(y_1^{s*}(\omega_1, \tau, h), 0, \tau, h) \\ & + \sigma_2^s(\omega, \tau, h)\; T^1(0, y_2^{s*}(\omega_2, \tau, h), \tau, h) \end{aligned}$$

[4] Left-sided and right-sided derivatives of y^{*d} may not be identical when there are multiple optimal choices conditional on being a dual earner couple. Clarifying whether the left-sided or the right-sided derivative captures the behavioral responses to a tax reform then requires a more detailed analysis.

where, for any level of y, $T^1(y,\tau,h) := T^0(y)+\tau h(y)$, $r^0(\omega)$ denotes revenues in the status quo and

$$\sigma^d(\omega,\tau,h) \quad = \quad \int_0^{\hat{\varphi}_1(\omega,\tau,h)} \int_0^{\hat{\varphi}_2(\omega,\tau,h)} f^{\varphi}(\varphi_1,\varphi_2 \mid \omega)\, d\varphi_2\, d\varphi_1 \;,$$

is the fraction of dual earner couples among those with productive ability ω, and the density f^{φ} characterizes the joint distribution of φ_1 and φ_2, conditional on the couples' productivities being ω. The fractions of couples where only spouse 1 or only spouse 2 has positive earnings are analogously defined. The derivative of $r(\omega,\tau,h)$ with respect to τ shows how these behavioral responses affect tax revenue. Specifically,

$$r_\tau(\omega,\tau,h) = r^I_\tau(\omega,\tau,h) + r^E_\tau(\omega,\tau,h) \;, \tag{B.24}$$

where $r^I_\tau(\omega,\tau,h)$ captures the revenue implications of intensive margin responses and $r^E_\tau(\omega,\tau,h)$ captures the revenue implications of extensive margin responses.[5] An aggregation over all possible combinations of ω gives us the change in tax revenue due to the reform (τ,h):

$$R(\tau,h) = \mathbb{E}_\omega[r(\omega,\tau,h)\; \mathbf{1}(\omega \in \Omega^d(\tau,h))]$$

Upon taking a derivative with respect to τ and evaluating the resulting expression at the status quo, we obtain the marginal effect on tax revenue that is associated with reform direction h. Proposition 6 provides a characterization.

Proposition 6.

$$R_\tau(0,h) = \mathbb{E}_\omega[r_\tau(\omega,0,h)\; \mathbf{1}(\omega \in \Omega^d(0,h))]$$

The Proposition – see the end of this section for a formal proof – says that a reform direction's marginal effect on tax revenue is only due those couples who, in the status

[5]More specifically,

$$\begin{aligned} r^I_\tau(\omega,\tau,h) \quad = \quad & \sigma^d(\omega,\tau,h)\Big(T^1_{y_1}(y^{d*}(\cdot),\tau,h)y^{d*}_{1,\tau}(\cdot) + T^1_{y_2}(y^{d*}(\cdot),\tau,h)y^{d*}_{2,\tau}(\cdot) + h(y^{d*}(\cdot))\Big) \\ & +\sigma^s_1(\omega,\tau,h)\Big(T^1_{y_1}(y^{s*}_1(\cdot),0,\tau,h)y^{s*}_{1,\tau}(\cdot) + h(y^{s*}_1(\cdot))\Big) \\ & +\sigma^s_2(\omega,\tau,h)\Big(T^1_{y_2}(0,y^{s*}_2(\cdot),\tau,h)y^{s*}_{2,\tau}(\cdot) + h(y^{s*}_2(\cdot))\Big) \qquad \text{and} \end{aligned} \tag{B.25}$$

$$\begin{aligned} r^E_\tau(\omega,\tau,h) \quad = \quad & \sigma^d_\tau(\omega,\tau,h)\; T^1(y^{d*}(\omega,\tau,h),\tau,h) \\ & +\sigma^s_{1,\tau}(\omega,\tau,h)\; T^1(y^{s*}_1(\omega_1,\tau,h),0,\tau,h) \\ & +\sigma^s_{2,\tau}(\omega,\tau,h)\; T^1(0,y^{s*}_2(\omega_2,\tau,h),\tau,h)\;. \end{aligned} \tag{B.26}$$

quo, make choices in the "treated region" i.e. in the range of incomes with $h(y) \neq 0$. This observation greatly simplifies the characterization of conditional revenue functions in the subsequent section.[6]

Proof of Proposition 5. We show that under the premises of the Proposition, we have, for all $\omega >> 0$.

$$\hat{\varphi}_1(\omega, \tau, h) \; < \; \hat{\hat{\varphi}}_1(\omega_1, \tau, h) \quad \text{and} \quad \hat{\varphi}_2(\omega, \tau, h) \; < \; \hat{\hat{\varphi}}_2(\omega_2, \tau, h) \; .$$

We only show that $\hat{\varphi}_1(\cdot) \; < \; \hat{\hat{\varphi}}_1$, or, equivalently, that

$$v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h) \quad < \quad v_1^s(\omega_1, \tau, h) \; .$$

By assumption,

$$\lim_{\omega_2 \to 0} v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h) \quad = \quad v_1^s(\omega_1, \tau, h) \; .$$

It therefore suffices to show that $v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h)$ is a decreasing function of ω_2. An application of the Envelope Theorem reveals that

$$\frac{\partial}{\partial \omega_2} v^d(\omega, \tau, h) \; = \; -k_{2,2}(y_2^{*d}(\cdot), \omega_2)$$

and

$$\frac{\partial}{\partial \omega_2} v_2^s(\omega_2, \tau, h) \; = \; -k_{2,2}(y_2^{*s}(\cdot), \omega_2)$$

where, by assumption, $y_2^{*s}(\cdot) > y_2^{*d}(\cdot)$. Hence,

$$\begin{aligned} \tfrac{\partial}{\partial \omega_2}\Big(v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h)\Big) \; &= \; k_{2,2}(y_2^{*s}(\cdot), \omega_2) - k_{2,2}(y_2^{*d}(\cdot), \omega_2) \\ &= \; \textstyle\int_{y_2^{*d}(\cdot)}^{y_2^{*s}(\cdot)} k_{2,12}(y, \omega_2) \; dy \; , \end{aligned}$$

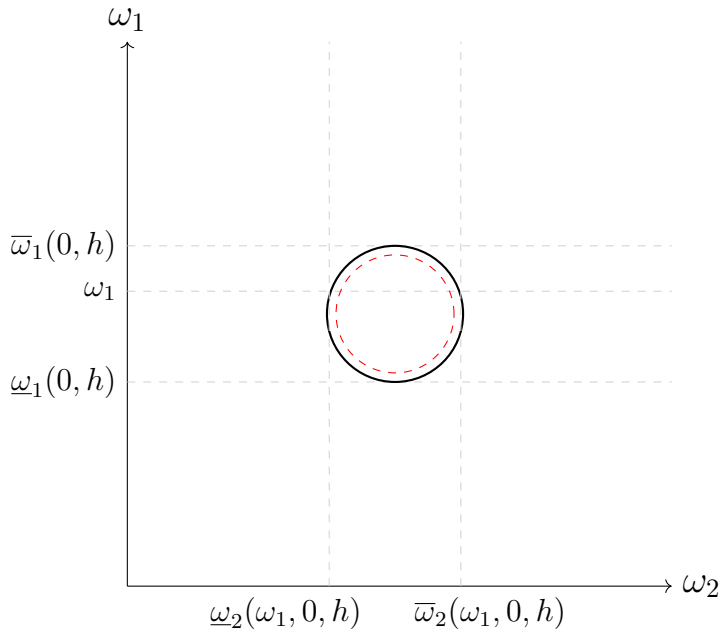
where $k_{2,12} < 0$ by the Spence-Mirrless single crossing property.

Proof of Proposition 6. Consider a reform direction h with $\mathcal{Y}_h = \{y \mid h(y) \neq 0\}$ in the interior of $\mathcal{Y}$ and with

$$\Omega^d(\tau, h) \quad = \quad \{\omega \mid y^{d*}(\omega, \tau, h) \in \mathcal{Y}_h\}$$

[6]To be clear, the set $\Omega^d(\tau, h)$ will generally depend on τ. Possibly a reform creates dominated regions in space of incomes and those regions may become larger as the reform intensifies, so that $\tau' > \tau$ implies that $\Omega^d(\tau', h)$ is a strict subset of $\Omega^d(\tau, h)$. As formally demonstrated in the proof of Proposition 6, this force is, however, inconsequential, at $\tau = 0$

Figure B.3. – A Couple's Discrete Choice Problem (II)



Note: The area in the black circle is the set $\Omega^d(0,h)$. (The circle is a generic shape. There is no presumption that the set takes this specific geometric form.) If for all $y \in \mathcal{Y}_h$, $h(y) > 0$, the treated area becomes less attractive and the set $\Omega^d(\tau,h)$ is therefore decreasing in τ, as indicated by the dashed red circle, which represents $\Omega^d(\tau,h)$ for some positive level of τ. At $\tau = 0$, the revenue loss from the fact that fewer couples make choices in the treated area is equal to zero. The reason is that every loss of a couple with a type ω on the black circle is multiplied with the change in their tax payment. At $\tau = 0$, this change is zero, $r(\omega, 0, h) = 0$.

Let

$$\underline{\omega}_1(\tau,h) \quad = \quad \min\{\omega_1 \mid \exists \omega_2 \text{ s.t } (\omega_1,\omega_2) \in \Omega^d(\tau,h)\}$$

and

$$\overline{\omega}_1(\tau,h) \quad = \quad \max\{\omega_1 \mid \exists \omega_2 \text{ s.t } (\omega_1,\omega_2) \in \Omega^d(\tau,h)\} \ .$$

For given $\omega_1 \in \Omega_1^d(\tau,h) := [\underline{\omega}_1(\tau,h), \overline{\omega}_1(\tau,h)]$, let

$$\underline{\omega}_2(\omega_1,\tau,h) \quad = \quad \min\{\omega_2 \mid (\omega_1,\omega_2) \in \Omega^d(\tau,h)\}$$

and

$$\overline{\omega}_2(\omega_1,\tau,h) \quad = \quad \max\{\omega_2 \mid (\omega_1,\omega_2) \in \Omega^d(\tau,h)\} \ .$$

Armed with this notation we can write

$$R(\tau,h) \quad = \quad \mathbb{E}_\omega[r(\omega,\tau,h)\mathbf{1}(\omega \in \Omega^d(\tau,h))]$$

$$\int_{\underline{\omega}_1(\tau,h)}^{\overline{\omega}_1(\tau,h)} \int_{\underline{\omega}_2(\omega_1,\tau,h)}^{\overline{\omega}_2(\omega_1,\tau,h)} r(\omega,\tau,h) f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \ f_{\omega_1}(\omega_1) d\omega_1 \ ,$$

where f_{ω_1} is the marginal distribution of ω_1 associated with F_θ and $f_{\omega_2}(\cdot \mid \omega_1)$ is the distribution of ω_2 conditional on ω_1. Differentiating with respect to τ and repeatedly applying of Leibniz' rule yields

$$\begin{aligned} R_\tau(\tau,h) &= \int_{\underline{\omega}_1(\tau,h)}^{\overline{\omega}_1(\tau,h)} \int_{\underline{\omega}_2(\omega_1,\tau,h)}^{\overline{\omega}_2(\omega_1,\tau,h)} r_\tau(\omega,\tau,h)\; f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \; f_{\omega_1}(\omega_1) d\omega_1 \;, \\ &\quad + A(\tau,h) \;+\; B(\tau,h) \;+\; C(\tau,h) \;+\; D(\tau,h) \;, \end{aligned}$$

where

$$A(\tau,h) = \left(\int_{\underline{\omega}_2(\overline{\omega}_1(\cdot),\tau,h)}^{\overline{\omega}_2(\overline{\omega}_1(\cdot),\tau,h)} r(\overline{\omega}_1(\cdot),\omega_2,\tau,h) f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \right) \frac{\partial}{\partial \tau} \overline{\omega}_1(\tau,h) \;,$$

$$B(\tau,h) = -\left(\int_{\underline{\omega}_2(\underline{\omega}_1(\cdot),\tau,h)}^{\overline{\omega}_2(\underline{\omega}_1(\cdot),\tau,h)} r(\underline{\omega}_1(\cdot),\omega_2,\tau,h) f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \right) \frac{\partial}{\partial \tau} \underline{\omega}_1(\tau,h) \;,$$

$$C(\tau,h) = \int_{\underline{\omega}_1(\tau,h)}^{\overline{\omega}_1(\tau,h)} r(\omega_1,\overline{\omega}_2(\cdot),\tau,h)\; \frac{\partial}{\partial \tau} \underline{\omega}_2(\omega_1,\tau,h)\; f_{\omega_1}(\omega_1) d\omega_1 \;,$$

and

$$D(\tau,h) = -\int_{\underline{\omega}_1(\tau,h)}^{\overline{\omega}_1(\tau,h)} r(\omega_1,\underline{\omega}_2(\cdot),\tau,h)\; \frac{\partial}{\partial \tau} \underline{\omega}_2(\omega_1,\tau,h)\; f_{\omega_1}(\omega_1) d\omega_1 \;.$$

Note that $\tau = 0$ implies that $T^1(y) = T^0(y)$, for all y. Hence $r(\omega,0,h) = 0$, for all ω and all h. This implies, moreover, that

$$A(0,h) \;=\; B(0,h) \;=\; C(0,h) \;=\; D(0,h) \;=\; 0 \;.$$

Therefore,

$$\begin{aligned} R_\tau(0,h) &= \int_{\underline{\omega}_1(0,h)}^{\overline{\omega}_1(0,h)} \int_{\underline{\omega}_2(\omega_1,0,h)}^{\overline{\omega}_2(\omega_1,0,h)} r_\tau(\omega,0,h)\; f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \; f_{\omega_1}(\omega_1) d\omega_1 \;, \\ &= \mathbb{E}_\omega[r_\tau(\omega,0,h)\; \mathbf{1}(\omega \in \Omega^d(0,h))] \;. \end{aligned}$$

B.1.6 Proof of Proposition 1

Notation. The joint distribution of productive abilities types is represented by a density $(\omega_1, \omega_2) \mapsto f_\omega(\omega_1, \omega_2)$, with marginal densities $\omega_1 \mapsto f_{\omega_1}(\omega_1)$ and $\omega_2 \mapsto f_{\omega_2}(\omega_2)$. We write $\omega_2 \mapsto f_{\omega_2}(\omega_2 \mid \omega_1)$ for the density of ω_2, conditional on ω_1, where

$$f_{\omega_2}(\omega_2 \mid \omega_1) \quad := \quad \frac{f_\omega(\omega_1, \omega_2)}{f_{\omega_1}(\omega_1)} \, .$$

The density $(\omega_1, \omega_2) \mapsto f_\omega(\omega_1, \omega_2)$ is an unobserved primitive of the economy. What is observed, by contrast, is the status quo distribution of earnings. We denote by $(y_1, y_2) \mapsto f_y(y_1, y_2)$ the density that represents the joint distribution of earnings, i.e. the cross-section distribution of $y_1^0(\theta)$ and $y_2^0(\theta)$. The marginal densities are denoted by f_{y_1} and f_{y_2} and the marginal *cdf*s by F_{y_1} and F_{y_2}. We write $y_2 \mapsto f_{y_2}(y_2 \mid y_1)$ and $y_2 \mapsto F_{y_2}(y_2 \mid y_1)$ for the conditional densities and *cdf*s.

Let Θ' be a measurable subset of $\Theta = \Theta_1 \times \Theta_2$. We use $P(\Theta')$ as a shorthand for

$$\int_\Theta \mathbf{1}(\theta \in \Theta') f(\theta) \, d\theta \, ,$$

where $\mathbf{1}$ is the indicator function, and the density f is associated with the joint distribution of θ_1 and θ_2.

Proof of Proposition 1

Recall that by Proposition 6

$$R_\tau(0, h) = \mathbb{E}_\omega[r_\tau(\omega, 0, h) \, \mathbf{1}(\omega \in \Omega^d(0, h))] \, .$$

Characterizing $\Omega^d(0, h)$. Fix $B_1 = [y_{1s}, y_{1s} + \ell_1]$ and $B_2 = [y_{2s}, y_{2s} + \ell_2]$ so that $\exists \omega : y^{d0}(\omega) \in B := B_1 \times B_2$, where $y^{d0}(\omega) = (y_1^{d0}(\omega), y_2^{d0}(\omega))$ is a shorthand for $y^{d*}(\omega, 0, h)$. Let

$$\underline{\omega}_1(y_{1s}) \quad := \quad \min \{\omega_1 \in \Omega_1 \mid \exists \, \omega_2 \text{ s.t. } y_1^{d0}(\omega_1, \omega_2) \geq y_{1s}\}$$

be the lowest value of ω_1 consistent with earnings of individual 1 in the bracket ranging from y_{1s} to $y_{1s} + \ell_1$. Analogously, define the largest value of ω_1 by

$$\overline{\omega}_1(y_{1s} + \ell_1) \quad := \quad \max \{\omega_1 \in \Omega_1 \mid \exists \, \omega_2 \text{ s.t. } y_1^{d0}(\omega_1, \omega_2) \leq y_{1s} + \ell_1\} \, .$$

We will sometimes write $\Omega_1(B) = [\underline{\omega}_1(y_{1s}), \overline{\omega}_1(y_{1s} + \ell_1)]$.

Henceforth, consider some fixed $\omega_1 \in [\underline{\omega}_1(y_{1s}), \overline{\omega}_1(y_{1s}+\ell_1)]$. There is then a non-empty subset $\Omega_2[B,\omega_1]$ of Ω_2 so that $\omega_2 \in \Omega_2[B,\omega_1]$ implies $y_1^{d0}(\omega_1,\omega_2) \in [y_{1s}, y_{1s}+\ell_1]$.

There is then a, possibly empty, subset $\Omega_2^{in}[B,\omega_1]$ of $\Omega_2[B,\omega_1]$ so that $\omega_2 \in \Omega_2^{in}[B,\omega_1]$ implies $y_2^0(\omega_1,\omega_2) \in [y_{2s}, y_{2s}+\ell_2]$. If the set is non-empty, denote by $\underline{\omega}_2^{in}(y_{2s} \mid \omega_1)$ and $\overline{\omega}_2^{in}(y_{2s}+\ell_2 \mid \omega_1)$ the minimal and the maximal element of this set.

There also is a, possibly empty, subset $\Omega_2^{+}[B,\omega_1]$ of $\Omega_2[B,\omega_1]$ so that $\omega_2 \in \Omega_2^{+}[B,\omega_1]$ implies $y_2^0(\omega_1,\omega_2) \geq y_{2s}+\ell_2$. If the set is non-empty, denote by $\underline{\omega}_2^{+}(y_{2s}+\ell_2 \mid \omega_1)$ and $\overline{\omega}_2^{+}(B \mid \omega_1)$ the minimal and the maximal element of this set.

Characterizing $r_\tau(\omega,0,h)$. From equations (B.24) - (B.26) in Section B.1.5 it follows that

$$\begin{aligned} r_\tau(\omega,0,h) \;=\; & \sigma^{d0}(\omega)\Big(T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot) + h(y^{d0}(\omega))\Big) \\ & +\sigma_1^{s0}(\omega)\Big(T^0_{y_1}(y_1^{s0}(\omega_1),0)y^{s0}_{1,\tau}(\cdot) + h(y_1^{s0}(\omega_1),0)\Big) \\ & +\sigma_2^{s0}(\omega)\Big(T^0_{y_2}(0,y_2^{s0}(\omega_2))y^{s0}_{2,\tau}(\cdot) + h(0,y_2^{s0}(\omega_2))\Big) \\ & +\sigma^{d0}_{\tau}(\omega,h)\; T^0(y^{d0}(\omega)) \\ & +\sigma^{s0}_{1,\tau}(\omega,h)\; T^0(y_1^{s0}(\omega_1),0) \\ & +\sigma^{s0}_{2,\tau}(\omega,h)\; T^0(0,y_2^{s0}(\omega_2))\;. \end{aligned} \tag{B.27}$$

where $\sigma^{d0}(\omega)$, $\sigma_1^{s0}(\omega)$ and $\sigma_2^{s0}(\omega)$ are, respectively, the fractions of dual and single earner couples amongst those with productive abilities ω in the status quo, and the expressions involving a subscript of τ and a superscript of 0 refer to derivatives with respect to τ that are evaluated at $\tau = 0$. We introduce shorthands – capturing, respectively, behavioral responses at the intensive and the extensive margin, β_I^{d0} and β_E^{d0}, and mechanical effects on tax revenue due to changes in the tax functions for dual or single-earner couples, μ^{d0}, μ_1^{0s}, μ_2^{0s} – so that we can write more compactly

$$\begin{aligned} r_\tau(\omega,0,h) \;&=\; \beta_I^{d0}(\omega) \quad + \quad \mu^0(\omega) \\ &=\; \beta_I^{d0}(\omega) + \beta_E^{d0}(\omega) \quad + \quad \mu^{d0}(\omega) + \mu_1^{0s}(\omega) + \mu_2^{0s}(\omega) \end{aligned}$$

where

$$\begin{aligned} \beta_I^{d0}(\omega) \;:=\; & \sigma^{d0}(\omega)\Big(T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot)\Big) \\ & +\sigma_1^{s0}(\omega)\; T^0_{y_1}(y_1^{s0}(\omega_1),0)y^{s0}_{1,\tau}(\cdot) \\ & +\sigma_2^{s0}(\omega)\; T^0_{y_2}(0,y_2^{s0}(\omega_2))y^{s0}_{2,\tau}(\cdot)\;, \end{aligned}$$

$$\begin{aligned}\beta_E^{d0}(\omega) &:= +\sigma_\tau^{d0}(\omega,h)\; T^0(y^{d0}(\omega)) \\ &\quad +\sigma_{1,\tau}^{s0}(\omega,h)\; T^0(y_1^{s0}(\omega_1),0) \\ &\quad +\sigma_{2,\tau}^{s0}(\omega,h)\; T^0(0,y_2^{s0}(\omega_2))\;,\end{aligned}$$

$$\mu^{d0}(\omega) := \sigma^{d0}(\omega)\; h(y^{d0}(\omega))\;,$$

$$\mu_1^{0s}(\omega) := \sigma_1^{s0}(\omega)\; h(y_1^{s0}(\omega_1),0) \quad \text{and} \quad \mu_2^{0s}(\omega) := \sigma_2^{s0}(\omega)\; h(0,y_2^{s0}(\omega_1))\;.$$

A Simple Reform Affecting only Dual Earner Couples. We are considering a simple reform that affects marginal tax rates only in the interior of $\mathcal{Y}$. Hence $h(y)=0$, whenever $y_1=0$ or $y_2=0$ and therefore also $y_{1,\tau}^{s0}(\cdot)=y_{2,\tau}^{s0}(\cdot)=0$. Thus, for $\omega\in\Omega^d(0,h)$ we have

$$r_\tau(\omega,0,h) = \beta_I^{d0}(\omega)+\beta_E^{d0}(\omega) \quad + \quad \mu^{d0}(\omega) \tag{B.28}$$

where

$$\beta_I^{d0}(\omega) = \sigma^{d0}(\omega)\Big(T_{y_1}^0(y^{d0}(\omega))y_{1,\tau}^{d0}(\cdot)+T_{y_2}^0(y^{d0}(\omega))y_{2,\tau}^{d0}(\cdot)\Big)\;. \tag{B.29}$$

Behavioral Responses at the Extensive Margin for a Simple Reform Affecting only Dual Earner Couples: Characterization. For ease of notation, in this paragraph we take the reform direction h as given and write cutoff types and other endogenous objects (indirect utilities, earnings) as functions of ω and τ, as opposed to writing them as a function of ω, τ and h. Thus, we write $\hat\varphi(\omega,\tau)=(\hat\varphi_1(\omega,\tau),\hat\varphi_2(\omega,\tau))$. Henceforth we denote by $\sigma^d(\hat\varphi(\omega,\tau),\omega)$ the fraction of type ω couples so that both are generating income. Analogously, we denote by $\sigma_1^s(\hat\varphi(\omega,\tau),\omega)$ and $\sigma_2^s(\hat\varphi(\omega,\tau),\omega)$ the fractions of type ω couples so that only spouse 1 or only spouse 2 is generating income. Note that

$$\sigma^d(\hat\varphi(\omega,\tau),\omega) = \int_0^{\hat\varphi_1(\omega,\tau)}\int_0^{\hat\varphi_2(\omega,\tau)} f^\varphi(\varphi_1,\varphi_2\mid\omega)\;d\varphi_2\;d\varphi_1\;,$$

where the density f^φ characterizes the joint distribution of φ_1 and φ_2, conditional on the couples' productivities ω. Moreover,

$$\begin{aligned}\sigma_1^s(\hat\varphi(\omega,\tau),\omega) &= \textstyle\int_0^{\hat\varphi_1(\omega,\tau)}\int_{\hat\varphi_2(\omega,\tau)}^\infty f^\varphi(\varphi_1,\varphi_2\mid\omega)\;d\varphi_2\;d\varphi_1 \\ &\quad + \textstyle\int_{\hat\varphi_1(\omega,\tau)}^{\hat{\hat\varphi}_1}\int_{\varphi_1-\Delta v^s(\omega)}^\infty f^\varphi(\varphi_1,\varphi_2\mid\omega)\;d\varphi_2\;d\varphi_1\;.\end{aligned}$$

Analogously,

$$\begin{aligned}\sigma_2^s(\hat\varphi(\omega,\tau),\omega) &= \textstyle\int_{\hat\varphi_1(\omega,\tau)}^\infty\int_0^{\hat\varphi_2(\omega,\tau)} f^\varphi(\varphi_1,\varphi_2\mid\omega)\;d\varphi_2\;d\varphi_1 \\ &\quad + \textstyle\int_{\hat\varphi_2(\omega,\tau)}^{\hat{\hat\varphi}_2}\int_{\Delta v^s(\omega)+\varphi_2}^\infty f^\varphi(\varphi_1,\varphi_2\mid\omega)\;d\varphi_1\;d\varphi_2\;.\end{aligned}$$

Note that a reform that affects only dual earner couples, does neither affect Δv^s nor $\hat{\hat{\varphi}}_1$ and $\hat{\hat{\varphi}}_2$, see Figure B.3 for an illustration. Hence, for the purpose of this paragraph, we treat these objects only as functions of ω, but not as functions of τ. (This will be different below, when we consider reforms that affect single-earner couples.)

For later use, we compute the derivatives of σ^d, σ_1^s and σ_2^s with respect to τ. For any of these quantities we indicate a derivative with respect to the first argument $\hat{\varphi}_1$ with a subscript $\hat{\varphi}_1$, and a derivative with respect to the second argument $\hat{\varphi}_2$ with a subscript $\hat{\varphi}_2$. Repeated applications of both the chain rule and Leibniz' rule yield:

$$\sigma^d_\tau(\hat{\varphi}(\tau,\omega),\omega) \;:=\; \hat{\varphi}_{2,\tau}(\omega,\tau)\;\sigma^d_{\hat{\varphi}_2}(\hat{\varphi}(\omega,\tau),\omega) + \hat{\varphi}_{1,\tau}(\omega,\tau)\;\sigma^d_{\hat{\varphi}_1}(\hat{\varphi}(\omega,\tau),\omega)$$

where

$$\sigma^d_{\hat{\varphi}_2}(\hat{\varphi}(\omega,\tau),\omega) := \int_0^{\hat{\varphi}_1(\omega,\tau)} f^\varphi(\varphi_1,\hat{\varphi}_2(\omega,\tau)\mid\omega)\;\; d\varphi_1$$

and

$$\sigma^d_{\hat{\varphi}_1}(\hat{\varphi}(\omega,\tau),\omega) := \int_0^{\hat{\varphi}_2(\omega,\tau)} f^\varphi(\hat{\varphi}_1(\omega,\tau),\varphi_2\mid\omega)\; d\varphi_2 \quad .$$

Analogously,

$$\sigma^s_{1,\tau}(\hat{\varphi}(\omega,\tau),\omega) \;:=\; -\hat{\varphi}_{2\tau}(\omega,\tau)\;\sigma^d_{\hat{\varphi}_2}(\hat{\varphi}(\omega,\tau),\omega)$$

and

$$\sigma^s_{2,\tau}(\hat{\varphi}(\omega,\tau),\omega) \;:=\; -\hat{\varphi}_{1\tau}(\omega,\tau)\;\sigma^d_{\hat{\varphi}_1}(\hat{\varphi}(\omega,\tau),\omega) \quad .$$

Note in particular that

$$\sigma^d_\tau(\hat{\varphi}(\tau,\omega),\omega) \;=\; -\sigma^s_{1,\tau}(\hat{\varphi}(\tau,\omega),\omega) - \sigma^s_{2,\tau}(\hat{\varphi}(\tau,\omega),\omega)\;.$$

From the definitions $\hat{\varphi}_1(\omega,\tau) := v^d(\omega,\tau) - v_2^s(\omega_2,\tau)$ and $\hat{\varphi}_2(\omega,\tau) := v^d(\omega,\tau) - v_1^s(\omega_1,\tau)$ and the envelope theorem it follows that

$$-\hat{\varphi}_{1,\tau}(\omega,\tau) \;=\; -\hat{\varphi}_{2,\tau}(\omega,\tau) \;=\; h(y^*(\tau,\omega)) \;=$$

$$\begin{cases} y_2^*(\tau,\omega) - y_{2s}, & \text{if} \quad y_1^*(\tau,\omega)\in B_1 \text{ and } y_2^*(\tau,\omega)\in B_2 \\ \ell_2, & \text{if} \quad y_1^*(\tau,\omega)\in B_1 \text{ and } y_2^*(\tau,\omega)\geq y_{2s}+\ell_2 \\ 0, & \text{otherwise,} \end{cases}$$

where, for ease of notation, we have suppressed the dependence of y^* on the reform direction h. If we now evaluate all these terms in the status quo, i.e. at $\tau = 0$, we obtain the expressions that are collected in the following Lemma.

Lemma 5.

$$\sigma_{\tau}^{d0}(\omega) = -h(y^0(\omega))\Big(\sigma_{\hat{\varphi}_2}^{d}(\cdot) + \sigma_{\hat{\varphi}_1}^{d}(\cdot)\Big)_{|\tau=0} , \tag{B.30}$$

$$\sigma_{1,\tau}^{s0}(\omega) = h(y^0(\omega))\ \sigma_{\hat{\varphi}_2}^{d}(\cdot)_{|\tau=0} , \tag{B.31}$$

and

$$\sigma_{2,\tau}^{s0}(\omega) = h(y^0(\omega))\ \sigma_{\hat{\varphi}_1}^{d}(\cdot)_{|\tau=0} . \tag{B.32}$$

Upon substituting these expressions in to the above definition of $\beta_E^{d0}(\omega)$ we obtain the following Corollary.

Corollary 5.

$$\beta_E^{d0}(\omega) = -h(y^{d0}(\omega))\left(\sigma_{\hat{\varphi}_1}^{d}(\cdot)_{|\tau=0}\Delta_{-1}^{T0}(\omega)\ +\ \sigma_{\hat{\varphi}_2}^{d}(\cdot)_{|\tau=0}\Delta_{-2}^{T0}(\omega)\right) , \tag{B.33}$$

where

$$\Delta T_{-1}^{0}(\omega) \quad := \quad T^0(y^{d0}(\omega)) - T^0(0, y_2^{s0}(\omega_2))$$

and

$$\Delta T_{-2}^{0}(\omega) \quad = \quad T^0(y^{d0}(\omega)) - T^0(y_1^{s0}(\omega_1), 0) .$$

Ceteris paribus, $\hat{\varphi}_2^0$ is one-by-one decreasing in $\Delta_{-2}^{T0}(\omega)$ and that $\hat{\varphi}_1^0$ is one-by-one decreasing in $\Delta_{-1}^{T0}(\omega)$.[7] Therefore, we can as well interpret $\sigma_{\hat{\varphi}_1}^{d}(\cdot)$ as a derivative of $\sigma^d(\cdot)$ with respect to $\Delta T_{-1}^{0}(\omega)$ and $\sigma_{\hat{\varphi}_2}^{d}(\cdot)$ as a derivative of $\sigma^d(\cdot)$ with respect to $\Delta T_{-2}^{0}(\omega)$. We write

$$\chi_{-1}^{d0}(\omega) \quad := \quad \frac{\sigma_{\hat{\varphi}_1}^{d}(\cdot)_{\tau=0}}{\sigma^{d0}(\omega)}\ \Delta T_{-1}^{0}(\omega) .$$

---

[7] To see this, note that, $\hat{\varphi}_1(\omega,\tau) := v^d(\omega,\tau) - v_2^s(\omega_2,\tau)$ can be written as

$$\begin{aligned} \hat{\varphi}_1(\omega,\tau) &= \Delta\mathbf{s}_{-1}(\omega,\tau)\ -\ \Big(T^0(y^d(\omega,\tau)) - T^0(0, y_2^s(\omega_2,\tau)) + \tau\ h(y^d(\omega,\tau))\Big) \\ &= \Delta\mathbf{s}_{-1}(\omega,\tau) - \Big(\Delta T_{-1}(\tau,\omega) + \tau\ h(y^d(\omega,\tau))\Big) \end{aligned} \tag{B.34}$$

where $\mathbf{s}_{-i}(\omega,\tau) := \mathbf{s}^d(\omega,\tau) - \mathbf{s}_i^s(\omega,\tau)$ and $\mathbf{s}^d(\omega,\tau) := y^d(\omega,\tau) - k_1(y_1^d(\omega,\tau),\omega_1) - k_2(y_2^d(\omega,\tau),\omega_2)$ is the surplus (i.e. output minus variable cost) generated by a dual-earner couple with productive abilities ω, whereas $\mathbf{s}_i^s(\omega,\tau) := y_i^s(\omega_i,\tau) - k_i(y_i^s(\omega_i,\tau),\omega_i)$ is the surplus generated by a couple with spouse i as the single-earner. Thus, $\mathbf{s}_{-i}(\omega,\tau)$ is the surplus that is lost when spouse $-i$ no longer opts for positive earnings. In the status quo, (B.34) becomes

$$\hat{\varphi}_1^0(\omega) \quad = \quad \Delta\mathbf{s}_{-1}^0(\omega)\ -\ \Delta T_{-1}^0(\omega) . \tag{B.35}$$

for the elasticity of σ^d with respect to $\Delta_{-1}^{T0}(\omega)$ in the status quo. Analogously, we define

$$\chi_{-2}^{d0}(\omega) \quad := \quad \frac{\sigma_{\hat{\varphi}_2}^d(\cdot)_{\tau=0}}{\sigma^{d0}(\omega)} \, \Delta T_{-2}^0(\omega) \, .$$

Armed with this notation we rewrite the expression for $\beta_E^{d0}(\omega)$ obtained in Corollary 5

$$\beta_E^{d0}(\omega) \quad := \quad -h(y^{d0}(\omega)) \, \sigma^{d0}(\omega) \Big(\chi_{-1}^{d0}(\omega) + \chi_{-2}^{d0}(\omega) \Big) \, .$$

Upon letting

$$\bar{\chi}^{d0}(\omega) \quad := \quad \chi_{-1}^{d0}(\omega) + \mu_{-2}^{d0}(\omega)$$

and interpreting this object as the exit rate – the percentage change of $\sigma_{d0}(\omega)$ that results when joint income is taxed more – we can write this even more concisely, see the following Lemma which summarizes the analysis in this paragraph.

Lemma 6.

$$\beta_E^{d0}(\omega) := -h(y^{d0}(\omega)) \;\; \sigma^{d0}(\omega) \; \bar{\chi}^{d0}(\omega) \, . \tag{B.36}$$

Corollary 6.

$$\beta_E^{d0}(\omega) \; + \; \mu^{d0}(\omega) \quad = \quad \sigma^{d0}(\omega)(1 - \bar{\chi}^{d0}(\omega)) h(y^{d0}(\omega)) \, . \tag{B.37}$$

The corollary illuminates the relation between a setup that has only behavioral responses at the intensive margin and one that has behavioral responses both at intensive and the extensive margin: Behavioral responses at the extensive margin response are as if there were only intensive margin responses but weaker mechanical effects of a tax reform. For ease of reference, we write henceforth

$$r_\tau(\omega, 0, h) \quad = \quad \beta_I^{d0}(\omega) \quad + \quad \mu_E^{d0}(\omega) \; h(y^{d0}(\omega)) \tag{B.38}$$

where

$$\beta_I^{d0}(\omega) \quad = \quad \sigma^{d0}(\omega) \Big(T_{y_1}^0(y^{d0}(\omega)) y_{1,\tau}^{d0}(\cdot) + T_{y_2}^0(y^{d0}(\omega)) y_{2,\tau}^{d0}(\cdot) \Big) \, . \tag{B.39}$$

and

$$\mu_E^{d0}(\omega) \quad = \quad \sigma^{d0}(\omega)(1 - \bar{\chi}^{d0}(\omega)) \tag{B.40}$$

is our notation for extensive-margin inclusive mechanical effects, per dollar of tax increase.

Aggregation. Note that $\beta_I^{d0}(\omega)$ is zero for $\omega_2 \in \Omega_2^+[B, \omega_1]$ In this case, marginal tax rates do not change so that there is no behavioral response at the intensive margin. We can now write

$$\mathbb{E}_\omega[r_\tau(\omega, 0, h)\ \mathbf{1}(\omega \in \Omega^d(0,h))] \quad =: \quad \mathbf{B}(y_{1s}, \ell_1, y_{2s}, \ell_2) + \mathbf{M}(y_{1s}, \ell_1, y_{2s}, \ell_2)\ ,$$

where

$$\mathbf{B}_I(y_{1s}, \ell_1, y_{2s}, \ell_2) \quad = \quad \int\limits_{\Omega_1(B)} \int\limits_{\Omega_2^{in}(B,\omega_1)} \beta_I^{d0}(\omega)\, dF(\omega)\ ,$$

and

$$\begin{aligned}\mathbf{M}(y_{1s}, \ell_1, y_{2s}, \ell_2) \quad &= \quad \int\limits_{\Omega_1(B)} \int\limits_{\Omega_2^{in}(B,\omega_1)} \mu_E^{d0}(\omega)\Big\{y_2^0(\omega) - y_{2s}\Big\} dF(\omega) \\ &\quad + \ell_2 \int\limits_{\Omega_1(B)} \int\limits_{\Omega_2^{+}(B,\omega_1)} \mu_E^{d0}(\omega) dF(\omega)\ .\end{aligned}$$

and $dF(\omega)$ is a shorthand for $f_{\omega_2}(\omega_2 \mid \omega_1)d\omega_2\ f_{\omega_1}(\omega_1)d\omega_1$.

Revenue Implications for $\ell_2 \to 0$. We now compute the derivative of

$$\mathbf{R}(y_{1s}, \ell_1, y_{2s}, \ell_2) \quad := \quad \mathbf{B}(y_{1s}, \ell_1, y_{2s}, \ell_2) + \mathbf{M}(y_{1s}, \ell_1, y_{2s}, \ell_2)$$

with respect to ℓ_2 and evaluate the resulting expressions at $\ell_2 = 0$. Note that in changing ℓ_2 we are changing the length of the bracket B_2 and therefore also the types $\overline{\omega}_2^{in}(y_{2s} + \ell_2 \mid \omega_1)$ and $\underline{\omega}_2^{+}(y_{2s} + \ell_2 \mid \omega_1)$ on the boundaries of $\Omega_2^{in}(B, \omega_1)$ and $\Omega_2^{+}(B, \omega_1)$. Note that the set $\Omega_1(B)$ does not change as ℓ_2 varies. This yields

$$\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \quad = \quad \mathbf{B}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \ + \ \mathbf{M}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)\ .$$

where

$$\mathbf{M}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \ = \int\limits_{\Omega_1(B)} \int\limits_{\Omega_2^{+}(y_{2s},\omega_1)} \mu_E^{d0}(\omega) dF(\omega)\ .$$

and

$$\Omega_2^{+}(y_{2s}, \omega_1) \quad := \quad \{\omega_2 \mid y_2^{d0}(\omega) \geq y_{2s}\ , \text{for } \omega_1 \in \Omega_1(B)\}\ .$$

Moreover,

$$\mathbf{B}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \ =$$

$$\int\limits_{\underline{\omega}_1(y_{1s})}^{\overline{\omega}_1(y_{1s}+\ell_1)} \beta_I^{d0}(\omega_1, \hat{\omega}_2) f_{\omega_2}(\hat{\omega}_2 \mid \omega_1)\ \left(\tfrac{\overline{\omega}_2^{in}(y_{2s}+\ell_2|\omega_1)}{\partial \ell_2}\right) |_{\ell_2=0}\ f_{\omega_1}(\omega_1) d\omega_1\ .$$

for $\hat{\omega}_2 = \overline{\omega}_2^{in}(y_{2s} \mid \omega_1)$.

Rewriting $\mathbf{M}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$. Note that $\mathbf{M}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$ can equivalently be written as

$$\mathbf{M}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) = P(\Theta') \, \mathbb{E}_\omega \left[\mu_E^{d0}(\omega) \mid \theta \in \Theta' \right] ,$$

where

$$\begin{aligned} P(\Theta') &= P(\{\theta \mid y_1^0(\theta) \in [y_{1s}, y_{1s} + \ell_1] \text{ and } y_2^0(\theta) \geq y_{2s}\}) \\ &= P(\{\theta \mid y_2^0(\theta) \geq y_{2s}\} \mid y_1^0(\theta) \in [y_{1s}, y_{1s} + \ell_1]) \Big(F_{y_1}(y_{1s} + \ell_1) - F_{y_1}(y_{1s}) \Big) \\ &= \Big(1 - F_{y_2}(y_{2s} \mid y_1^0(\theta) \in [y_{1s}, y_{1s} + \ell_1]) \Big) \Big(F_{y_1}(y_{1s} + \ell_1) - F_{y_1}(y_{1s}) \Big) . \end{aligned}$$

Rewriting $\mathbf{B}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$. Recall that

$$\begin{aligned} &\mathbf{B}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \\ &= \int_{\underline{\omega}_1(y_{1s})}^{\overline{\omega}_1(y_{1s}+\ell_1)} \beta_I^{d0}(\omega_1, \overline{\omega}_2^{in}(y_{2s} \mid \cdot)) f\omega_2(\overline{\omega}_2^{in}(y_{2s} \mid \cdot) \mid \omega_1) \left(\tfrac{\overline{\omega}_2^{in}(y_{2s}+\ell_2 \mid \cdot)}{\partial \ell_2} \right) \mid_{\ell_2 = 0} f_{\omega_1}(\omega_1) d\omega_1 . \end{aligned}$$

For any y_2', and any ω_1, let

$$F_{y_2}(y_2' \mid \omega_1) \quad := \quad F_{\omega_2}(\bar{\omega}_2(y_2' \mid \omega_1) \mid \omega_1) ,$$

where $\bar{\omega}_2(y_2' \mid \omega_1)$ is the largest value of ω_2 so that $y_2^{d0}(\omega_1, \omega_2) \leq y_2'$. Upon differentiating with respect to y_2', we obtain

$$f_{y_2}(y_2' \mid \omega_1) \quad := \quad f_2(\bar{\omega}_2(y_2' \mid \omega_1) \mid \omega_1) \frac{\partial}{\partial y_2'} \bar{\omega}_2(y_2' \mid \omega_1) .$$

We can therefore write

$$\begin{aligned} &\mathbf{B}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \\ &= \int_{\underline{\omega}_1(y_{1s})}^{\overline{\omega}_1(y_{1s}+\ell_1)} \beta_I^{d0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1) \, f_{\omega_1}(\omega_1) d\omega_1 , \\ &= \Big(F^{y_1}(y_{1s} + \ell_1) - F^{y_1}(y_{1s}) \Big) \mathbb{E}_{\omega_1} \Big[\beta_I^{d0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f^{y_2}(y_{2s} \mid \omega_1) \mid y_1^0(\theta) \in [y_{1s}, y_{1s} + \ell_1] \Big] . \end{aligned}$$

Rewriting $\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$. Recall that

$$\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \;=\; \mathbf{B}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \;+\; \mathbf{M}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)\;.$$

Using the results above this can be written as

$$\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \;=\; \Big(F^{y_1}(y_{1s}+\ell_1) - F^{y_1}(y_{1s})\Big)\Big(\mathcal{B}(y_{1s}, \ell_1, y_{2s}) + \mathcal{M}(y_{1s}, \ell_1, y_{2s})\Big)\;,$$

where

$$\mathcal{B}(y_{1s}, \ell_1, y_{2s}) \;:=\; \mathbb{E}_{\omega_1}\Big[\beta_I^{d0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1^{d0}(\omega) \;\in\; B_1\Big]$$

and

$$\mathcal{M}(y_{1s}, \ell_1, y_{2s}) \;:=\; \Big(1 - F_{y_2}(y_{2s} \mid y_1^{d0}(\omega) \;\in\; B_1)\Big)\mathbb{E}_{\omega}\left[\mu_E^{d0}(\omega) \mid \omega \in \Omega'\right]$$

Henceforth, for brevity, we will write $\mathcal{R}(y_{1s}, \ell_1, y_{2s})$ rather than $\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$.

Taking the Derivative with Respect to ℓ_1 and Evaluating at $\ell_1 = 0$. This yields

$$\mathcal{R}_{\ell_1}(y_{1s}, 0, y_{2s}) \;=\; f^{y_1}(y_{1s})\Big(\mathcal{B}(y_{1s}, 0, y_{2s}) + \mathcal{M}(y_{1s}, 0, y_{2s})\Big)\;,$$

where

$$\mathcal{B}(y_{1s}, 0, y_{2s})) \;=\; \mathbb{E}_{\omega_1}\Big[\beta_I^{d0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1^{0}(\theta) \;=\; y_{1s}\Big]$$

and

$$\mathcal{M}(y_{1s}, 0, y_{2s})) \;=\; \Big(1 - F_{y_2}(y_{2s} \mid y_{1s})\Big)\mathbb{E}_{\omega}\left[\mu_E^{d0}(\omega) \mid \theta \in \Theta''\right]$$

where

$$\Theta'' \;:=\; \{\theta \mid y_1^0(\theta) = y_{1s} \text{ and } y_2^0(\theta) \geq y_{2s}\}\;.$$

Upon writing, more compactly,

$$\mathcal{B}(y_{1s}, 0, y_{2s}) \;=:\; \bar{\beta}_I^{d0}(y_{1s}, y_{2s})\;,$$

$$\mathbb{E}_{\omega}\left[\mu_E^{d0}(\omega) \mid \theta \in \Theta''\right] \;=:\; \bar{\mu}_E^{d0}(y_{1s}, y_{2s})\;,$$

and

$$\mathcal{R}(y_{2s} \mid y_{1s}) \;:=\; \mathcal{R}_{\ell_1}(y_{1s}, 0, y_{2s})\;,$$

we ultimately obtain

$$\mathcal{R}(y_{2s} \mid y_{1s}) \quad = \quad f^{y_1}(y_{1s}) \left(\bar{\beta}_I^{d0}(y_{1s}, y_{2s}) + \Big(1 - F^{y_2}(y_{2s} \mid y_{1s})\Big) \bar{\mu}_E^{d0}(y_{1s}, y_{2s}) \right) .$$

B.1.7 Conditional Revenue Functions: Single-Earner Couples

The revenue functions for single earner couples are shaped by similar forces as the ones for dual earner couples. For concreteness, we characterize the revenue function $y_2 \mapsto \mathcal{R}_2^s(y_2)$. Recall that $\mathcal{R}_2^s(y_{2s})$ is the marginal effect on tax revenue associated with an increase on the earnings of spouse 2, conditional on spouse 2 having earnings in narrow bracket that starts at y_{2s} and conditional on spouse 2 being a single earner.[8]

The Intensive Margin. Specifically, the revenue implications of intensive margin behavioral response are captured by

$$\beta_{2,I}^{s0}(\omega) \quad := \quad \sigma_2^{s0}(\omega)\, T_{y_2}^0(0, y_2^{s0}(\omega_2))\, y_{2,\tau}^{s0}(\cdot) , \quad \text{and}$$

$$\bar{\beta}_{2,I}^{s0}(y_{2s}) \quad := \quad \mathbb{E}_{\omega_1}\Big[\beta_{2,I}^{s0}(\omega_1, \bar{\omega}_2(y_{2s})) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1 = 0 \Big] .$$

The term $\beta_{2,I}^{s0}(\omega)$ looks at the set of couples with productive abilities ω. Of these couples, a fraction $\sigma_2^{s0}(\omega)$ are single earner couples with $y_1 = 0$ and y_2 in the range where marginal tax rates change. Then, $\bar{\beta}_{2,I}^{s0}(y_{2s})$ is obtained as the average of $\beta_{2,I}^{s0}(\omega)$ among all such couples, again, taking account of the possibility that the status quo tax schedule has a kink at $y = (0, y_{2s})$.

Mechanical Effects and the Extensive Margin. The reform comes with an increase of the tax burden for single earner couples with $y_1 = 0$ and $y_2 \geq y_{2s}$. Consider couples with productive ability ω, who conditional on being a single earner couple with $y_1 = 0$, choose $y_2 \geq y_{2s}$ in the status quo. The reform implies that some of these couples become dual-earner couples, some are turned into couples with spouse 1 as the single earner and some are turned into couples with no earnings, see Figure B.3. The percentage change in the mass of these single earner couples is denoted by $\mathcal{E}_{x2}^s(y_{2s})$.

Proposition 7.

$$\mathcal{R}_2^s(y_{2s}) \quad = \quad F_{y_1}(0) \Big(\bar{\beta}_{2,I}^{s0}(y_{2s}) + (1 - F_{y_2}(y_{2s} \mid 0))\ (1 - \mathcal{E}_{x2}^s(y_{2s})) \Big) ,$$

[8] A characterization of $y_1 \mapsto \mathcal{R}_1^s(y_1)$ can be obtained along the same lines.

where $y_1 \mapsto F^{y_1}(y_1)$ *is the cdf associated with the marginal distribution of* y_1 *and* $y_2 \mapsto F^{y_2}(y_{2s} \mid 0)$ *is the cdf of* y_2*, conditional on* y_1 *being equal to* 0*.*

Proof of Proposition 7

We are now considering a reform (τ, h), with

$$h(y) \;=\; \begin{cases} y_2 - y_{2s}, & \text{if} \quad y \in \{0\} \times B_2(y_{2s}, \ell_2) \;, \\ \ell_2, & \text{if} \quad y_2 \geq y_{2s} + \ell_2 \quad \text{and} \quad y_1 = 0 \;, \\ 0, & \text{otherwise} \;, \end{cases}$$

that raises marginal tax rates on the earnings of spouse 2 conditional on $y_2 \in B_2(y_{2s}, \ell_2)$ and spouse 1 having no earnings, $y_1 = 0$.

Consider the set of couples with productivity $\omega = (\omega_1, \omega_2)$ that includes a subset with $y_1^0(\theta) = 0$ and $y_2^0(\theta) \geq y_{2s}$ in the status quo. Figure B.3 shows how the reform affects the solution to the couple's discrete choice problem: In particular, the blue line indicates how the mass of couples with spouse 2 as a single earner shrinks, this is the extensive margin response. In addition, there is an intensive margin behavioral response as the earnings incentives of spouse 2, conditional on being a single earner, change. Equations (B.27) therefore simplifies to

$$r_\tau(\omega, 0, h) \;=\; \beta_{2,I}^{s0}(\omega) \;+\; \mu_2^{s0}(\omega) \;+\; \beta_{2,E}^{s0}(\omega) \;, \tag{B.41}$$

where

$$\begin{aligned} \beta_{2,I}^{s0}(\omega) \quad &:= \quad \sigma_2^{s0}(\omega)\Big(T_{y_2}^0(0, y_2^{s0}(\omega_2)) y_{2,\tau}^{s0}(\cdot)\Big) \;, \\ \mu_2^{s0}(\omega) \quad &:= \quad \sigma_2^{s0}(\omega) h(0, y_2^{s0}(\omega_2)) \;, \end{aligned}$$

and

$$\begin{aligned} \beta_{2,E}^{s0}(\omega) \;:=\; & \sigma_\tau^{d0}(\omega, h)\; T^0(y^{d0}(\omega)) \\ & + \sigma_{1,\tau}^{s0}(\omega, h)\; T^0(y_1^{s0}(\omega_1), 0) \\ & + \sigma_{2,\tau}^{s0}(\omega, h)\; T^0(0, y_2^{s0}(\omega_2)) \;. \end{aligned} \tag{B.42}$$

Comparative Statics at the Extensive Margin. Henceforth, we denote by

$$\begin{aligned} & \sigma_2^s(\hat{\varphi}_1(\tau, \omega), \hat{\hat{\varphi}}_2(\tau, \omega), \Delta v^s(\tau, \omega), \omega) \\ & = \int_{\hat{\varphi}_1(\tau,\omega)}^{\infty} \int_0^{\hat{\varphi}_2} f^\varphi(\varphi_1, \varphi_2 \mid \omega)\; d\varphi_2\; d\varphi_1 \\ & \quad + \int_{\hat{\varphi}_2}^{\hat{\hat{\varphi}}_2(\tau,\omega)} \int_{\Delta v^s(\tau,\omega)\,+\,\varphi_2}^{\infty} f^\varphi(\varphi_1, \varphi_2 \mid \omega)\; d\varphi_1\; d\varphi_2 \;. \end{aligned}$$

the mass of couples where only spouse 2 generates income. Let $\sigma^s_{2,\hat{\varphi}_1}(\cdot)$, $\sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot)$ and $\sigma^s_{2,\Delta v^s}(\cdot)$ denote, respectively, the derivative of $\sigma^s_2(\cdot)$ with respect to its first, second and third argument. An application of Leibniz' rule yields

$$\begin{aligned}&\sigma^s_{2\tau}(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta v^s(\tau,\omega),\omega)\\&=\sigma^s_{2,\hat{\varphi}_1}(\cdot)\;\hat{\varphi}_{1\tau}(\tau,\omega)\quad+\quad\sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot)\;\hat{\hat{\varphi}}_{2\tau}(\tau,\omega)\quad+\quad\sigma^s_{2,\Delta v^s}(\cdot)\;\Delta v^s_\tau(\tau,\omega)\;.\end{aligned}$$

where

$$\begin{aligned}\sigma^s_{2,\hat{\varphi}_1}(\cdot)\quad&=\quad-\int_0^{\hat{\varphi}_2}f^\varphi(\varphi_1(\tau,\omega),\varphi_2\mid\omega)d\varphi_2\;,\\\sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot)\quad&=\quad\int_{\hat{\hat{\varphi}}_1}^{\infty}f^\varphi(\varphi_1,\hat{\hat{\varphi}}_2(\tau,\omega)\mid\omega)d\varphi_1\;,\end{aligned}$$

and

$$\sigma^s_{2,\Delta v^s}(\cdot)\quad=\quad-\int_{\hat{\varphi}_2}^{\hat{\hat{\varphi}}_2(\tau,\omega)}f^\varphi(\Delta v^s(\tau,\omega)+\varphi_2,\varphi_2\mid\omega)d\varphi_2\;.$$

We denote by $\sigma^s_1(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta v^s(\tau,\omega),\omega)$ and $\sigma^d(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta v^s(\tau,\omega),\omega)$, respectively, the mass of couples where only spouse 1 generates income and the mass where both spouses generate income. We note that

$$\sigma^d_\tau(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta_s(\tau,\omega),\omega)\quad=\quad-\sigma^s_{2,\hat{\varphi}_1}(\cdot)\;\hat{\varphi}_{1\tau}(\tau,\omega)$$

and that

$$\sigma^s_{1\tau}(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta v^s(\tau,\omega),\omega)\quad=\quad-\sigma^s_{2,\Delta v^s}(\cdot)\;\Delta v^s_\tau(\tau,\omega)$$

Moreover, from the envelope theorem it follows that

$$\hat{\varphi}_{1\tau}(\tau,\omega)\;=\Delta v^s_\tau(\tau,\omega)\;=\;-\hat{\hat{\varphi}}_{2\tau}(\tau,\omega)\;=\;h(0,y_2^{s*}(\tau,\omega))\;=$$

$$\begin{cases}0, & \text{if}\quad y_2^{s*}(\tau,\omega)\le y_{2s}\\ y_2^{s*}(\tau,\omega)-y_{2s}, & \text{if}\quad y_2^{s*}(\tau,\omega)\in[y_{2s},y_{2s}+\ell_2]\\ \ell_2, & \text{if}\quad y_2^{s*}(\tau,\omega)\ge y_{2s}+\ell_2\end{cases}$$

where, for ease of notation, we have suppressed the dependence of y_2^{s*} on the reform direction h. If we now evaluate all this expressions in the status quo, i.e. at $\tau=0$, we obtain the following Lemma.

Lemma 7.

$$\sigma^{d0}_\tau(\omega)=-h(0,y_2^{s0}(\omega_2))\;\sigma^s_{2,\hat{\varphi}_1}(\cdot)_{|\tau=0}\;,\tag{B.43}$$

$$\sigma^{s0}_{1,\tau}(\omega)=-h(0,y_2^{s0}(\omega_2))\;\sigma^s_{2,\Delta v^s}(\cdot)_{|\tau=0}\;,\tag{B.44}$$

and

$$\sigma^{s0}_{2,\tau}(\omega) = h(0, y^{s0}_2(\omega_2))\Big(\sigma^s_{2,\hat\varphi_1}(\cdot) + \sigma^s_{2,\Delta v^s}(\cdot) + \sigma^s_{2,\hat{\hat\varphi}_2}(\cdot)\Big)_{|\tau=0} . \tag{B.45}$$

Corollary 7.

$$\begin{aligned}\beta^{d0}_E(\omega) \;=\; -h(0, y^{s0}_2(\omega_2))\Big(&\sigma^s_{2,\hat\varphi_1}(\cdot)_{|\tau=0}\Delta^{T0}_{-1}(\omega) \;+\; \sigma^s_{2,\Delta v^s}(\cdot)_{|\tau=0}\Delta^{T0}_{2-1}(\omega)\\ &-\sigma^s_{2,\hat{\hat\varphi}_2}(\cdot)_{|\tau=0}T^0(0, y^{s0}_2(\omega_2))\Big) ,\end{aligned} \tag{B.46}$$

where

$$\begin{aligned}\Delta T^0_{-1}(\omega) \;&:=\; T^0(y^{d0}(\omega)) - T^0(0, y^{s0}_2(\omega_2)) , \quad \text{and}\\ \Delta T^0_{2-1}(\omega) \;&:=\; T^0(y^{s0}_1(\omega_1), 0) - T^0(0, y^{s0}_2(\omega_2)) .\end{aligned}$$

Note that, ceteris paribus, $\hat\varphi^0_1$ and $\Delta v^s(\tau, \omega)$ are one-by-one increasing in $T_0(0, y^{s0}_2(\omega_2))$ and that $\hat{\hat\varphi}^0_2$ is one-by-one decreasing in $T_0(0, y^{s0}_2(\omega_2))$. This follows from the definitions of these cutoff types. We can therefore as well interpret $\sigma^s_{2,\hat\varphi_1}(\cdot)$ as a derivative of $\sigma^s_2(\cdot)$ with respect to $-\Delta^{T0}_{-1}(\omega)$, $\sigma^s_{2,\Delta v^s}(\cdot)$ as a derivative of σ^s_2 with respect to $-\Delta^{T0}_{2-1}(\omega)$ and, finally, $\sigma^s_{2,\hat{\hat\varphi}_2}(\cdot)$ as a derivative with respect to $T_0(0, y^{s0}_2(\omega_2))$, see also Footnote 7.We write

$$\chi^{s0}_{2,-1}(\omega) \;:=\; \frac{\sigma^s_{2,\hat\varphi_1}(\cdot)_{\tau=0}}{\sigma^{s0}_2(\omega)}\,\Delta T^0_{-1}(\omega) ,$$

$$\chi^{s0}_{2,2-1}(\omega) \;:=\; \frac{\sigma^s_{2,\Delta v^s}(\cdot)_{\tau=0}}{\sigma^{s0}_2(\omega)}\,\Delta T^0_{2-1}(\omega) ,$$

and

$$\chi^{s0}_{2,-2}(\omega) \;:=\; \frac{\sigma^s_{2,\hat{\hat\varphi}_2}(\cdot)_{|\tau=0}}{\sigma^{s0}_2(\omega)}\,T_0(0, y^{s0}_2(\omega_2)) .$$

Upon collection terms and defining

$$\bar\chi^{s0}_2(\omega) \;:=\; \chi^{s0}_{2,-1}(\omega) + \chi^{s0}_{2,2-1}(\omega) - \chi^{s0}_{2,-2}(\omega)$$

we summarize the analysis in this paragraph with the following Lemma.

Lemma 8.

$$\beta^{s0}_{2,E}(\omega) := -h(0, y^{s0}_2(\omega_2))\;\; \sigma^{s0}_2(\omega)\;\bar\chi^{s0}_2(\omega) . \tag{B.47}$$

Corollary 8.

$$\beta^{s0}_{2,E}(\omega) \;+\; \mu^{s0}_2(\omega) \;=\; \sigma^{s0}_2(\omega)(1 - \bar\chi^{s0}_2(\omega))h(0, y^{s0}_2(\omega_2)) . \tag{B.48}$$

The corollary illuminates the relation between a setup that has only behavioral responses at the intensive margin and one that has behavioral responses both at intensive and the extensive margin: Behavioral responses at the extensive margin response are as if there were only intensive margin responses but weaker mechanical effects of a tax reform. For ease of reference, we write henceforth

$$r_\tau(\omega, 0, h) \;=\; \beta^{s0}_{2,I}(\omega) \quad + \quad \mu^{s0}_{2,E}(\omega)\; h(0, y^{s0}_2(\omega_2)) \tag{B.49}$$

where

$$\beta^{s0}_{2,I}(\omega) \;=\; \sigma^{s0}_2(\omega)\; T^0_{y_2}(0, y^{s0}_2(\omega_2))\; y^{s0}_{2,\tau}(\cdot)\;, \tag{B.50}$$

and

$$\mu^{s0}_{2,E}(\omega) \;=\; \sigma^{s0}_2(\omega)(1 - \bar{\chi}^{s0}_2(\omega)) \tag{B.51}$$

is our notation for extensive-margin inclusive mechanical effects, per dollar of tax increase.

Aggregation. Let $B_2 = [y_{2s}, y_{2s} + \ell_2]$. We can now write

$$\mathbb{E}_\omega[r_\tau(\omega, 0, h)\; \mathbf{1}(\omega \in \Omega^s_2(0, h))] \;=:\; \mathbf{B}(\{0\}, y_{2s}, \ell_2) + \mathbf{M}(\{0\}, y_{2s}, \ell_2)\;,$$

where

$$\mathbf{B}(\{0\}, y_{2s}, \ell_2) \;=\; \int_{\Omega_1} \int_{\Omega^{in}_2(B_2)} \beta^{s0}_{2,I}(\omega)\, dF(\omega)\;,$$

and

$$\begin{aligned}\mathbf{M}(\{0\}, y_{2s}, \ell_2) \;=\; & \int_{\Omega_1} \int_{\Omega^{in}_2(B_2)} \mu^{s0}_{2,E}(\omega)\Big\{y^0_2(\omega) - y_{2s}\Big\} dF(\omega) \\ & + \ell_2 \int_{\Omega_1} \int_{\Omega^{+}_2(B_2)} \mu^{s0}_{2,E}(\omega) dF(\omega)\;.\end{aligned}$$

Revenue Implications for $\ell_2 \to 0$. We now compute the derivative of

$$\mathbf{R}(\{0\}, y_{2s}, \ell_2) \quad := \quad \mathbf{B}(\{0\}, y_{2s}, \ell_2) + \mathbf{M}(\{0\}, y_{2s}, \ell_2)$$

with respect to ℓ_2 and evaluate the resulting expressions at $\ell_2 = 0$. Note that in changing ℓ_2 we are changing the length of the bracket B_2 and therefore also the types $\overline{\omega}^{in}_2(y_{2s} + \ell_2)$ and $\underline{\omega}^{+}_2(y_{2s} + \ell_2)$ on the boundaries of $\Omega^{in}_2(B_2, \omega)$ and $\Omega^{+}_2(B_2)$. This yields

$$\mathbf{R}_{\ell_2}(\{0\}, y_{2s}, 0) \;=\; \mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0) \;+\; \mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0)\;,$$

where

$$\mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0) \;= \int\limits_{\Omega_1} \int\limits_{\Omega_2^+(y_{2s})} \mu_{2,E}^{s0}(\omega) dF(\omega) \; .$$

and

$$\Omega_2^+(y_{2s}) \quad := \quad \{\omega_2 \mid y_2^{s0}(\omega_2) \geq y_{2s}\} \; .$$

Moreover,

$$\mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0) \;=$$

$$\int\limits_{\Omega_1} \beta_{2,I}^{s0}(\omega_1, \hat{\omega}_2) f_{\omega_2}(\hat{\omega}_2 \mid \omega_1) \; \left(\frac{\overline{\omega}_2^{in}(y_{2s}+\ell_2)}{\partial \ell_2} \right) |_{\ell_2=0} \; f_{\omega_1}(\omega_1) d\omega_1 \; .$$

for $\hat{\omega}_2 = \overline{\omega}_2^{in}(y_{2s})$.

Rewriting $\mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0)$. Note that $\mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0)$ can equivalently be written as

$$\mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0) = P\left(\Theta'\right) \mathbb{E}_{\omega} \left[\mu_{2,E}^{s0}(\omega) \mid \theta' \in \Theta' \right] \; ,$$

where

$$\begin{aligned} P\left(\Theta'\right) \;&=\; P\left(\{\theta \mid y_1^0(\theta) = 0 \;\text{ and }\; y_2^0(\theta) \geq y_{2s}\}\right) \\ &=\; P\left(\{\theta \mid y_2^0(\theta) \geq y_{2s}\} \mid y_1^0(\theta) = 0\right) F_{y_1}(0) \\ &=\; \Big(1 - F_{y_2}(y_{2s} \mid y_1^0(\theta) = 0)\Big) F_{y_1}(0) \; . \end{aligned}$$

Rewriting $\mathbf{B}_{\ell_2}((\{0\}, y_{2s}, 0)$. Recall that

$$\begin{aligned} &\mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0) \\ &= \int\limits_{\Omega_1} \beta_{2,I}^{s0}(\omega_1, \overline{\omega}_2^{in}(y_{2s})) f_{\omega_2}(\overline{\omega}_2^{in}(y_{2s}) \mid \omega_1) \; \left(\frac{\overline{\omega}_2^{in}(y_{2s}+\ell_2)}{\partial \ell_2} \right) |_{\ell_2=0} \; f_{\omega_1}(\omega_1) d\omega_1 \; . \end{aligned}$$

For any y_2', and any ω_1, let

$$F_{y_2}(y_2' \mid \omega_1) \quad := \quad F_{\omega_2}(\bar{\omega}_2(y_2') \mid \omega_1) \; ,$$

where $\bar{\omega}_2(y_2')$ is the largest value of ω_2 so that $y_2^{s0}(\omega_2) \leq y_2'$. Upon differentiating with respect to y_2', we obtain

$$f_{y_2}(y_2' \mid \omega_1) \quad := \quad f_2(\bar{\omega}_2(y_2') \mid \omega_1) \frac{\partial}{\partial y_2'} \bar{\omega}_2(y_2') \; .$$

We can therefore write

$$\begin{aligned}
&\mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0) \\
&= \int_{\Omega_1} \beta_{2,I}^{s0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1)\ f_{\omega_1}(\omega_1) d\omega_1\ , \\
&= F^{y_1}(0)\ \mathbb{E}_{\omega_1}\Big[\beta_{2,I}^{s0}(\omega_1, \bar{\omega}_2(y_{2s})) f^{y_2}(y_{2s} \mid \omega_1) \mid y_1^0(\theta) = 0\Big]\ .
\end{aligned}$$

Rewriting $\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$. Recall that

$$\mathbf{R}_{\ell_2}(\{0\}, y_{2s}, 0) \ = \ \mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0) \ + \ \mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0)\ .$$

Using the results above this can be written as

$$\mathbf{R}_{\ell_2}(\{0\}, y_{2s}, 0) \ = \ F^{y_1}(0)\Big(\mathcal{B}(\{0\}, y_{2s}) + \mathcal{M}(\{0\}, y_{2s})\Big)\ ,$$

where

$$\mathcal{B}(\{0\}, y_{2s}) \ := \ \mathbb{E}_{\omega_1}\Big[\beta_{2,I}^{s0}(\omega_1, \bar{\omega}_2(y_{2s})) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1^0(\theta) = 0\Big]$$

and

$$\mathcal{M}(\{0\}, y_{2s}) \ := \ \Big(1 - F_{y_2}(y_{2s} \mid y_1^0(\theta) = 0)\Big)\mathbb{E}_{\omega}\left[\mu_E^{d0}(\omega) \mid \theta \in \Theta'\right]$$

Upon writing

$$\bar{\beta}_{2,I}^{s0}(y_{2s}) \quad := \quad \mathbb{E}_{\omega_1}\Big[\beta_{2,I}^{s0}(\omega_1, \bar{\omega}_2(y_{2s})) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1^0(\theta) = 0\Big]$$

and

$$\bar{\mu}_{2,E}^{s0}(y_{2s}) \quad := \quad \mathbb{E}_{\omega}\left[\mu_{2,E}^{s0}(\omega) \mid \theta \in \Theta'\right]$$

we ultimately, obtain

$$\begin{aligned}
\mathcal{R}_2^s(y_{2s}) \ &:= \ \mathbf{R}_{\ell_2}(\{0\}, y_{2s}, 0) \\
&= \ F_{y_1}(0)\Big(\bar{\beta}_{2,I}^{s0}(y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_1^0(\theta) = 0))\ \bar{\mu}_{2,E}^{s0}(y_{2s})\Big)
\end{aligned}$$

B.1.8 Intensive Margin Responses of Dual Earner Couples

First Order Conditions

We consider a dual earner couple whose choices in the status quo satisfy the first order conditions

$$1 - T^0_{y_1}(y^*_1, y^*_2) = k_{1,y_1}(y_1, \omega_1) \tag{B.52}$$

and

$$1 - T^0_{y_2}(y^*_1, y^*_2) = k_{2,y_2}(y_1, \omega_1) \tag{B.53}$$

We now consider the possibility that the marginal tax rate on the earnings of spouse 1 increases by τ_1 or that the the marginal tax rate on the earnings of spouse 1 increases by τ_2. To capture this possibility, we let $\tau \in \{\tau_1, \tau_2\}$ and interpret both y^*_1 and y^*_2 as functions of τ. Specifically, $y^*_1(\tau)$ and $y^*_2(\tau)$ solve

$$1 - T^0_{y_1}(y^*_1(\tau), y^*_2(\tau)) - \tau \, \mathbf{1}(\tau = \tau_1) = k_{1,y_1}(y^*_1(\tau), \omega_1) \tag{B.54}$$

and

$$1 - T^0_{y_2}(y^*_1(\tau), y^*_2(\tau)) - \tau \, \mathbf{1}(\tau = \tau_2) = k_{2,y_2}(y^*_2(\tau), \omega_1) \; . \tag{B.55}$$

Differentiating these two equations with respect to τ, evaluating the resulting expressions in the status quo, i.e. for $\tau = 0$, and using (B.52) and (B.53), yields the following system of equations

$$\rho^1_{y_1}(y^*) \; \delta y^*_1 \; + \; \rho^1_{y_2}(y^*) \; \delta y^*_2 \; - \mathbf{1}(\tau = \tau_1) = \kappa^1(y^*_1, \omega_1) \; \delta y^*_1 \; , \tag{B.56}$$

where

$$\delta y^*_1 \; := \; \frac{y^*_{1,\tau}}{y^*_1} \quad \text{and} \quad \delta y^*_2 \; := \; \frac{y^*_{2,\tau}}{y^*_2}$$

are the changes of y^*_1 and y^*_2 in percent and

$$\rho^1_{y_1}(y) \quad := \quad \frac{-T^0_{y_1 y_1}(y)}{1 - T^0_{y_1}(y)} \; y_1 \; ,$$

$$\rho^1_{y_2}(y) \quad := \quad \frac{-T^0_{y_1 y_2}(y)}{1 - T^0_{y_1}(y)} \; y_2 \; ,$$

are the elasticities of the net of tax or retention rate on the earnings of spouse 1 with respect to y_1 and y_2, respectively. Note that a tax system that is progressive in that marginal tax rates on y_1 are non-decreasing in y_1 has $\rho^1_{y_1}(y) \leq 0$. A tax system that

has positive jointness so that the marginal tax rates on y_1 are non-decreasing in y_2 has $\rho^1_{y_2}(y) \leq 0$. Finally,

$$\kappa^1(y_1, \omega_1) \quad := \quad \frac{k_{1,y_1y_1}(y_1, \omega_1)}{k_{1,y_1}(y_1, \omega_1)}\, y_1$$

is the elasticity of spouse 1's marginal effort costs with respect to the earnings of spouse 1. Note that, with an iso-elastic effort cost function,

$$k_1(y_1, \omega_1) = \frac{1}{1 + \frac{1}{\varepsilon_1}} \left(\frac{y_1}{\omega_1}\right)^{1+\frac{1}{\varepsilon_1}} ,$$

$$\kappa^1(y_1, \omega_1) \quad = \quad \frac{1}{\varepsilon_1} .$$

Analogously, we obtain

$$\rho^2_{y_1}(y^*)\, \delta y_1^* \; + \; \rho^2_{y_2}(y^*)\, \delta y_2^* \; - \mathbf{1}(\tau = \tau_2) = \kappa^2(y_2^*, \omega_2)\, \delta y_2^* , \tag{B.57}$$

Equations (B.56) and (B.57) can be used to solve for δy_1^* and δy_2^* this yields

$$\delta y_1^* = \frac{1}{D(y^*)} \left(\Big(\rho^2_{y_1}(y^*) - \kappa^2(y_2^*, \omega_1)\Big) \frac{\mathbf{1}(\tau = \tau_1)}{1 - T^0_{y_1}(y)} - \rho^1_{y_2}(y^*) \frac{\mathbf{1}(\tau = \tau_2)}{1 - T^0_{y_2}(y)} \right) \tag{B.58}$$

and

$$\delta y_2^* = \frac{1}{D(y^*)} \left(\Big(\rho^1_{y_2}(y^*) - \kappa^1(y_1^*, \omega_1)\Big) \frac{\mathbf{1}(\tau = \tau_2)}{1 - T^0_{y_2}(y)} - \rho^2_{y_1}(y^*) \frac{\mathbf{1}(\tau = \tau_1)}{1 - T^0_{y_1}(y)} \right) \tag{B.59}$$

where

$$D(y^*) \quad := \quad \Big(\rho^1_{y_1}(y^*) - \kappa^1(y_1^*, \omega_1)\Big)\Big(\rho^2_{y_2}(y^*) - \kappa^2(y_2^*, \omega_1)\Big) \; - \; \rho^1_{y_2}(y^*)\, \rho^2_{y_1}(y^*) .$$

Piece-Wise Linear Tax System. If the tax system is piecewise linear, then, locally, $\rho^1_{y_1}(y^*) = \rho^2_{y_1}(y^*) = \rho^2_{y_2}(y^*) = \rho^2_{y_1}(y^*) = 0$. In this case, (B.58) and (B.59) imply

$$\delta y_1^* = -\frac{1}{\kappa^1(y_1^*, \omega_1)} \, \frac{\mathbf{1}(\tau = \tau_1)}{1 - T^0_{y_1}(y)} \tag{B.60}$$

and

$$\delta y_2^* = -\frac{1}{\kappa^2(y_2^*, \omega_1)} \, \frac{\mathbf{1}(\tau = \tau_2)}{1 - T^0_{y_2}(y)} . \tag{B.61}$$

With iso-elastic effort costs this becomes, moreover,

$$\delta y_1^* = -\varepsilon_1 \, \frac{\mathbf{1}(\tau = \tau_1)}{1 - T^0_{y_1}(y)} \tag{B.62}$$

and

$$\delta y_2^* = -\varepsilon_2 \, \frac{\mathbf{1}(\tau = \tau_2)}{1 - T_{y_2}^0(y)} \, . \tag{B.63}$$

Family Taxation. If there is a system of family taxation, then, for every y there is a number $p(y)$ so that

$$\frac{\rho_{y_1}^1(y)}{y_1} \;=\; \frac{\rho_{y_1}^2(y)}{y_1} \;=\; \frac{\rho_{y_2}^2(y)}{y_2} \;=\; \frac{\rho_{y_2}^1(y)}{y_2} \quad = \quad p(y) \, .$$

Implications for Tax Revenue

Recall that behavioral responses of a dual earner couples affect marginal tax rates via

$$\beta_I^0(\omega) \;\; = \;\; \sigma^{d0}(\omega)\Big(T_{y_1}^0(y^{d0}(\omega))y_{1,\tau}^{d0}(\cdot) + T_{y_2}^0(y^{d0}(\omega))y_{2,\tau}^{d0}(\cdot)\Big)$$

We now use the characterization of $y_{1,\tau}^{d0}(\cdot)$ and $y_{2,\tau}^{d0}(\cdot)$ in the previous section – in particular, equations (B.58) and (B.59) – for an analysis of

$$\begin{aligned} r_\tau^d(y^*) \;\; &:= \;\; T_{y_1}^0(y^*)\; y_{1,\tau}^*(\cdot) \quad + \quad T_{y_2}^0(y^*)\; y_{2,\tau}^*(\cdot) \\ &= \;\; T_{y_1}^0(y^*)\; \delta y_1^*\; y_1^* \quad + \quad T_{y_2}^0(y^*)\; \delta y_2^*\; y_2^* \, . \end{aligned}$$

Straightforward computations yield

$$\begin{aligned} r_\tau^d(y^*) \;\; := \;\; & \tfrac{\mathbf{1}(\tau=\tau_1)}{(1-T_{y_1}^0(y^*))D(y^*)}\Big(T_{y_1}^0(y^*)\, y_1^*\, (\rho_{y_1}^2(y^*) - \kappa^2(y_2^*,\omega_1)) \;-\; T_{y_2}^0(y^*)\, y_2^*\, \rho_{y_1}^2(y^*)\Big) \\ & + \tfrac{\mathbf{1}(\tau=\tau_2)}{(1-T_{y_2}^0(y^*))D(y^*)}\Big(T_{y_2}^0(y^*)\, y_2^*\, (\rho_{y_2}^1(y^*) - \kappa^1(y_2^*,\omega_1)) \;-\; T_{y_1}^0(y^*)\, y_1^*\, \rho_{y_2}^1(y^*)\Big) \, . \end{aligned}$$

Piece-Wise Linear Tax System and Iso-Elastic Effort Costs. In this case, this simplifies to

$$r_\tau^d(y^*) \;\; := \;\; -\mathbf{1}(\tau=\tau_1)\, \tfrac{T_{y_1}^0(y^*)}{1-T_{y_1}^0(y^*)}\; y_1^*\, \varepsilon_1 \quad - \quad \mathbf{1}(\tau=\tau_2)\, \tfrac{T_{y_2}^0(y^*)}{1-T_{y_2}^0(y^*)}\; y_2^*\, \varepsilon_2 \, .$$

Family Taxation. In this case, for every y, there is a number $T'(y)$ so that $T^0_{y_1}(y) = T^0_{y_2}(y) = T'(y)$. Therefore

$$\begin{aligned} r^d_\tau(y^*) \;\; := \;\; & \tfrac{\mathbf{1}(\tau=\tau_1)T'(y^*)}{(1-T'(y^*))D(y^*)}\Big(y^*_1\,(\rho^2_{y_1}(y^*) - \kappa^2(y^*_2,\omega_1)) \;-\; y^*_2\,\rho^2_{y_1}(y^*)\Big) \\ & \tfrac{\mathbf{1}(\tau=\tau_2)T'(y^*)}{(1-T'(y^*))D(y^*)}\Big(y^*_2\,(\rho^1_{y_2}(y^*) - \kappa^1(y^*_2,\omega_1)) \;-\; y^*_1\,\rho^1_{y_2}(y^*)\Big)\,. \end{aligned}$$

Characterization of $\beta^{d0}_{I,2}(\omega)$.

Recall that

$$\beta^{d0}_{I,2}(\omega) \quad = \quad T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot)\,. \tag{B.64}$$

For $\mathbf{1}(\tau = \tau_2) = 1$, it follows from the analysis in the previous section that

$$y^{d0}_{1,\tau}(\omega) \quad = \quad -y^{d0}_1(\omega)\,\frac{\rho^1_{y_2}(y^{d0}(\omega))}{(1-T^0_{y_2}(y^{d0}(\omega))D(y^{d0}(\omega))} \tag{B.65}$$

and

$$y^{d0}_{2,\tau}(\omega) \quad = \quad y^{d0}_2(\omega)\,\frac{\rho^1_{y_2}(y^{d0}(\omega)) - \kappa^1(y^{d0}_1(\omega),\omega_1)}{(1-T^0_{y_2}(y^{d0}(\omega))D(y^{d0}(\omega))}\,. \tag{B.66}$$

By substituting (B.65) and (B.66) into (B.64) we obtain the reform's marginal impact on the couple's tax payment. The resulting expression simplifies when (i) there is a system of family taxation so that, in the status quo, there is a joint marginal tax rate for the earnings of spouse 1 and spouse 2, henceforth denoted by $T^{0\prime}(y^{d0}(\omega)) = T^0_{y_1}(y^{d0}(\omega)) = T^0_{y_2}(y^{d0}(\omega))$, (ii) the status quo tax system is piece-wise-linear so that

$$\rho^1_{y_1}(y) \quad = \quad \rho^1_{y_2}(y) \quad = \quad \rho^2_{y_2}(y) \quad = \quad \rho^2_{y_1}(y) \quad = \quad 0\,,$$

for all y where marginal tax rates are well-defined, and (iii) effort costs are iso-elastic. Then,

$$T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot) \quad = \quad -\frac{T^{0\prime}(y^{d0}(\omega))}{1-T^{0\prime}(y^{d0}(\omega))}\;y^{d0}_2(\omega)\;\varepsilon_2\,.$$

Hence, only the earnings and the Frisch elasticity of spouse 2 matter for the loss of tax revenue due to behavioral responses at the intensive margin.

B.1.9 Proof of Proposition 3

We use the conditions $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$ to derive a candidate solution for optimal tax rates. We then verify that, under the candidate solution it is indeed the case, that $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$.

Fix $(y_1, y_2) \in Y'$. With extensive margin elasticities of zero, and using part i) of Assumption 2, the two optimality conditions $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$ can be written as

$$\bar{\beta}^d_{I,1}(y_1, y_2) + \Big(1 - \lambda_1^* \, \bar{g}_1\Big)\Big(1 - F_{y_1}(y_1 \mid y_2)\Big) = 0 \,, \tag{B.67}$$

and

$$\bar{\beta}^d_{I,2}(y_1, y_2) + \Big(1 - \lambda_2^* \, \bar{g}_2\Big)\Big(1 - F_{y_2}(y_2 \mid y_1)\Big) = 0 \,. \tag{B.68}$$

Part iii) of Assumption 2, implies that there is only one ω with $y^{d0}(\omega) = (y_1, y_2)$. Thus,

$$\bar{\beta}^d_{I,1}(y_1, y_2) = \beta^{d0}_{I,1}(\omega) f_{y_1}(y_1 \mid y_2) \,, \tag{B.69}$$

and

$$\bar{\beta}^d_{I,2}(y_1, y_2) = \beta^{d0}_{I,1}(\omega) f_{y_1}(y_1 \mid y_2) \tag{B.70}$$

Substituting (B.69) and (B.70) into (B.67) and (B.68), using part (ii) of Assumption yields

$$\beta^{d0}_{I,1}(\omega) \frac{1}{y_1} = \Big(1 - \lambda_1^* \, \bar{g}_1\Big) \frac{1}{\bar{\alpha}_1} \,, \tag{B.71}$$

and

$$\beta^{d0}_{I,2}(\omega) \frac{1}{y_2} = \Big(1 - \lambda_2^* \, \bar{g}_2\Big) \frac{1}{\bar{\alpha}_2} \,. \tag{B.72}$$

Recall from part B.1.8 of the Appendix that

$$\beta^{d0}_{I,1}(\omega) = \Big(T^0_{y_1}(y^{d0}(\omega)) y^{d0}_{1,\tau_1}(\cdot) + T^0_{y_2}(y^{d0}(\omega)) y^{d0}_{2,\tau_1}(\cdot)\Big) \,. \tag{B.73}$$

where

$$y^{d0}_{1,\tau_1}(\cdot) = y_1 \, \frac{1}{D(y^*)} \Big(\rho^2_{y_1}(y^*) - \kappa^2(y_2^*, \omega_2)\Big) \frac{1}{1 - T^0_{y_1}(y)} \tag{B.74}$$

and

$$y^{d0}_{2,\tau_1}(\cdot) = -y_2 \, \frac{1}{D(y^*)} \rho^2_{y_1}(y^*) \frac{1}{1 - T^0_{y_1}(y)} \tag{B.75}$$

Analogously,

$$\beta^{d0}_{I,2}(\omega) = \Big(T^0_{y_1}(y^{d0}(\omega)) y^{d0}_{1,\tau_2}(\cdot) + T^0_{y_2}(y^{d0}(\omega)) y^{d0}_{2,\tau_2}(\cdot)\Big) \,. \tag{B.76}$$

where

$$y^{d0}_{1,\tau_2}(\cdot) = -y_1 \, \frac{1}{D(y^*)} \rho^1_{y_2}(y^*) \frac{1}{1 - T^0_{y_2}(y)} \tag{B.77}$$

and

$$y^{d0}_{2,\tau_2}(\cdot) = y_2 \, \frac{1}{D(y^*)} \Big(\rho^1_{y_2}(y^*) - \kappa^1(y^*_1, \omega_1) \Big) \frac{1}{1 - T^0_{y_2}(y)} \tag{B.78}$$

Substituting all these expressions into equations (B.71) and (B.72) yields a system of partial differential equations (recall that the second derivatives of the tax function appear in all the ρ-terms.) Rather than working with that system, we hypothesize that, under Assumption 2, there is a solution that is piecewise linear. Under this hypothesis all the ρ-terms are zero with the implications that

$$\begin{aligned} y^{d0}_{2,\tau_1}(\cdot) \quad &= \quad y^{d0}_{1,\tau_2}(\cdot) \quad = \quad 0 \\ y^{d0}_{1,\tau_1}(\cdot) \quad &= \quad -y_1 \, \frac{\varepsilon_1}{1 - T^0_{y_1}(y)} \end{aligned}$$

and

$$y^{d0}_{2,\tau_2}(\cdot) \quad = \quad -y_2 \, \frac{\varepsilon_2}{1 - T^0_{y_2}(y)} \, .$$

Therefore,

$$\beta^{d0}_{I,1}(\omega) = -\frac{T^0_{y_1}(y)}{1 - T^0_{y_1}(y)} \, y_1 \, \varepsilon_1 \, , \tag{B.79}$$

and

$$\beta^{d0}_{I,2}(\omega) = -\frac{T^0_{y_2}(y)}{1 - T^0_{y_2}(y)} \, y_2 \, \varepsilon_2 \, . \tag{B.80}$$

Substituting (B.79) and (B.80) into (B.71) and (B.72) yields

$$\frac{T^0_{y_1}(y)}{1 - T^0_{y_1}(y)} \quad = \quad \Big(1 - \lambda^*_1 \, \bar{g}_1 \Big) \, \frac{1}{\bar{\alpha}_1} \, \frac{1}{\varepsilon_1} \, , \tag{B.81}$$

and

$$\frac{T^0_{y_2}(y)}{1 - T^0_{y_2}(y)} \quad = \quad \Big(1 - \lambda^2_1 \, \bar{g}_2 \Big) \, \frac{1}{\bar{\alpha}_2} \, \frac{1}{\varepsilon_2} \, . \tag{B.82}$$

and hence

$$T^0_{y_1}(y) = t^*_1 \quad \text{and} \quad T^0_{y_2}(y) = t^*_2 \, . \tag{B.83}$$

To complete the proof, we finally note that if we choose for all $y \in Y'$, $T^0_{y_1}(y) = t^*_1$ and $T^0_{y_1}(y) = t^*_1$, then this implies that the tax system is indeed piecewise linear: When

marginal tax rates are constant, second derivatives are zero. Moreover, equations (B.71) and (B.72) hold. Under Assumption 2 this implies that $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$, for all $y \in Y'$, as had to be shown.

B.1.10 Alternative Modeling Choices

Alternative Assumptions on the Preferences of Married Individuals. The couples' preferences specified above are such that the couples' disposable income is a public good for the spouses in a couple. Moreover, the spouses' preferences are quasi-linear in disposable income. Furthermore, in the couples' objective function, each spouse's effort costs receive the same weight. Theorems 1 - 3 below could be proven under much weaker assumptions: There could be a separate utility function for each spouse that, moreover, is not quasi-linear in disposable income. The disposable income could be treated as resource that is split between the couples – as opposed to being a public good. Who receives how much of that private good, and also who contributes how much to the couple's pre-tax income could be determined via Nash bargaining. Finally, the Nash bargaining might also be over the allocation of household duties, see Bierbrauer et al., 2023b for a formal treatment under these alternative assumptions. More specific assumptions are needed, however, in the context of an empirical application that uses sufficient statistics for the purpose of checking whether the conditions in the Theorems hold in the data. To ease the exposition, we base the whole analysis in this paper on the most simple setup which is prominent also in the related literature. In the empirical application, we will, moreover, invoke the assumption that the variable costs of productive effort are iso-elastic so that, for $i = 1, 2$,

$$k_i(y_i, \omega_i) = \frac{1}{1 + \frac{1}{\varepsilon_i}} \left(\frac{y_i}{\omega_i}\right)^{1+\frac{1}{\varepsilon_i}} ,$$

where ε_i can be interpreted as the Frisch elasticity of spouse i's labor supply.

Nash Bargaining. The formalism above nests the possibility that couples engage in Nash bargaining and, moreover, that there is heterogeneity in the bargaining weights of the spouses with index 1 and the spouses with index 2. To see this, suppose that spouse i has a utility function

$$u(c_m, y_i, \theta_i) \quad = \quad c_m - k_i(y_i, \omega_i) - \varphi_i \, \mathbf{1}(y_i > 0) \, .$$

and that the outcome of Nash bargaining is as if the couples were maximizing

$$\gamma_1\, u(c_m, y_1, \theta_1) + \gamma_2\, u(c_m, y_2, \theta_2) =$$
$$c_m - \gamma_1\, k_1(y_1, \omega_1) - \gamma_1\, \varphi_2\, \mathbf{1}(y_1 > 0) - \gamma_2\, k_2(y_2, \omega_2) - \gamma_2\, \varphi_2\, \mathbf{1}(y_2 > 0)\ ,$$

where γ_1 and $\gamma_2 = 1 - \gamma_2$ are, respectively, the bargaining weights of spouse 1 and spouse 2. With iso-elastic cost functions the resulting behavior is as if $\gamma_1 = \gamma_2 = 1$, while, the productive ability and fixed costs of spouse i are respectively equal to $\gamma_i^{1+\frac{1}{\varepsilon_i}}\omega_i$ and $\gamma_i\,\varphi_i$. Thus, working with an arbitrary distribution of θ, F_θ encompasses the possibility that couples differ in three dimensions: bargaining weights, productive abilities and fixed costs.

Singles and Couples. We look at married couples in isolation. In particular, we do not consider tax reforms that jointly alter the tax treatment of singles and the tax treatment of couples, see Bierbrauer et al., 2023b for an extensive discussion of this possibility. For the analysis of Pareto-improving reform directions this is without loss of generality: A Pareto-improving reform of an overall tax system – consisting of a tax function that applies to singles and one that applies to married couples – is possible if and only if one of the following conditions is met: (i) there is a Pareto-improving reform of the tax function for singles, while holding the tax function for couples fixed, and (ii) there is a Pareto-improving reform of the tax function for couples, while holding the tax function for singles fixed. Welfare-improving reforms of an overall system involve tradeoffs between singles and married individuals; e.g. between single mothers and married women. In line with the literature on optimal welfare-maximizing taxes, we focus on welfare-improvements amongst couples, while holding the tax function for singles fixed.[9]

[9]There is the possibility that both the tax function for singles and the tax function for couples are welfare-optimal on a stand-alone basis, while there are still welfare gains, say, from extracting more tax revenues from couples so as to raise the intercept of the single's consumption function.

B.2 APPENDIX FOR THE EMPIRICAL PART

This section of the appendix contains additional explanations, supplementary graphs and tables on the empirical part of the paper in Section 2.3. We first present details on the data preparation (Section B.2.1). Subsequently, supplementary insights are presented separately for each of the respective subset of results, i.e., for results on empirical conditional revenue functions (Section B.2.2), welfare effects of jointness-reducing reforms (Section B.2.3), tax reforms at the bottom of the income distribution (Section B.2.4), and optimal top tax rates (Section B.2.5).

B.2.1 Data

In Section 2.3 of the main text, we use household micro data from the Current Population Survey (CPS) and tabulated income tax return data from the Statistics of Income (SOI) program. CPS data is used for the estimation of conditional revenue functions (Section 2.3.1), and for analyzing the welfare effects of reducing jointness (Section 2.3.3), as well as tax reforms at the bottom of the income distribution (Section 2.3.4). For the calibration of optimal top tax rates in Section 2.3.5, we use tabulated income tax return data. In the following, we describe the details of the two data sources and respective data preparation.

Current Population Survey (CPS)

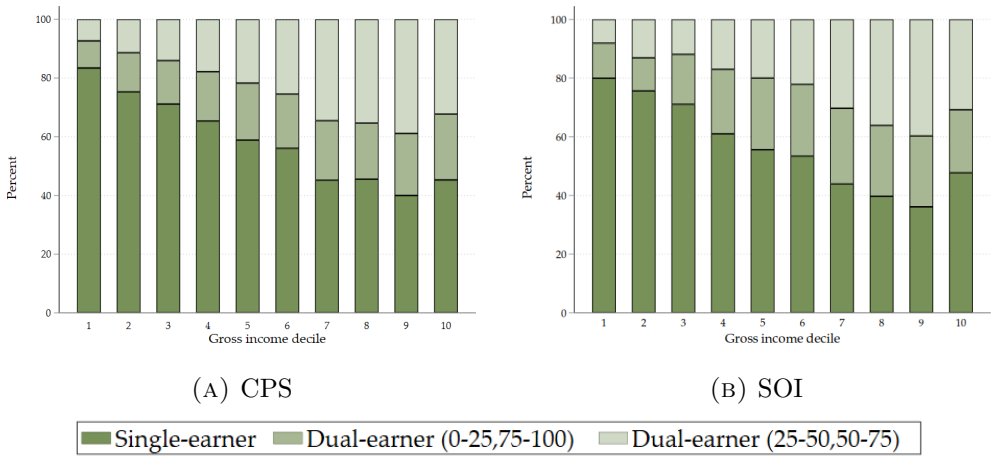
The Current Population Survey (CPS) is conducted by the US Census Bureau and the Bureau of Labor Statistics and contains nationally representative cross-sectional survey data from 1962 onward. We use data from the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC).[10] The sample size of CPS-ASEC increased from around 30,000 households in 1962 to more than 90,000 in the most recent wave. In contrast to tax return micro data such as the public use files (IRS-SOI PUF) from the Statistics of Income (SOI) division of the Internal Revenue Service (IRS), as, e.g., used by Bargain et al. (2015) or Bierbrauer et al. (2021), the CPS data contain exact information about the incomes of primary and secondary earners of the tax unit.[11]

[10] See Flood et al. (2021) and https://cps.ipums.org for a detailed description of CPS data.

[11] In the IRS-SOI PUF, the relevant information on salaries and wages from the W2-form of the primary and secondary earner is only available for the year 1974 and imputed for all other years using an undocumented procedure. For 1974, in which reliable information is available, the distribution of different couple types across per capita income distribution is very similar to the CPS data (see Figure B.4). Moreover, Bargain et al. (2015) compare inequality measures as well as the direct effects of tax policies based on CPS and SOI-PUF data and show that results are very similar (except for the very top of the distribution).

To adapt the CPS to the input requirements of the microsimulation model, we transform the CPS from a household-level data set to a tax unit level data set. For this purpose, we form tax units by joining all married spouses with their dependent children. Single individuals and unmarried spouses form separate tax units. Children of single individuals are in most cases allocated to the household head. Adult individuals with a total income below the year-specific personal exemption threshold are assumed to reflect dependents of the household head. Table B.1 illustrates in detail the correspondence between variables utilized in NBER TAXSIM and variables in the CPS data.

Figure B.4. – Comparison of CPS and SOI Data (1974), Couple Types



(A) CPS (B) SOI

Note: This figure displays for the tax year of 1974 the distribution of married couple types across deciles of the per capita income distribution. The figure compares the distribution based on the CPS data (Figure B.4a) to the distribution based on the IRS-SOI PUF tax return micro data (Figure B.4b). All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old.
Source: Our calculations based on CPS-ASEC and SOI PUF.

Treatment of Top Incomes. In the CPS data, information on top incomes is limited by (i) public topcoding, and (ii) internal censoring. We address both limitations by harmonizing the treatment of top incomes across the different survey years and by following Piketty and Saez (2003) and Piketty (2003) in assuming that top incomes are well represented by a Pareto distribution.

In a first step, we address the challenge that public topcoding methods vary over time. In most recent years (since 2011), the Census Bureau uses a rank proximity swapping procedure to preserve the privacy for top income earners while maintaining the internal distribution of top incomes. In this procedure, values at or above a specific swap thresh-

old are switched against other top income values within a bounded interval. For previous years, however, the CPS data originally contains top income values that are based on different procedures, in particular traditional topcoding (1962-1995), and a replacement value system procedure (1996-2010). To be able to consistently analyze the data, we apply the most recent method of rank proximity swapping also to previous years using supplementary files provided by IPUMS.[12] Thereby, we preserve the internally used distribution of top incomes whenever possible.

In a second step, we address the challenge that top incomes are also internally censored based on the value range limits of the income variables. As shown by Larrimore et al. (2008), since these censoring thresholds have changed discretely at specific points in time, the share of individuals affected by censoring varies and can reach up to one percent in specific years. To address the unequal representation of censored incomes, we replace censored incomes by random draws from a Pareto distribution. In particular, we first identify for every year and every income type the highest possible income T assigned in a given year. Based on this censoring threshold, we generate for every year and every income type the parameter α of a Pareto distribution with density $f(Y) = \alpha * T^{\alpha} * Y^{-\alpha-1}$. We thereby assume that incomes above the 99th percentile follow a Pareto distribution and thus estimate the shape parameter α as

$$\alpha = \frac{ln\left(\frac{N_{Y \geq p99}}{N_{Y=T}}\right)}{ln\left(\frac{Y_T}{Y_{p99}}\right)}$$

where $N_{Y \geq p99}$ is the number of individuals with an income above the 99th percentile of the income distribution, $N_{Y=T}$ is the number of individuals at the highest income, and Y_T and Y_{p99} are the top income and the income at the 99th percentile respectively.[13] Finally, we use the distribution to replace the top incomes T by random draws from this calibrated distribution.[14]

[12]For details on the treatment of top incomes in general and the data used for rank proximity swapping, see https://cps.ipums.org/cps/topcodes_tables.shtml and https://cps.ipums.org/cps/income_cell_means.shtml.

[13]Discussions of different estimation methods for the shape parameter of the Pareto distribution can be found in Armour et al. (2016) and Blanchet et al. (2022).

[14]To reduce the impact of random sampling on our results, we use quantiles of the distribution. The number of quantiles utilized depends on the number of individuals at the top income. For instance, if we observe 25 individuals at the top income, we assign these individuals income levels that correspond to the 25 quantiles of the randomly drawn values from the calibrated Pareto distribution. Thereby, we preserve the properties of the distribution while limiting the influence of random draws.

Sample Restrictions. In this paper, we focus on married couples.[15] We assume that married couples always file jointly. While married couples can also file separately, this filing status is usually not beneficial (see Figure B.6) and is chosen by less than two percent of all tax units (see Figure B.5).[16] Similarly, we abstract from the qualifying widow(er) filing status that gives widowed individuals a preferential tax treatment in the two years following the spouses' death. Given our sample restriction, the occurrence of widow(er)s is negligible (see also Figure B.5). If not indicated otherwise, we restrict the sample to tax units in which primary and secondary taxpayer are between 25 and 55 years old and have non-negative gross income. This sample restriction is guided by (i) our model that considers neither education nor retirement decisions, and (ii) the assumptions on labor supply responses to taxation that are not valid for young and old people with weak labor force attachment.

Throughout the analysis, we calculate tax payments as well as average and marginal tax rates based on the federal income tax and abstract from state income tax and social security payroll taxes. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income.

IRS Income Tax Tabulations

For the estimation of Pareto coefficients, we use tabular data from the Statistics of Income (SOI) program.[17] The data contains detailed tabular information on joint return tax payers with wage income, separated by the size of wage income, gender, and the share of total wage income. It includes all primary and secondary taxpayers with wages, tips and other compensation as reported in the wage and tax statement of income tax return form W-2.

Data Preparation. We prepare the tables for 2008-2018 in two steps. First, since some cells have been merged by the IRS to avoid information disclosure for specific taxpayers, we need to impute some information. Such an imputation is not necessary to estimate Pareto coefficients for a particular group. However, to aggregate information across groups, e.g.,

[15] Our companion paper Bierbrauer et al. (2025) discusses the relationship between singles and married couples explicitly.

[16] Filing separately can be beneficial in very particular circumstances that we do not observe, i.e., in the case of substantial itemizable deductions (e.g. high medical expenses or student loan repayments).

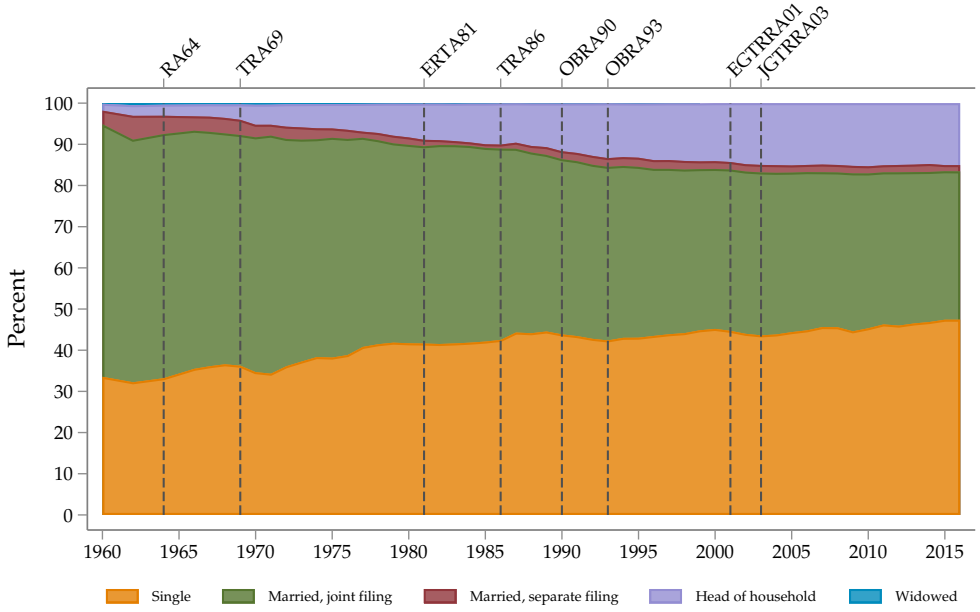
[17] See IRS, Statistics of Income Division, Form W-2 study, June 2019.

Table B.1. – TAXSIM Variables and CPS Application

TAXSIM	Explanation	CPS Application
taxsimid	Case ID	N/A
year	Tax year	ASEC income reference year
state	State	State of residence
mstat	Marital Status	Marital status (married vs. unmarried)
page	Age of primary taxpayer	Age of husband
sage	Age of spouse	Age of spouse
depx	Number of dependents	Number of children below and of age 18 + additional dependents
dep13	Number of children under 13	Number of children under 13
dep17	Number of children under 17	Number of children under 17
dep18	Number of qualifying children for EITC.	Number of children below and of age 18
pwages	Wage and salary income of Primary Taxpayer	Wage income + business income + farm income of husband
swages	Wage and salary income of Spouse	Wage income + business income + farm income of spouse
dividends	Dividend income	Income from dividends
intrec	Interest Received	Income from interest
stcg	Short Term Capital Gains or losses	N/A
ltcg	Long Term Capital Gains or losses.	Capital gains - capital losses
otherprop	Other property income	Income from rent
nonprop	Other non-property income	Income from other Source not specified + income from alimony
pensions	Taxable Pensions and IRA distributions	Retirement income
gssi	Gross Social Security Benefits	Social Security income
ui	Unemployment compensation received	Income from unemployment benefits
transfers	Other non-taxable transfer Income	Welfare (public assistance) income + income from worker's compensation + income from veteran benefits + income from survivor benefits + income from disability benefits + income from child support + income from educational assistance + income from SSI + income from assistance
rentpaid	Rent Paid	N/A
proptax	Real Estate taxes paid	Annual property taxes
otheritem	Other Itemized deductions	Indirect calculation via difference between adjusted gross income and taxable income calculated by the Census Bureau's tax model.
childcare	Child care expenses	N/A
mortgage	Deductions not included in otheritem	N/A

Note: This table displays the variables utilized as part of the tax calculation via the NBER TAXSIM (v32) microsimulation model and the corresponding information from the CPS used for the respective variables. For details on TAXSIM (v32) see Feenberg and Coutts (1993) and https://users.nber.org/~taxsim/.
Source: NBER TAXSIM and CPS-ASEC.

Figure B.5. – Filing Status According to SOI Data



Note: This figure shows the distribution of filing status from 1960 to 2016. Filing statuses are based on the IRS-SOI PUF administrative tax return micro data.
Source: Bierbrauer et al. (2025), calculations based on SOI PUF.

to obtain data on all secondary earners within a specific income range, the income ranges need to be harmonized. To fill in missing values for unmerged cells, we distribute the merged number of tax payers and wages according to the distribution of all tax payers of a specific gender. For instance, for the tax year 2018, among women with a wage share between 25 and 50 percent, there are 127 tax payers who earn between 1.5 million and 10 million (merged cell). To allocate these tax payers to the more detailed income brackets of 1.5-2 / 2-5 / 5-10 million (unmerged cells), we use information on how tax payers are distributed within the merged income cell among all women. This allocation procedure is applied for both the number of tax payers and the sum of wages, separately for men and women.

Second, we combine the cell information for men and women with different wage shares to data on primary and secondary earners, i.e. we combine the number of taxpayers and the sum of wages for men and women with wage shares between 0 and 50 percent (secondary earners) and for men and women with wage shares between 50 and 100 percent (primary earners). The final table contains separate primary and secondary earner data on the number of tax payers and the sum of wages for different income ranges. We further produce tables for single earner couples and equal earning dual earner couples, for which the primary (secondary) earner's income share lies between 50 (25) and 75 (50) percent.

Figure B.6. – Married Couples Filing Jointly and Separately (2019)

Note: This figure shows how the average tax rate of a couple with specific gross earnings differs between whether this couple files separately or jointly. In addition, the figure also shows the average tax rate of two singles with the same joint income. The figure differentiates further by the type of couple: single earner couples (95% / 5%), unequal dual earner couples (75% / 25%) and dual earner couples with equal incomes (50% / 50%).
Source: Bierbrauer et al. (2025), calculations based on NBER TAXSIM and CPS-ASEC.

In case of inconsistencies, for instance when the imputation yields an average income that is lower than the bracket threshold, we abstract from this data point for the Pareto interpolation.

Pareto Interpolation. We use the tabular information on the number of taxpayers and the sum of wage income in a particular wage income bracket to generate empirical quantiles associated with the lower threshold and average wage income of a specific wage income bracket. We then follow the generalized Pareto interpolation technique by Blanchet et al. (2022) to obtain information on specific percentiles of the distribution and the associated Pareto coefficient. In contrast to standard Pareto interpolation, generalized Pareto interpolation in principle allows the Pareto coefficient to vary across the distribution and thereby provide a more flexible way to describe distributions based on tabular data.

B.2.2 Conditional Revenue Functions

In Section 2.3.1 of the main text, we compute empirical conditional revenue functions for the analysis of Pareto-improving tax reforms in the domain of couples taxation. In this section of the appendix, we provide more detailed explanations on how empirical revenue

functions are constructed (Section B.2.2). We also show supplementary results on the empirical ingredients of these conditional revenue functions (Section B.2.2), robustness tests regarding behavioral responses at the extensive and intensive margin (Section B.2.2), and results for alternative choices of the conditioning brackets (Section B.2.2).

From Theory to Data

As discussed in the main text, to bring conditional revenue functions to the data, we condition on income ranges as described in equation (2.10). In the main text, we showed in equation (2.11) that the adjusted equation for secondary earners based on Proposition 1 reads

$$\mathcal{R}_2(y_{2s} \mid B_1) := s_1(B_1)\Big(I_2^d(y_{2s}, B_1) \; + \; M_2^{xd}(y_{2s}, B_1)\Big) \tag{B.84}$$

where $s_1(B_1) = F_{y_1}(\bar{b}_1) - F_{y_1}(\underline{b}_1)$ is the mass of couples with $y_1 \in B_1 = [\underline{b}_1, \bar{b}_1]$,

$$I_2^d(y_{2s}, B_1) \quad := \quad -\mathbb{E}_{y_1}\left[f_{y_2}(y_{2s} \mid y_1)\frac{T_{y_2}^0(y_1, y_{2s})}{1 - T_{y_2}^0(y_1, y_{2s})}\; y_{2s}\; \varepsilon_2 \mid y_1 \in B_1\right]$$

gives the revenue effect of behavioral responses at the intensive margin, where

$$M_2^{xd}(y_{2s}, B_1) \quad := \quad \mathbb{E}_{y_1}\Big[(1 - F_{y_2}(y_{2s} \mid y_1))(1 - \mathcal{E}_x^d(y_{2s} \mid y_1)) \mid y_1 \in B_1\Big] \; .$$

For the empirical application, we assume that the conditional distribution of secondary earnings does not vary within the specific bracket of primary earnings that we are looking at. This assumption is necessary, since we require enough data points to estimate the secondary earner income distribution. We also assume that the extensive margin responses do not vary across the conditioning earnings variable. Consequently, the empirically estimated revenue function simplifies to

$$\mathcal{R}\left(y_{2s} \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q\right) = \underbrace{s_1^Q}_{\substack{\text{Share Inc.} \\ \text{Range}}} \Bigg(\underbrace{-\bar{\beta}_{I,2}^d\left(y_{1s}, y_{2s}\right) f^{y_2}(y_{2s} \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q)}_{\text{Intensive Margin}} + \underbrace{\left(1 - F^{y_2}\left(y_{2s} \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q\right)\right)}_{\text{Sec. Earnings Cond. CDF}} \underbrace{\left(1 - \mathcal{E}_x^d\right)}_{\text{Extensive Margin}} \Bigg) \tag{B.85}$$

$$\bar{\beta}_{I,2}^d = \mathbb{E}_{y_{1s}}\left[\frac{T^{0\prime}\left(y^{d0}\right)}{1 - T^{0\prime}\left(y^{d0}\right)} y_{2s}^{d0} \varepsilon_2 \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q\right]$$

where $\underline{y}_1^Q$ and $\overline{y}_1^Q$ indicate lower and upper thresholds for the respective primary earner income range Q, upon which the revenue functions are conditioned.

The equation provides an intuitive understanding of the mechanics behind the conditional revenue function. The term s_1^Q reflects the share of dual earner couples whose primary earnings fall in the respective income range, and are thus affected by a tax reform. The first part of the equation captures the change in tax revenue through behavioral responses at the intensive margin. The intensive margin behavioral response affects all secondary earners at the income level y_{2s} with primary earnings in the respective income range. $\bar{\beta}_{I,2}^d = \mathbb{E}_{y_{1s}}\left[\frac{T^{0\prime}\left(y^{d0}\right)}{1-T^{0\prime}\left(y^{d0}\right)} y_{2s}^{d0} \varepsilon_2 \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q\right]$ describes the reaction of these secondary earners, and is based on an average at every secondary earnings level y_{2s}, because marginal tax rates faced by secondary earners, as well as elasticities, might vary across primary earnings in the bracket. The second part of the equation captures the mechanical change in tax revenue. This mechanical change in tax revenue is downscaled by the behavioral response at the extensive margin. As a whole, the value and shape of the empirical conditional revenue function characterize the potential for Pareto improving tax reforms for secondary earners with a spouse in specific income ranges (see Theorem 1).

We also compare *conditional* revenue functions to the respective *unconditional* revenue function. The unconditional revenue function illustrates the potential for Pareto improvements through tax rate changes that are independent from the earnings of the other spouse. For secondary earners in dual earner couples, this unconditional revenue function reads

$$\mathcal{R}\left(y_{2s}\right) = \underbrace{-\mathbb{E}_{y_{1s}}\left[\frac{T^{0\prime}\left(y^{d0}\right)}{1 - T^{0\prime}\left(y^{d0}\right)} y_{2s}^{d0} \varepsilon_2\right] f^{y_2}(y_{2s})}_{\text{Intensive Margin}} + \underbrace{\left(1 - F^{y_2}\left(y_{2s}\right)\right)}_{\text{Mechanical Effect}} \underbrace{\left(1 - \mathcal{E}_x^d\right)}_{\text{Extensive Margin}} . \tag{B.86}$$

A comparison between conditional and unconditional revenue functions reveals whether inefficiencies in the tax code can be cured by changing marginal tax rates for *all* secondary earners or whether curing inefficiencies requires the joint consideration of secondary and primary earnings.

Ingredients

The estimation of revenue functions requires assumptions about labor supply elasticities and conditioning brackets for the reform. In addition, we need estimates on the (conditional) income distributions and intensive margin response terms.

Table B.2 display three different scenarios that we use throughout the paper and that are informed by the literature listed in the table note. Table B.3 presents the upper thresholds of the income deciles for the primary and secondary earner income distributions that are used as conditioning brackets.

We estimate all income distributions by means of a kernel density estimation using a Gaussian kernel and 20,000 grid points. The income grid ranges from zero to one million dollars.

The intensive margin response terms in equation (B.85) and (B.86) demand knowledge of a conditional average. For instance, the secondary earner conditional revenue function contains an average of the intensive margin response at every secondary earnings income level conditional on primary earnings in a specific decile. We estimate these averages by means of fitting a local polynomial regression of degree zero (local mean smoothing) with a Gaussian kernel as a weighting function.

Figures B.7 - B.16 show the empirical ingredients to compute the conditional revenue functions for secondary earners, in particular marginal tax rates, conditional income distributions, and behavioral responses to taxation. Figures B.17-B.26 show the respective ingredients for the conditional revenue functions for primary earners.

Robustness

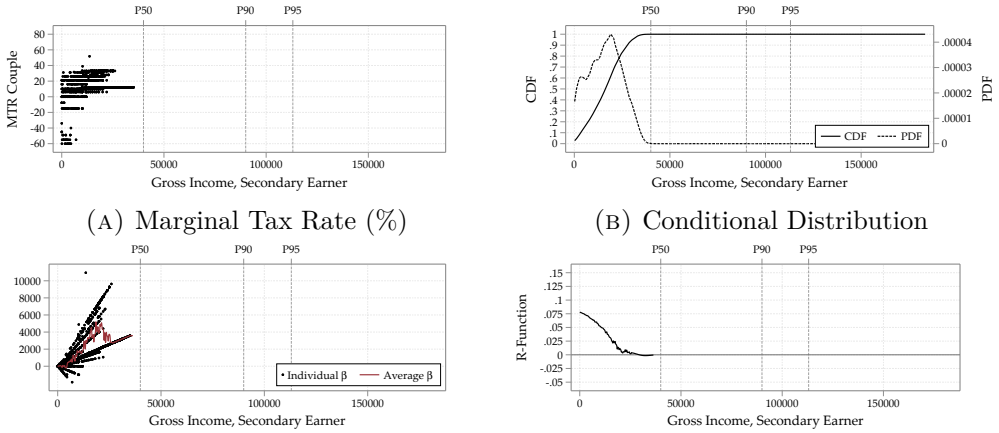
In the baseline, conditional revenue functions are computed under the assumption of extensive margin and moderate intensive margin responses. We provide robustness tests regarding the exclusion of margin responses (Figure B.27) and for the use of low (high) elasticity scenarios from Table B.2 in Figure B.28 (B.29).

Table B.2. – Assumptions about Labor Supply Elasticities

	Primary Earner	Secondary Earner
Low Elasticity Scenario	0.15	0.35
Baseline Elasticity Scenario	0.25	0.75
High Elasticity Scenario	0.5	1.5

Note: This table displays our assumptions about the labor supply elasticities for primary and secondary earners in married couples. Assumptions are guided by the range of estimates found in the literature, e.g. Blundell and Macurdy (1999), Blau and Kahn (2007), Eissa and Hoynes (2004), LaLumia (2008), Saez et al. (2012), Bargain et al. (2014), and Neisser (2021).

Figure B.7. – Cond. Revenue Functions, Secondary Earners, PE Q1



(A) Marginal Tax Rate (%)

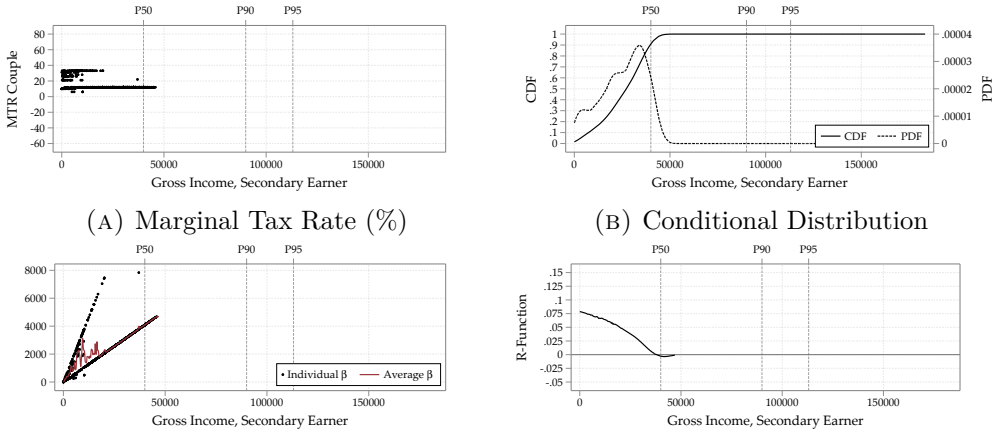
(B) Conditional Distribution

(C) β

(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q1. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.8. – Cond. Revenue Functions, Secondary Earners, PE Q2



(A) Marginal Tax Rate (%)

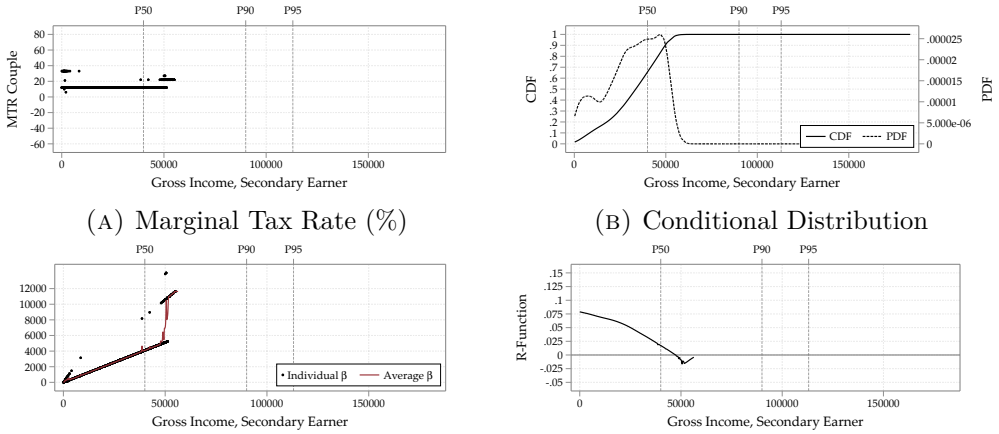
(B) Conditional Distribution

(C) β

(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q2. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B.9. – Cond. Revenue Functions, Secondary Earners, PE Q3



(A) Marginal Tax Rate (%)

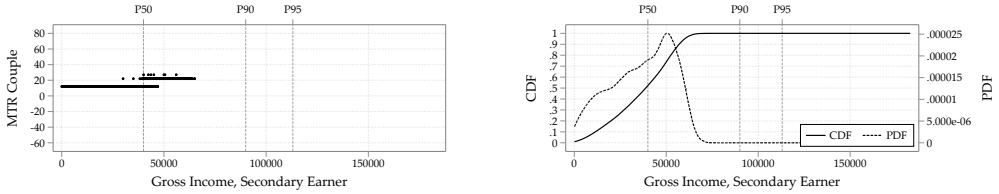
(B) Conditional Distribution

(C) β

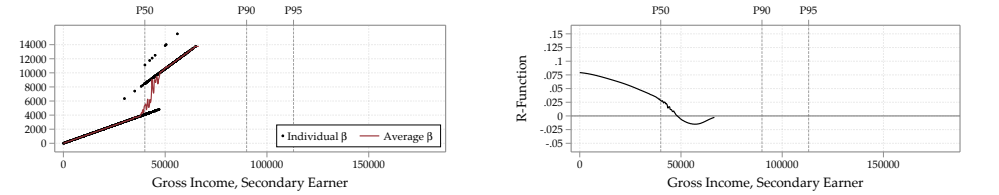
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q3. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.10. – Cond. Revenue Functions, Secondary Earners, PE Q4



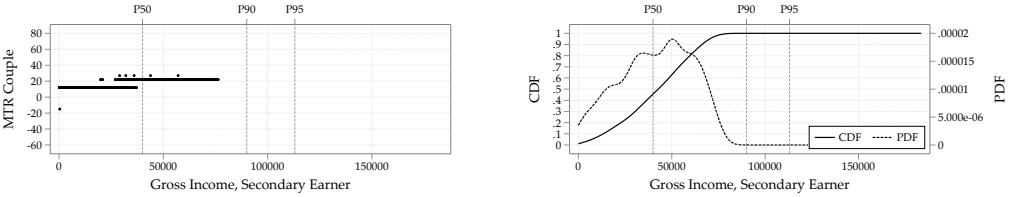
(A) Marginal Tax Rate (%) (B) Conditional Distribution



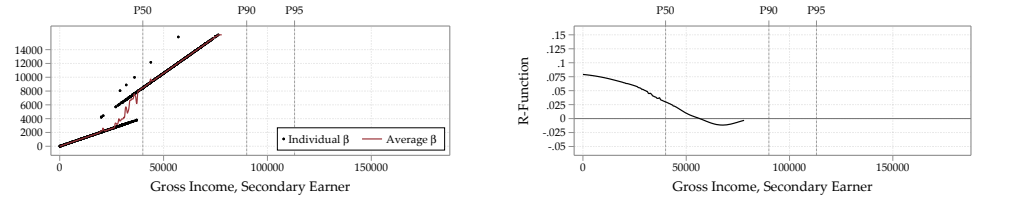
(C) β (D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q4. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B.11. – Cond. Revenue Functions, Secondary Earners, PE Q5



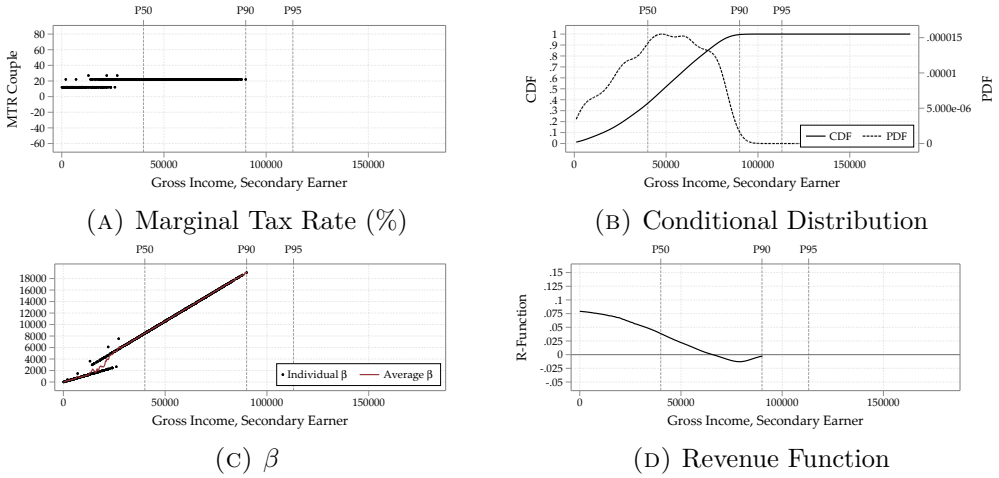
(A) Marginal Tax Rate (%) (B) Conditional Distribution



(C) β (D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q5. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.12. – Cond. Revenue Functions, Secondary Earners, PE Q6



(A) Marginal Tax Rate (%)

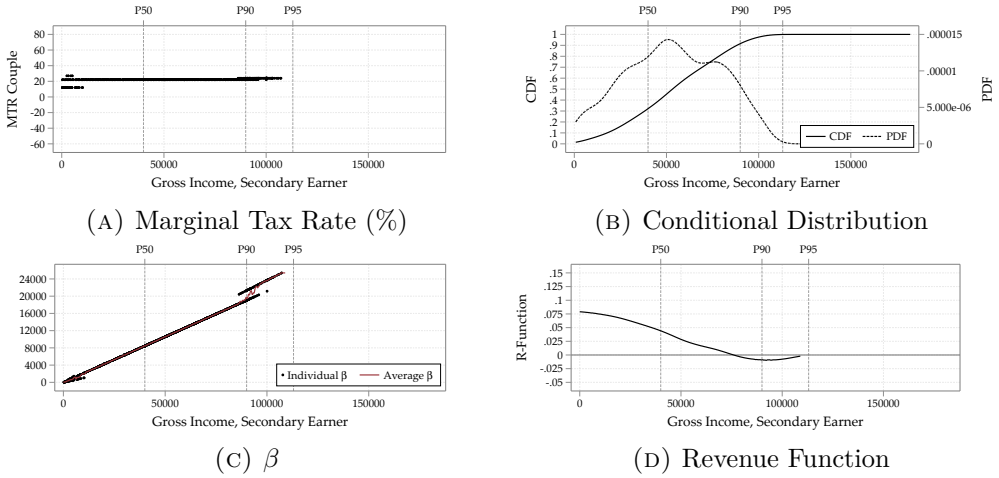
(B) Conditional Distribution

(C) β

(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q6. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B.13. – Cond. Revenue Functions, Secondary Earners, PE Q7



(A) Marginal Tax Rate (%)

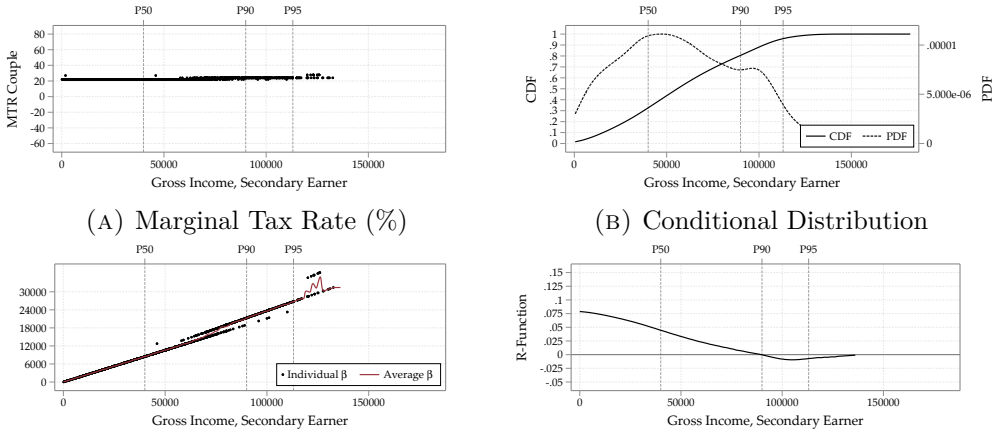
(B) Conditional Distribution

(C) β

(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q7. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.14. – Cond. Revenue Functions, Secondary Earners, PE Q8



(A) Marginal Tax Rate (%)

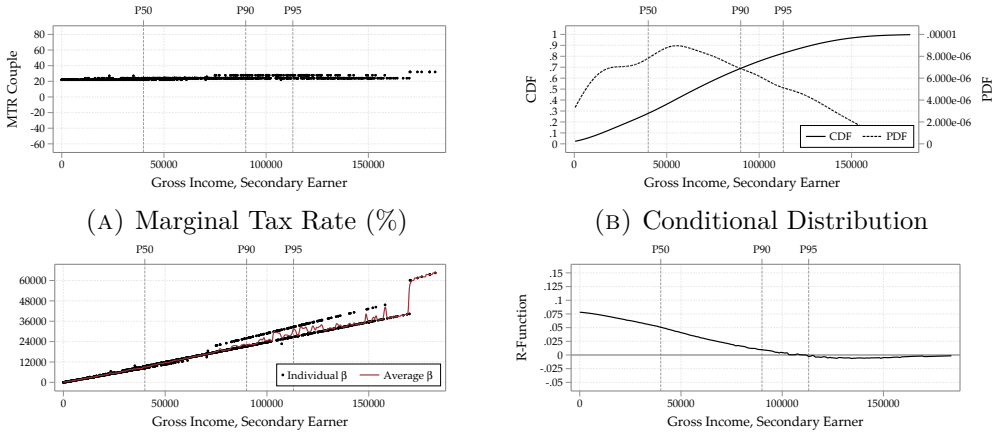
(B) Conditional Distribution

(C) β

(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q8. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.15. – Cond. Revenue Functions, Secondary Earners, PE Q9



(A) Marginal Tax Rate (%)

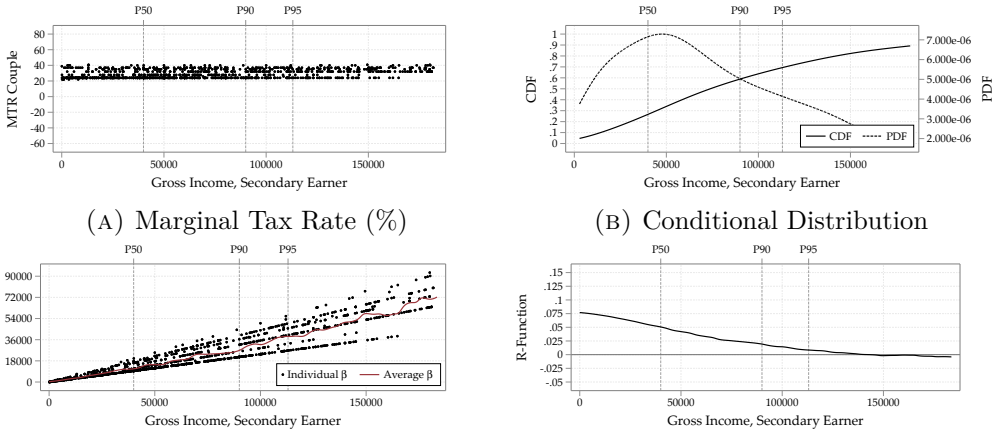
(B) Conditional Distribution

(C) β

(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q9. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.16. – Cond. Revenue Functions, Secondary Earners, PE Q10



(A) Marginal Tax Rate (%)

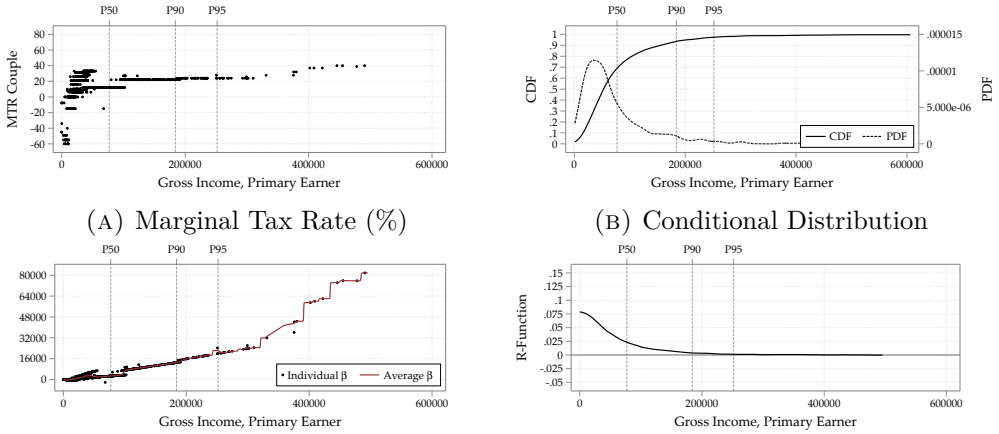
(B) Conditional Distribution

(C) β

(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q10. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.17. – Cond. Revenue Functions, Primary Earners, SE Q1



(A) Marginal Tax Rate (%)

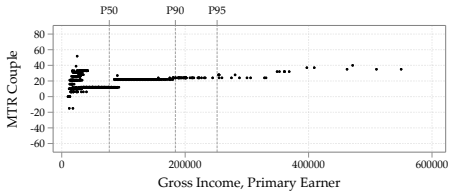
(B) Conditional Distribution

(C) β

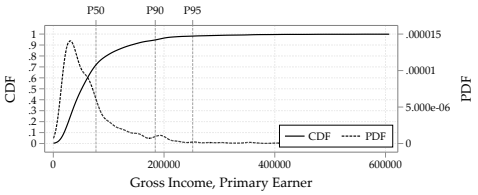
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q1. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

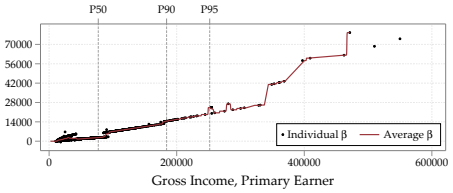
Figure B.18. – Cond. Revenue Functions, Primary Earners, SE Q2



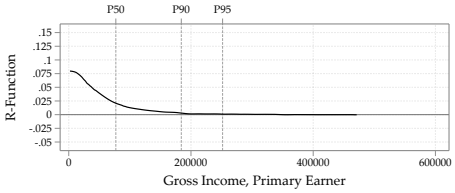
(A) Marginal Tax Rate (%)



(B) Conditional Distribution



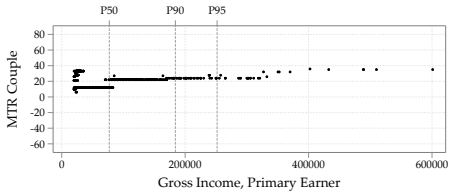
(C) β



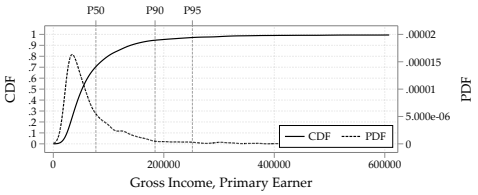
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q2. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

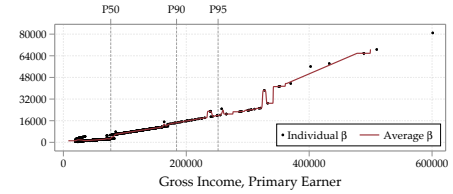
Figure B.19. – Cond. Revenue Functions, Primary Earners, SE Q3



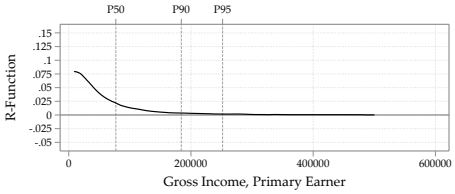
(A) Marginal Tax Rate (%)



(B) Conditional Distribution



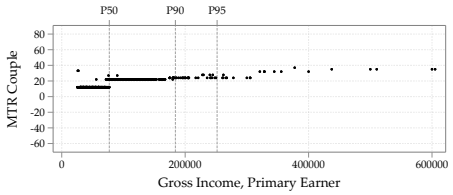
(C) β



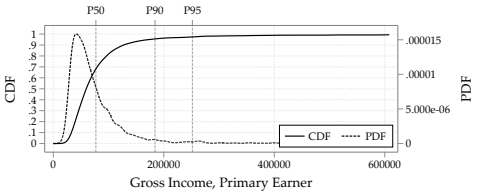
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q3. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

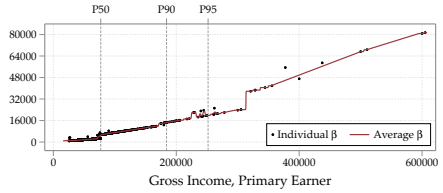
Figure B.20. – Cond. Revenue Functions, Primary Earners, SE Q4



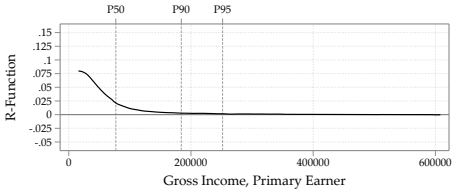
(A) Marginal Tax Rate (%)



(B) Conditional Distribution



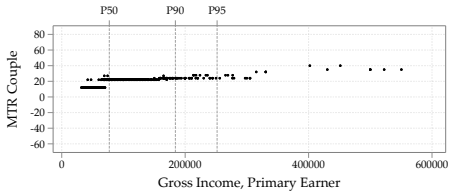
(C) β



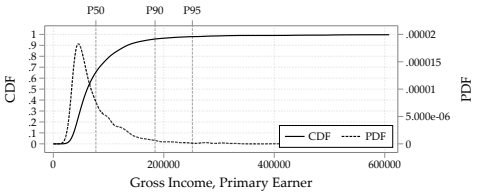
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q4. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

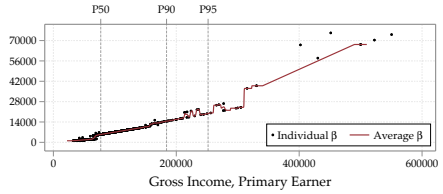
Figure B.21. – Cond. Revenue Functions, Primary Earners, SE Q5



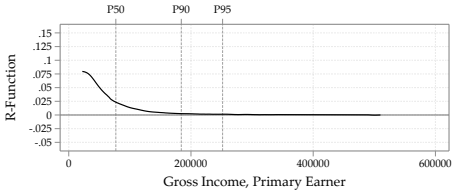
(A) Marginal Tax Rate (%)



(B) Conditional Distribution



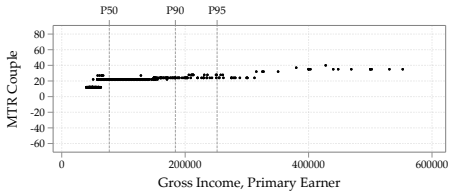
(C) β



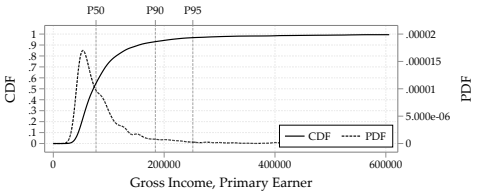
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q5. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

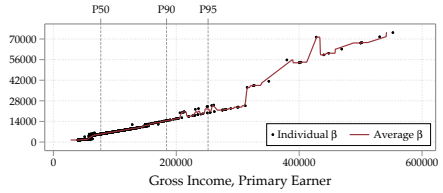
Figure B.22. – Cond. Revenue Functions, Primary Earners, SE Q6



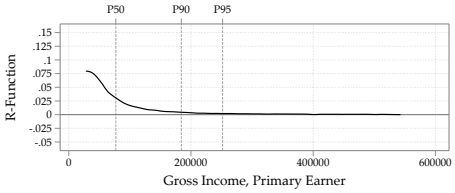
(A) Marginal Tax Rate (%)



(B) Conditional Distribution



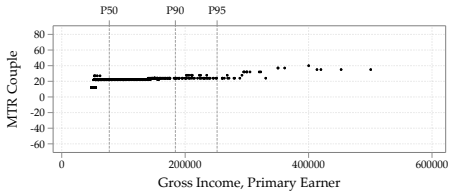
(C) β



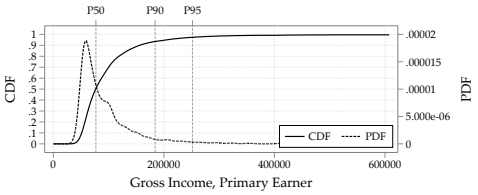
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q6. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

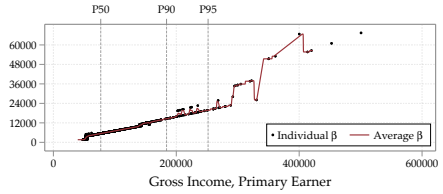
Figure B.23. – Cond. Revenue Functions, Primary Earners, SE Q7



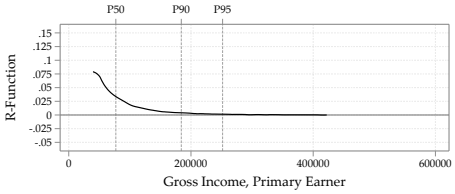
(A) Marginal Tax Rate (%)



(B) Conditional Distribution



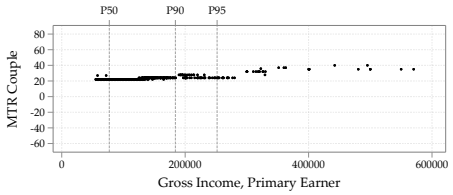
(C) β



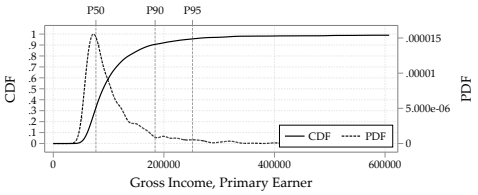
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q7. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

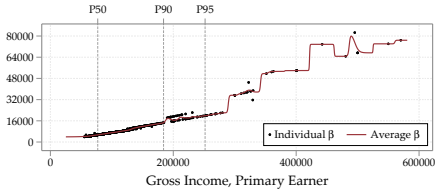
Figure B.24. – Cond. Revenue Functions, Primary Earners, SE Q8



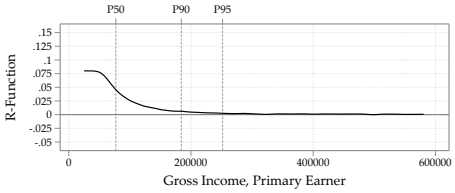
(A) Marginal Tax Rate (%)



(B) Conditional Distribution



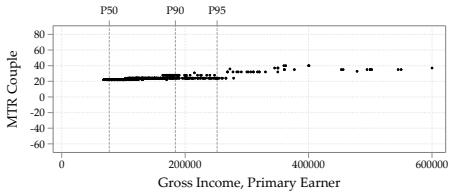
(C) β



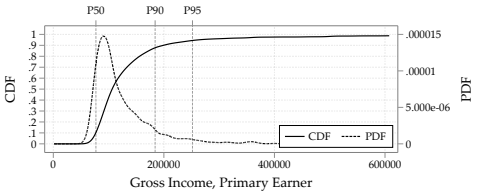
(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q8. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

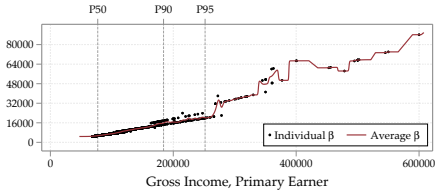
Figure B.25. – Cond. Revenue Functions, Primary Earners, SE Q9



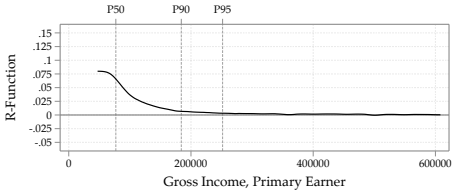
(A) Marginal Tax Rate (%)



(B) Conditional Distribution



(C) β



(D) Revenue Function

Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q9. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Table B.3. – Income Deciles (2019)

Decile	Primary Earnings	Secondary Earnings
Q1	35000	10001
Q2	46002	19000
Q3	55604	25015
Q4	65228	32005
Q5	77056	40000
Q6	91032	47300
Q7	108966	55009
Q8	135030	67785
Q9	184155	90001
Q10	18009756	1358632

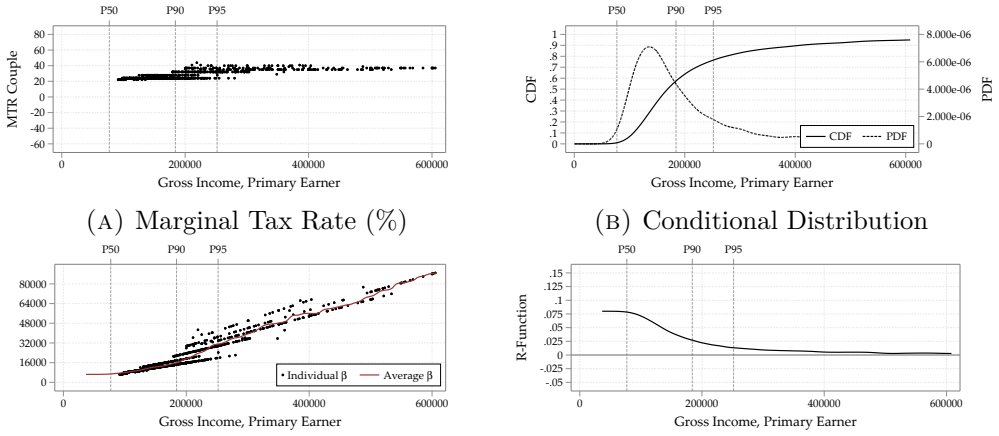
Note: This table shows the upper thresholds of the deciles for the primary and secondary earner income distribution among dual earner couples as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income.
Source: Own calculations based on CPS-ASEC (2019)

Alternative Conditioning Brackets

In the main text, we construct conditioning brackets by looking at deciles of the primary and secondary earner income distribution. Our framework, however, is flexible enough to capture any construction of conditioning brackets. Empirically, we are limited by data constraints, since we need enough observations to estimate conditional distributions. If we condition on wider brackets, we receive more precise results; however we potentially also average out more inefficiencies. If we condition on narrower brackets, we require less averaging, but also have less observations to estimate distributions. In this section, we provide results for wider brackets (quintiles in Figure B.30), and more narrow brackets (vingtiles in Figure B.31).

An alternative way of conditioning on primary or secondary earnings is to select brackets based on the statutory income tax schedule. Table B.4 displays the corresponding income thresholds and the share of primary and secondary earners with incomes in the respective tax bracket. Figure B.32 displays the estimated conditional revenue functions for secondary (primary) earners conditional on primary (secondary) earnings being in one of the seven income brackets in 2019. In contrast to the conditioning based on deciles, revenue functions for different tax brackets do not have a common intercept with the vertical axis, because the share of secondary earners in a particular income range is no longer the same. However, conditional revenue functions again show inefficiently high marginal tax rates for some "middle tax bracket" income levels (see Figure B.32a).

Figure B.26. – Cond. Revenue Functions, Primary Earners, SE Q10



(A) Marginal Tax Rate (%)

(B) Conditional Distribution

(C) β

(D) Revenue Function

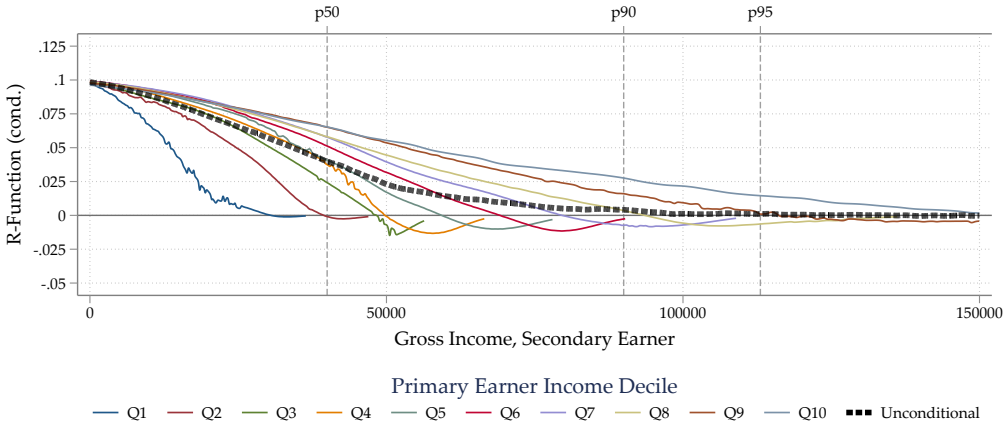
Note: This figure shows the ingredients necessary to estimate equation (B.85) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q10. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Table B.4. – Taxable Income Brackets (2019)

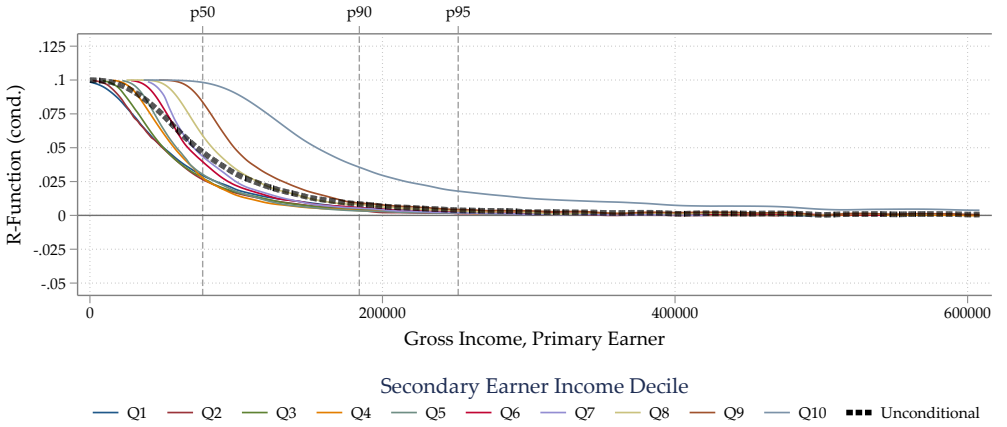
Bracket	Upper Threshold	MTR	Share Prim. Earner	Share Sec. Earner
1	19400	10	1.83%	20.41%
2	78950	12	48.89%	65.41%
3	168400	22	36.83%	12.75%
4	321450	24	9.46%	1.29%
5	408200	32	1.01%	0.02%
6	612350	35	1.01%	0.05%
7	.	37	0.98%	0.07%

Note: This table shows the upper thresholds (column 2) and marginal tax rates (column 3) for the statutory tax schedule for married couples filing jointly as of 2019. All threshold values refer to taxable income. Column 5 (6) displays the share of primary (secondary) earners whose primary (secondary) earnings lie in a respective bracket.
Source: Own calculations based on CPS-ASEC (2019) and U.S. Federal Individual Income Tax Rates and Brackets (2019)

Figure B.27. – Cond. Revenue Functions, Deciles (2019), No Extensive Margin



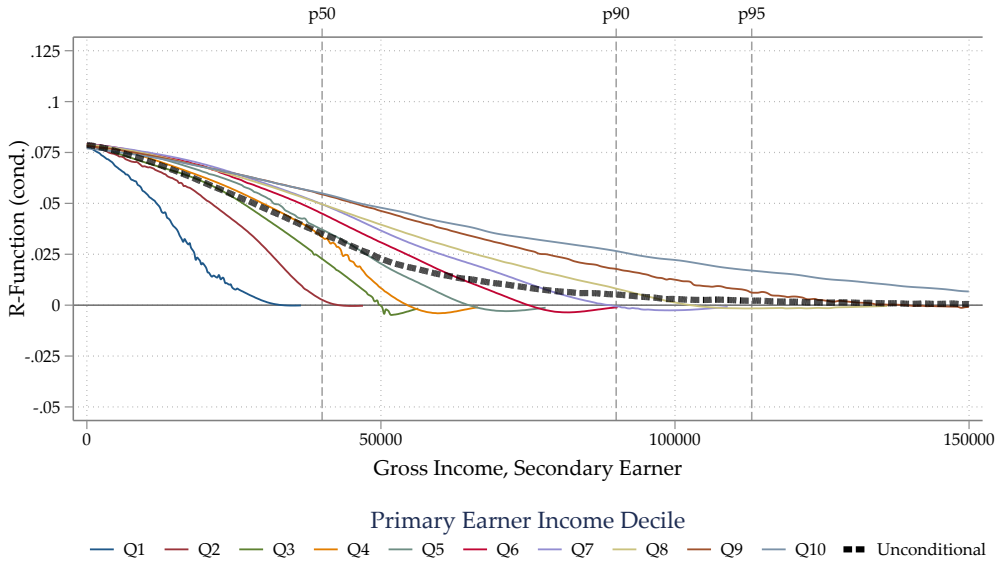
(A) Secondary Earners



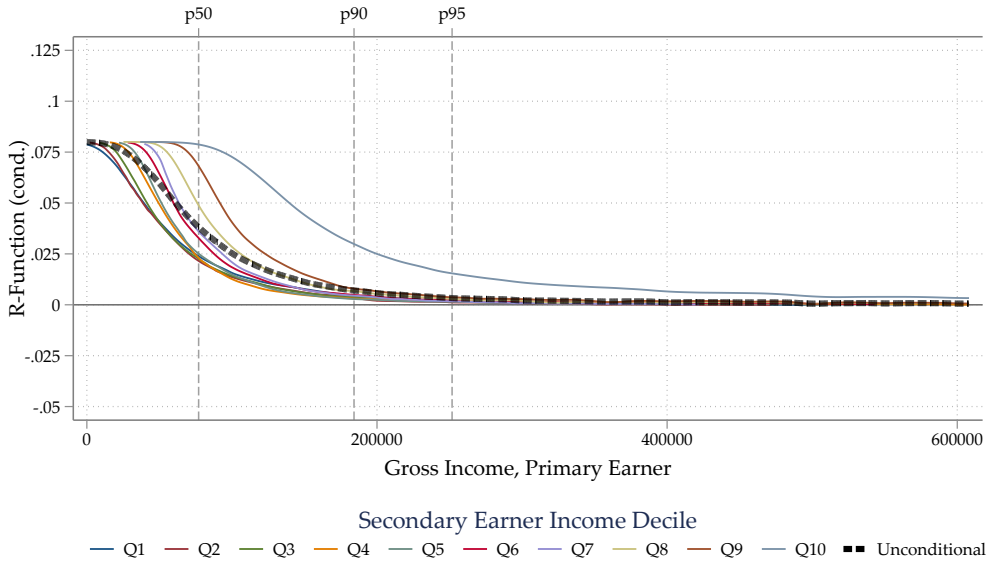
(B) Primary Earners

Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive margin into account. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners. The figure also displays modified unconditional revenue functions for secondary and primary earners where unconditional revenue functions have been scaled by 0.1 to facilitate comparability with the conditional revenue functions.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.28. – Cond. Revenue Functions, Deciles (2019), Low Elasticity



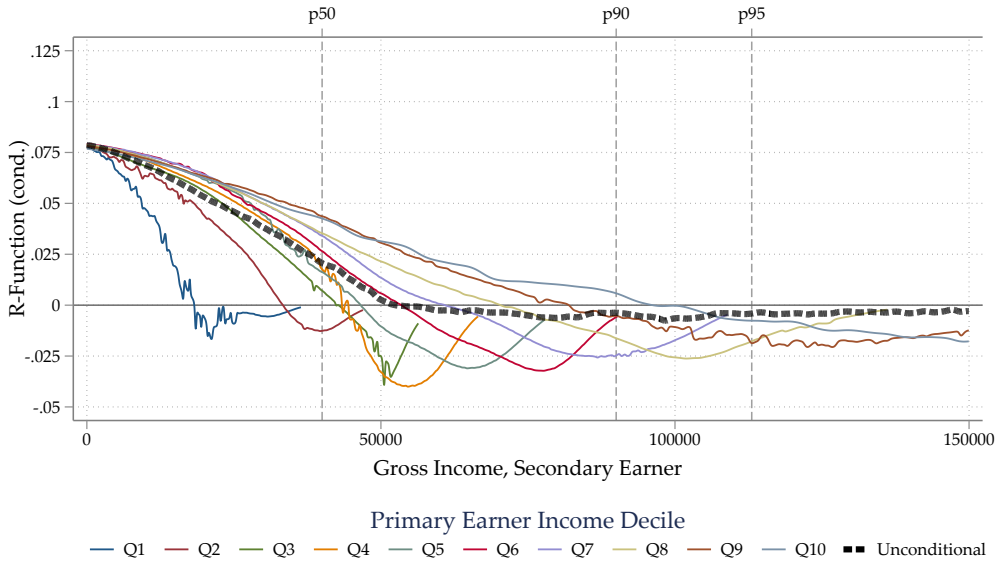
(A) Secondary Earners



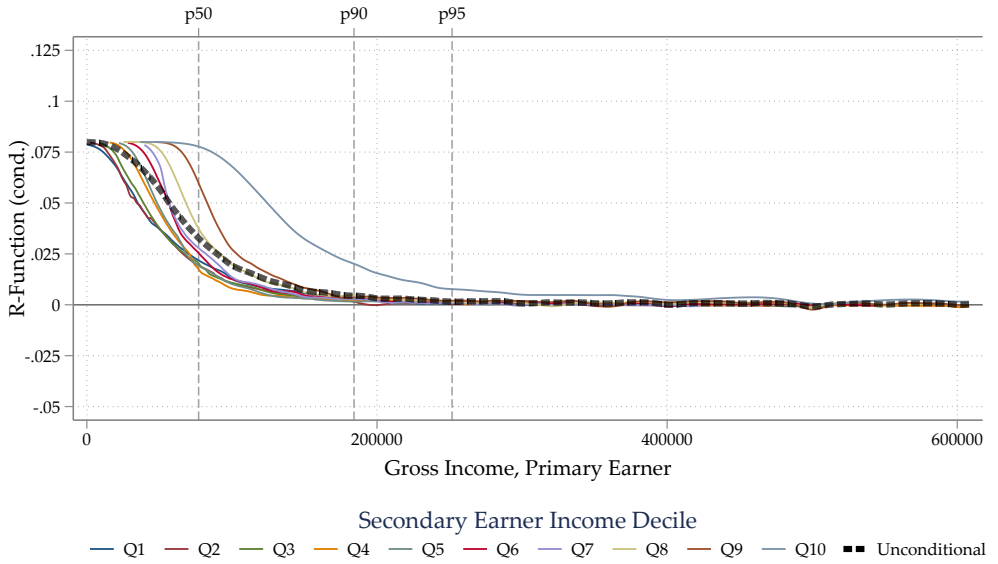
(B) Primary Earners

Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.15 (0.35) for primary (secondary) earners and an extensive margin elasticity of 0.2. The figure also displays modified unconditional revenue functions for secondary and primary earners where unconditional revenue functions have been scaled by 0.1 to facilitate comparability with the conditional revenue functions. Results for the baseline elasticity scenario are shown in Figure 2.5 of the main text.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.29. – Cond. Revenue Functions, Deciles (2019), High Elasticity



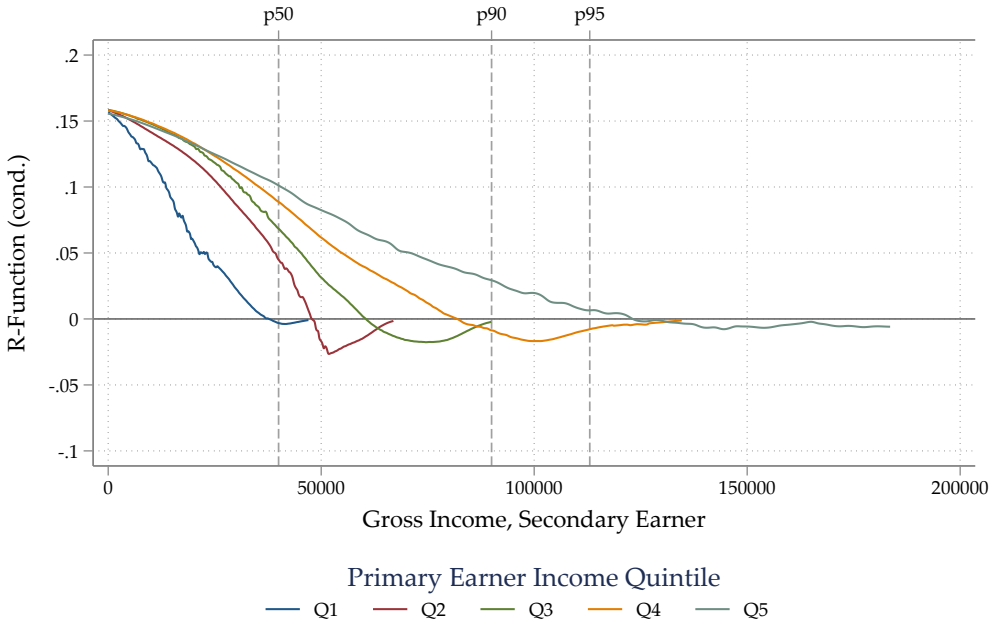
(A) Secondary Earners



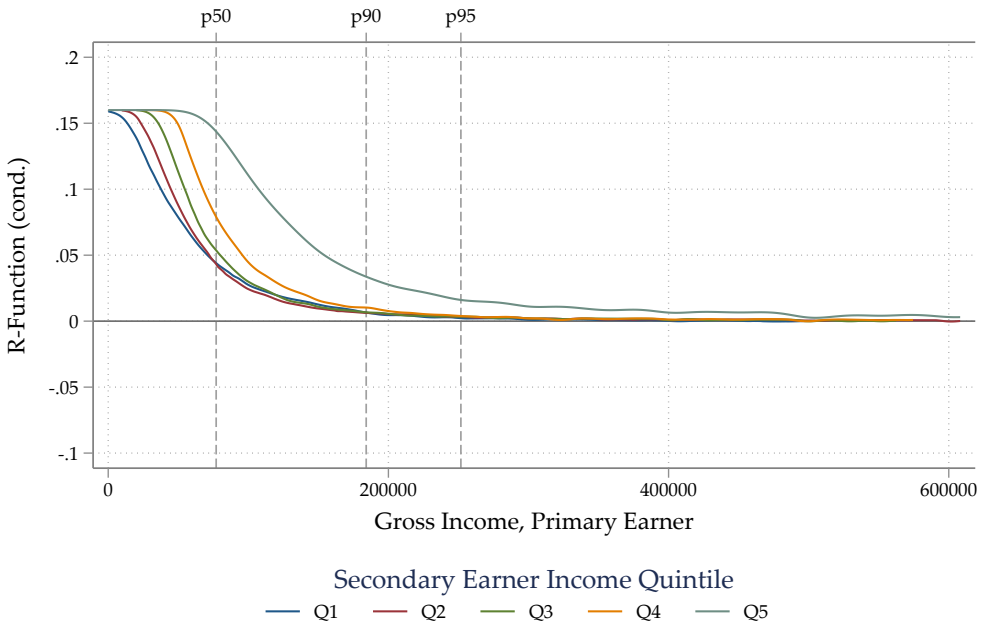
(B) Primary Earners

Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.5 (1.5) for primary (secondary) earners and an extensive margin elasticity of 0.2. The figure also displays modified unconditional revenue functions for secondary and primary earners where unconditional revenue functions have been scaled by 0.1 to facilitate comparability with the conditional revenue functions. Results for the baseline elasticity scenario are shown in Figure 2.5 of the main text.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.30. – Cond. Revenue functions, Quintiles (2019)



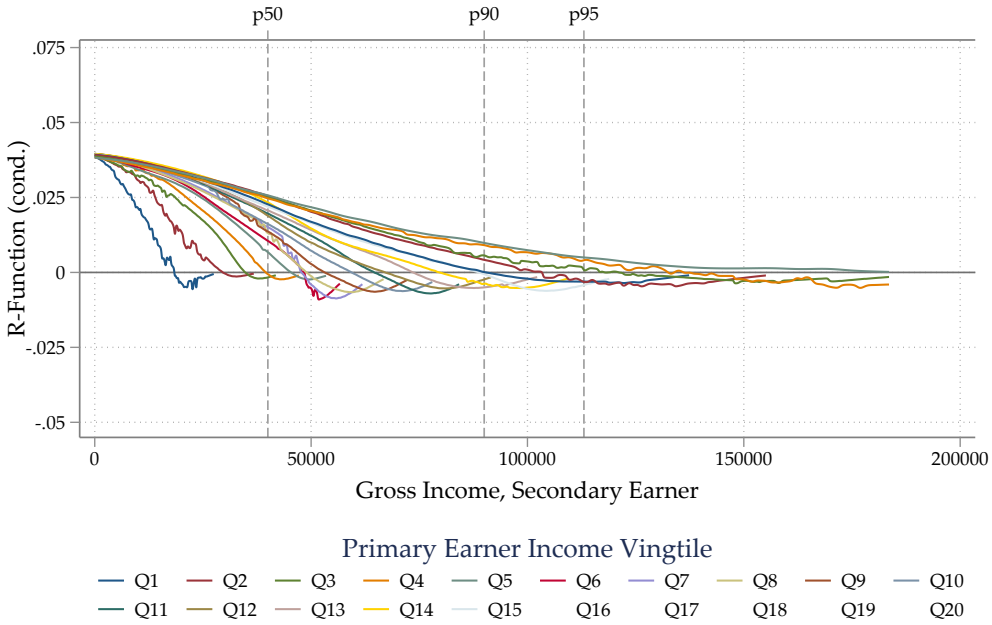
(A) Secondary Earners



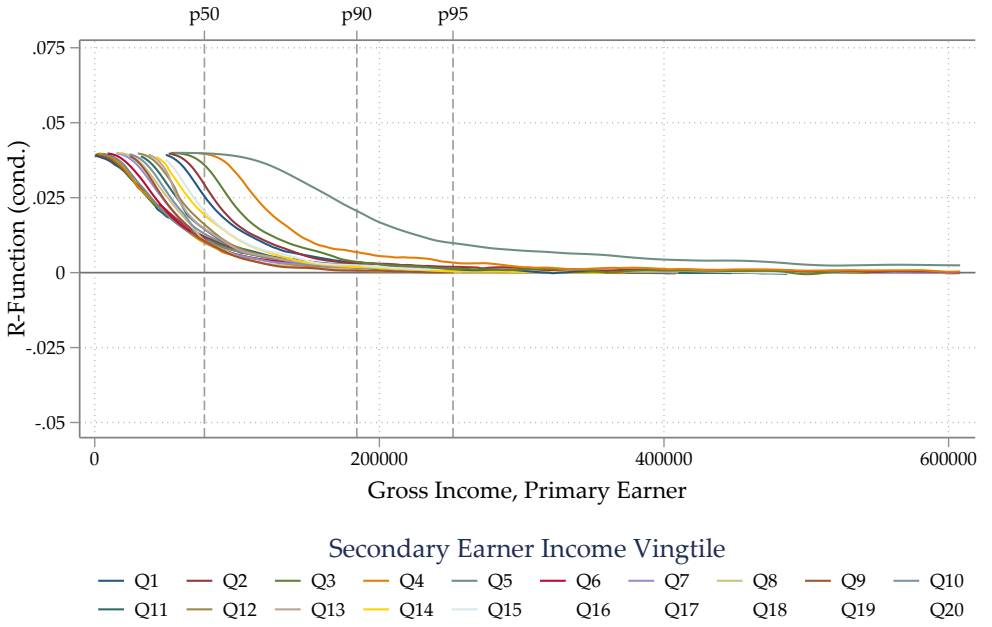
(B) Primary Earners

Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income quintiles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.31. – Cond. Revenue Functions, Vingtiles (2019)



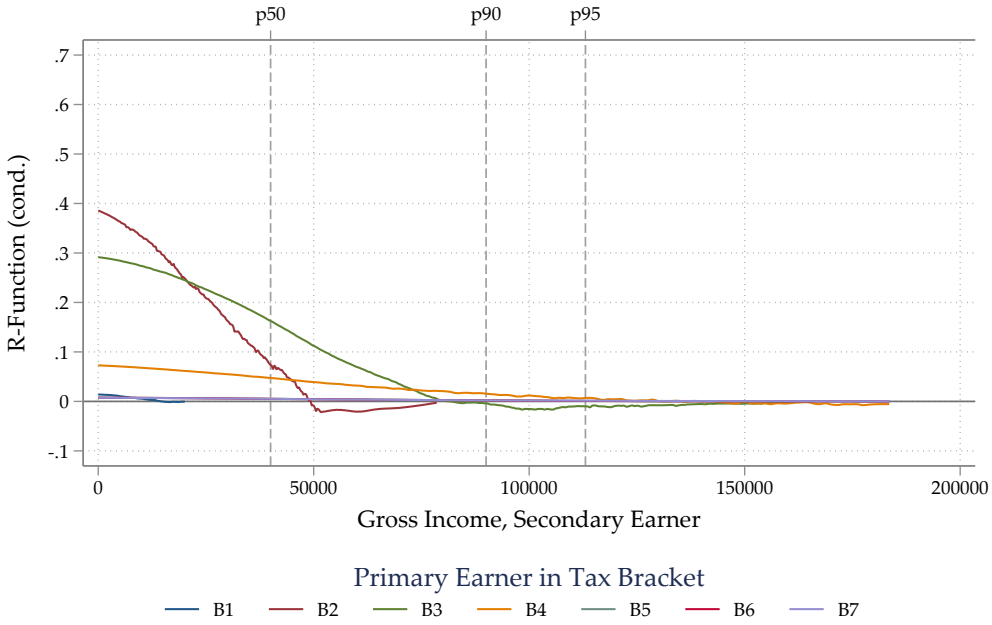
(A) Secondary Earners



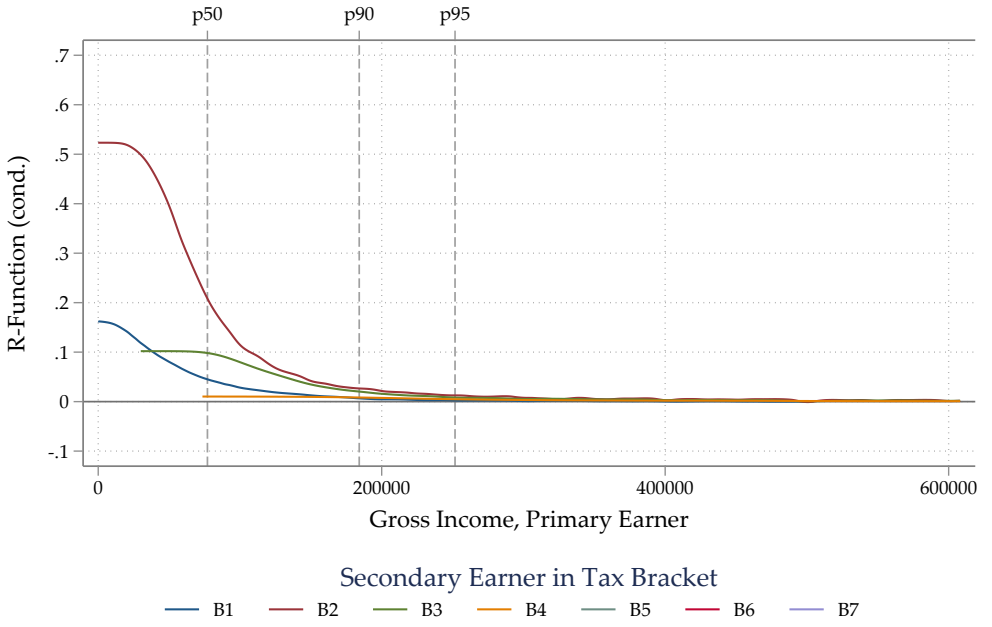
(B) Primary Earners

Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income vingtiles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.32. – Cond. Revenue Functions, Tax Brackets (2019)



(A) Secondary Earners



(B) Primary Earners

Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings in specific income tax brackets in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019).

B.2.3 Reducing Jointness

In Section 2.3.3 of the main text, we estimate the welfare effects of jointness-reducing reforms at different combinations of primary and secondary earnings. In this section of the appendix, we provide detailed information on the construction of empirical welfare assessments (Section B.2.3), and show supplementary results and robustness checks (Section B.2.3).

From Theory to Data

Following equation (2.13), the welfare effect of a reform can be approximated as

$$W = \tau\ \ell_1\ \ell_2 \Bigg[\mathcal{R}_2\left(y_{2s} \mid [y_{1s}, \bar{y}_1]\right) + \mathcal{R}_1\left(y_{1s} \mid [y_{2s}, \bar{y}_2]\right) - 2\ s\left(y_{1s}, y_{2s}\right) \mathcal{G}\left(y_{1s}, y_{2s}\right) \Bigg] \tag{B.87}$$

where the first two terms are revenue effects coming from the reduction in marginal tax rates for secondary (primary) earners conditional on primary (secondary) earnings ranges while the third term captures the adjustment required by specific welfare objectives. The conditional revenue functions are estimated as

$$\begin{aligned} \mathcal{R}_i\left(y_i \mid \left[\underline{b}_j, \bar{y}_j\right]\right) = & \underbrace{s(y_j \geq \underline{b}_j)}_{\text{Share in Inc. Range}} \Bigg(\underbrace{-\overline{\beta}^d_{I,i}\left(y_i, y_j\right) f^{y_i}(y_i \mid y_j \geq \underline{b}_j)}_{\text{Intensive Margin}} \Bigg) \\ + & \underbrace{s(y_j \geq \underline{b}_j)}_{\text{Share in Inc. Range}} \underbrace{\left(1 - F^{y_i}(y_i \mid y_j \geq \underline{b}_j)\right)}_{\text{Cond. CDF}} \underbrace{\left(1 - \mathcal{E}^d_x\right)}_{\text{Extensive Margin}} \end{aligned} \tag{B.88}$$

$$\overline{\beta}^d_{I,i} = \mathbb{E}_{y_i}\left[\frac{T^{0\prime}\left(y^{d0}\right)}{1 - T^{0\prime}\left(y^{d0}\right)} y_i^{d0} \varepsilon_i \mid y_j \geq \underline{b}_j \right] \quad , \quad (i,j) \in \{(1,2),(2,1)\} \quad . \tag{B.89}$$

Specifications. For consistency reasons, assumptions about behavioral responses follow the previous section (see Table B.2). We use three different specifications of welfare weights displayed in Table B.5. Reforms are carried out at bracket combinations defined by the primary earnings deciles and the respective conditional secondary earnings deciles (see Table B.6).

Table B.5. – Welfare Weights

Welfare Weight Type	Welfare Weight
Equal Weights	$g(y_1, y_2) = 1$
Quasi-Rawlsian Weights	$g(y_1, y_2) = \begin{cases} 1, & \text{for } y_1 + y_2 \leq \mathrm{P}_{10} \\ 0, & \text{for } y_1 + y_2 \geq \mathrm{P}_{10} \end{cases}$
Feminist Weights	$g(y_1, y_2) = \begin{cases} \frac{y_2}{y_1+y_2}, & \text{for } y_2 = y_{\text{woman}} \\ \frac{y_1}{y_1+y_2}, & \text{for } y_1 = y_{\text{woman}} \end{cases}$

Note: This table shows the specifications of different welfare weights used in the empirical analysis. For quasi-Rawlsian weights, P_{10} refers to the tenth percentile of the joint income distribution of dual earner couples. Welfare weights are normalized to mean one.

Table B.6. – Jointness Reforms, Selected Reform Brackets (2019)

Primary earnings bracket	$\underline{b}_1$	Secondary earnings bracket, $\underline{b}_2$									
		$B_2(1)$	$B_2(2)$	$B_2(3)$	$B_2(4)$	$B_2(5)$	$B_2(6)$	$B_2(7)$	$B_2(8)$	$B_2(9)$	$B_2(10)$
$B_1(1)$	1	1	3200	7488	10167	14001	16001	19001	20010	24001	27750
$B_1(2)$	35001	1	7500	15002	20001	24000	28002	30115	34001	36001	40001
$B_1(3)$	46003	1	8484	18006	25001	30001	34011	38001	41802	45603	50001
$B_1(4)$	55625	1	12001	20001	27002	33501	40001	45001	49792	52002	56004
$B_1(5)$	65255	1	14288	23150	30086	36040	42202	49550	53700	60006	65353
$B_1(6)$	77060	1	16102	26523	35012	42005	50001	55271	61100	70005	76401
$B_1(7)$	91036	1	18200	30005	39200	46225	52800	60124	70010	79501	88060
$B_1(8)$	109000	1	16200	28307	37757	47000	55028	65125	76150	90001	101051
$B_1(9)$	135031	1	16001	30005	42510	55001	65218	77252	91088	107345	128001
$B_1(10)$	184202	1	15106	30014	45200	59100	73500	92301	114400	140223	186300

Note: This table shows the lower bracket thresholds used for the analysis of different jointness-reducing reforms. Primary earner brackets (vertical) are chosen based on deciles of the primary earner income distribution. Secondary earner brackets (horizontal) are chosen based on the secondary earner deciles conditional on the respective primary earner decile. The red (blue) cells indicate the bracket combination, for which reforms are discussed in detail (Figure B.33 and Figure B.34). *Source:* Own calculations based on CPS-ASEC (2019).

Mechanics and Intuition. To understand the mechanics behind the welfare evaluation based on the selection of welfare weights in Table B.5, it can be helpful to rewrite and combine equations (B.87) - (B.89) and decompose the welfare effect as

$$
\begin{aligned}
W = \tau\, l(B_1)\, l(B_2) \Bigg[& \underbrace{s(y_1 \geq \underline{b}_1) f^{y_2}(\underline{b}_2 \mid y_1 \geq \underline{b}_1)}_{\text{Share Affected, Behavioral Effect}} \Bigg(\underbrace{-\overline{\beta}^d_{I,2}\left(y_1, \underline{b}_2\right)}_{\text{Intensive Margin}} \Bigg) \\
& + \underbrace{s(y_1 \geq \underline{b}_1, y_2 \geq \underline{b}_2)}_{\text{Share Affected, Mechanical Effect}} \Bigg(1 - \underbrace{\mathcal{E}^d_x}_{\text{Ext. margin}} - \underbrace{\mathcal{G}\left(\underline{b}_1, \underline{b}_2\right)}_{\text{Welfare}} \Bigg) \\
& + \underbrace{s(y_2 \geq \underline{b}_2) f^{y_1}(\underline{b}_1 \mid y_2 \geq \underline{b}_2)}_{\text{Share Affected, Behavioral Effect}} \Bigg(\underbrace{-\overline{\beta}^d_{I,1}\left(y_2, \underline{b}_1\right)}_{\text{Intensive Margin}} \Bigg) \\
& + \underbrace{s(y_2 \geq \underline{b}_2, y_1 \geq \underline{b}_1)}_{\text{Share Affected, Mechanical Effect}} \Bigg(1 - \underbrace{\mathcal{E}^d_x}_{\text{Extensive Margin}} - \underbrace{\mathcal{G}\left(\underline{b}_2, \underline{b}_1\right)}_{\text{Welfare}} \Bigg) \Bigg].
\end{aligned}
\tag{B.90}
$$

Equation (B.90) first decomposes the welfare effect into the welfare effect based on (i) changes of secondary marginal tax rates conditional on primary earnings (lines 1+2), and the welfare effect based on changes of primary marginal tax rates conditional on secondary earnings (lines 3+4). Each of the two effects can be further decomposed into the change in tax revenue coming from intensive margin behavioral responses (lines 1 and 3), and from the mechanical revenue effect corrected for both extensive margin responses and welfare evaluations (lines 2 and 4). For instance, an increase in secondary earner marginal tax rates triggers intensive margin behavioral responses for all couples with primary earnings above $\underline{b_1}$ and secondary earnings at $\underline{b_2}$. In contrast, all couples with primary earnings above $\underline{b_1}$ and secondary earnings above $\underline{b_2}$ pay now more taxes due to the reform. This mechanical effect is potentially downscaled by the extensive margin elasticity representing dual earner couples that decide to become single earner couples as a response to the increased tax liability. In addition, the mechanical effect is further downscaled by how much one cares about those couples that pay for the increase in revenue in terms of a higher tax liability.

Revisiting the social welfare objectives specified in Table B.5 can provide some additional intuition. If the social planner operates under a Rawlsian welfare objective, the welfare terms are zero if the reform is done at income levels that are above the tenth decile of the joint income distribution. Welfare evaluations then boil down to how much tax revenue is raised through the reform. Under equal welfare weights, in contrast, one would care about all couples equally, i.e. a one dollar higher mechanical tax revenue increase is completely counteracted by lower welfare of those that are paying for this revenue with a higher tax liability. Welfare evaluations then boil down to the intensive and extensive margin behavioral responses that are triggered by the reform.

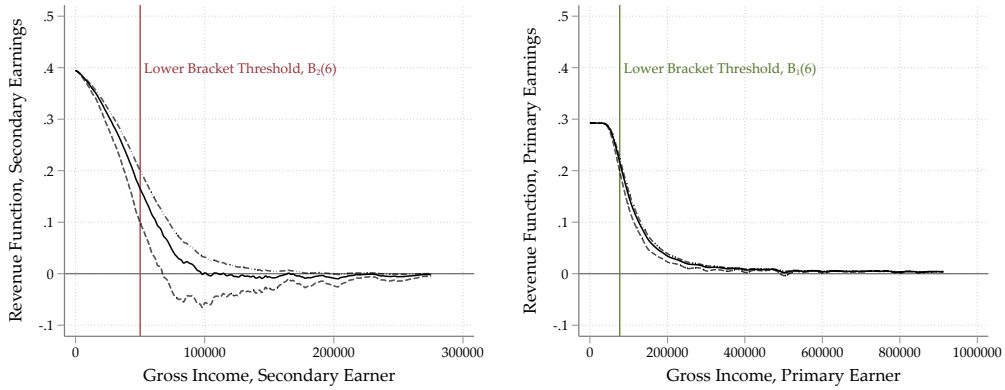
Supplementary Graphs and Robustness

Figure B.33 provides supplementary details on the empirical conditional revenue functions that are used to estimate the welfare effect for a reduction of jointness around the median in Figure 2.6. We also provide an illustration of the welfare effect for a bracket combination of mid levels of primary earnings and high secondary earnings that is discussed in the main text (Figure B.34). The effects of a reform at this bracket combination illustrates that the revenue effects from jointness-reducing reforms can move in opposite directions for the primary and secondary earner marginal tax rate changes. In the particular case, due to the inefficiencies in the tax treatment of secondary earnings, the tax cut for secondary earner marginal tax rates yields a revenue increase. However, the corresponding revenue loss

coming from the tax cut of primary earner marginal tax rates is higher. Consequently, despite the inefficiencies, a quasi-Rawlsian welfare function would still not agree to a jointness-reducing reform in this area.

Figure B.35 replicates the results on the welfare implications of changing jointness at different bracket combinations and illustrates, how results change under different elasticity scenarios. Figure B.36 provides the results without extensive margin responses.

Figure B.33. – Welfare, Reducing Jointness at $B_1(6)$, $B_2(6)$



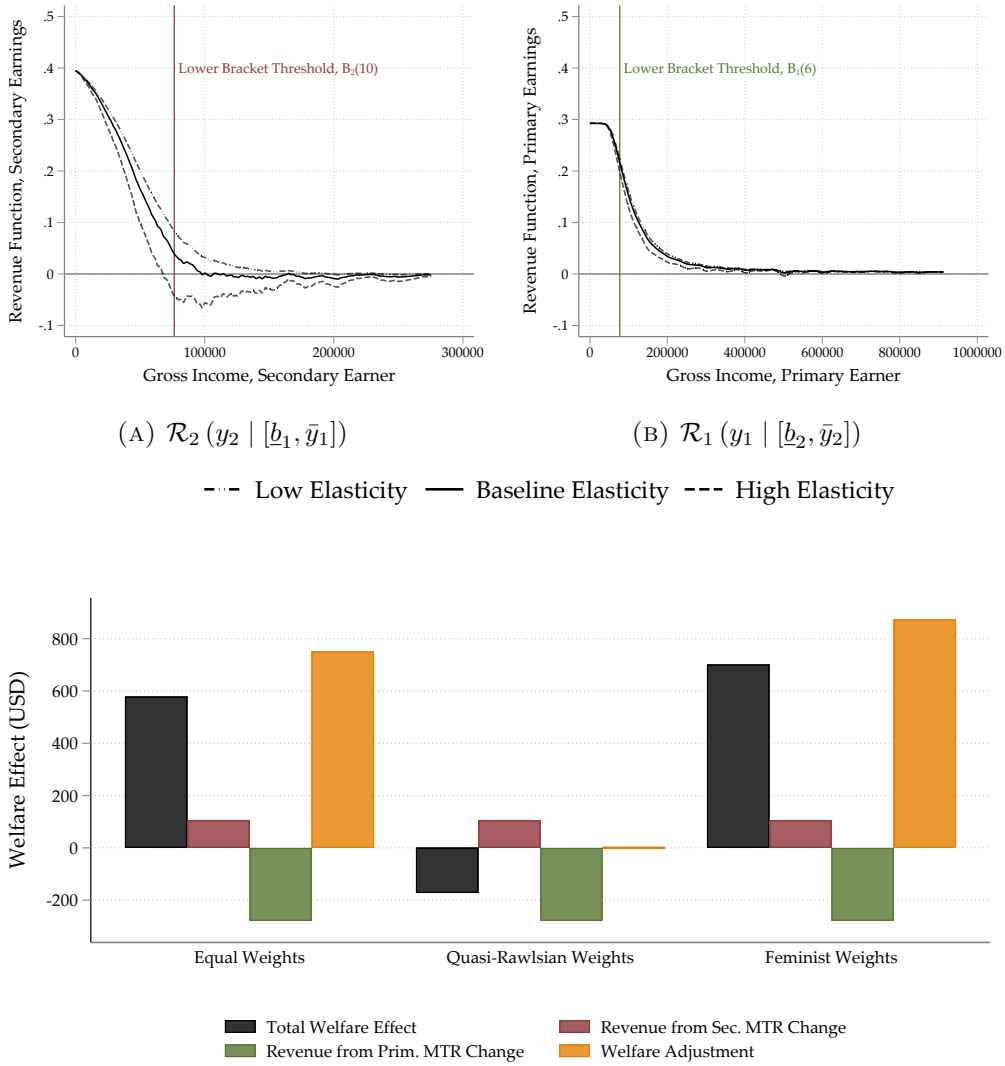
(A) $\mathcal{R}_2\left(y_2 \mid [\underline{b}_1, \bar{y}_1]\right)$ (B) $\mathcal{R}_1\left(y_1 \mid [\underline{b}_2, \bar{y}_2]\right)$

Note: This figure visualizes the estimation of the two revenue effects required for estimating the welfare effect implied by a reduction in jointness for $(y_1, y_2) \in B_1 times B_2$, where B_1 is the 6th-decile of the distribution of primary earnings and B_2 is the sixth decile in the distribution of secondary earnings conditional on $y_1 \in B_1$ (see Figure 2.6). Figure B.33b shows the relevant conditional revenue function describing the effect of increasing marginal tax rates for primary earners above a certain threshold conditional on secondary earnings being above the bracket threshold b_2 of the bracket $B_2(6)$. The solid red (green) lines indicate the lower bracket thresholds of secondary (primary) earnings that are relevant for the welfare evaluation, i.e. b_2 (b_1). Both revenue functions distinguish between different intensive margin elasticity scenarios and include extensive margin behavioral responses.
Source: Own calculations based on CPS-ASEC (2019).

B.2.4 Tax Reforms at the Bottom

In Section 2.3.4 of the main text, we estimate welfare effects of tax reforms at the bottom of the income distribution. This section presents graphs on the shape of welfare weights (Figure B.37) and robustness tests on the exclusion of extensive margin responses (Figure B.38) and the choice of elasticities (Figure B.39).

Figure B.34. – Welfare, Reducing Jointness at $B_1(6)$, $B_2(10)$



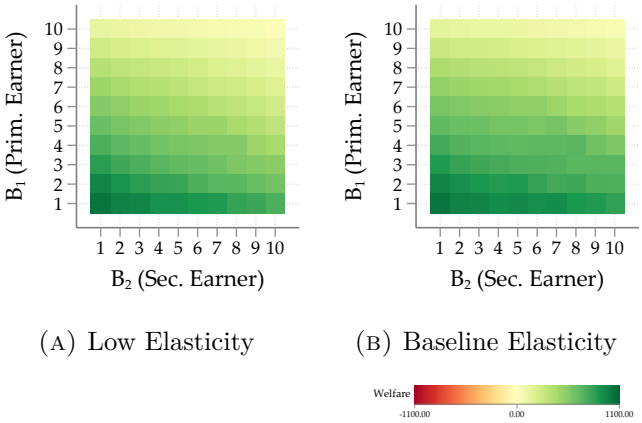
(A) $\mathcal{R}_2\left(y_2 \mid [\underline{b}_1, \bar{y}_1]\right)$ (B) $\mathcal{R}_1\left(y_1 \mid [\underline{b}_2, \bar{y}_2]\right)$

(C) Welfare effect

Note: This figure visualizes the welfare effects implied through a change in jointness at the bracket combination $B_1(6)$ and $B_2(10)$. Figure B.34a shows the relevant conditional revenue function describing the effect of increasing marginal tax rates for secondary earners above a certain threshold conditional on primary earnings being above the bracket threshold b_1 of the bracket $B_1(6)$. Figure B.34b shows the relevant conditional revenue function describing the effect of increasing marginal tax rates for primary earners above a certain threshold conditional on secondary earnings being above the bracket threshold b_2 of the bracket $B_2(10)$. The solid red (green) lines indicate the lower bracket thresholds of secondary (primary) earnings that are relevant for the welfare evaluation, i.e. b_2 (b_1). Both revenue functions distinguish between different intensive margin elasticity scenarios and include extensive margin behavioral responses. Figure B.34c displays the aggregate welfare effect and its components based on the bracket lengths $l(B_1) = l(B_2) = 500$ and $\tau = -0.01$, and for the high intensive margin elasticity scenario, and including extensive margin responses. The black bar indicates the aggregate welfare effect. The red (green) bar visualizes the revenue change coming from the decrease of secondary (primary) earnings. The orange bar illustrates the welfare adjustment.
Source: Own calculations based on CPS-ASEC (2019).

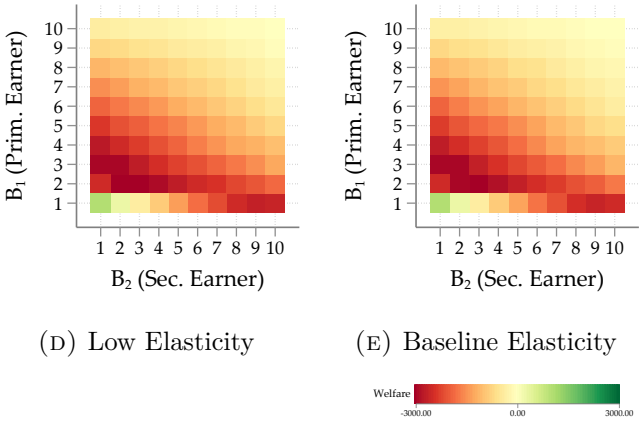
Figure B.35. – Decreasing Jointness, Varying Elasticities

Equal Weights



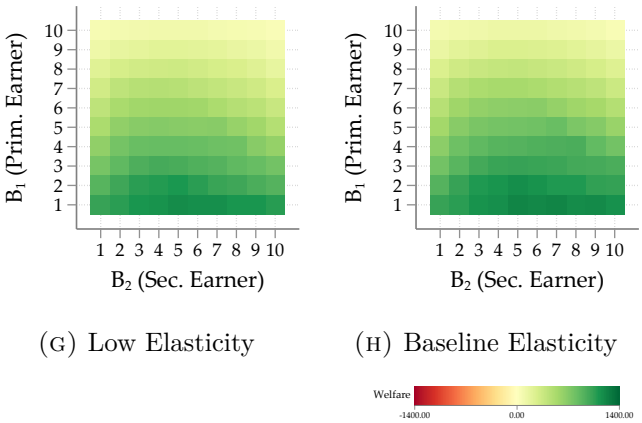
(A) Low Elasticity (B) Baseline Elasticity (C) High Elasticity

Quasi-Rawlsian Weights



(D) Low Elasticity (E) Baseline Elasticity (F) High Elasticity

Feminist Weights

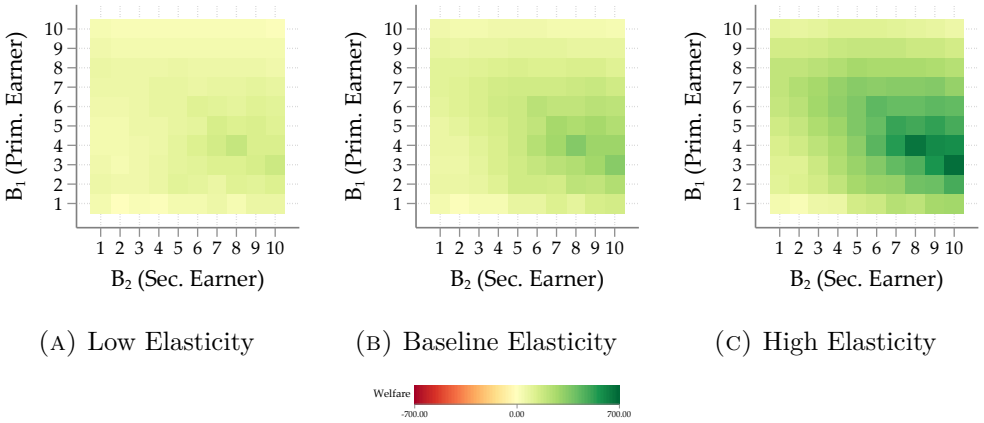


(G) Low Elasticity (H) Baseline Elasticity (I) High Elasticity

Note: This figure displays the welfare implications of decreasing jointness by decreasing secondary and primary marginal tax rates conditional on particular brackets of primary earnings and secondary earnings. Bracket thresholds are shown in Table B.6. The reform applies to a bracket of length $l(B_1) = l(B_2) = 500$ and has the magnitude $\tau = -0.01$. The figure distinguishes between three different forms of welfare weights (see Table B.5) and three elasticity scenarios (see Table B.2). All results are shown including extensive margin responses. Results without extensive margin responses are shown in Table B.36.
Source: Own calculations based on CPS-ASEC (2019).

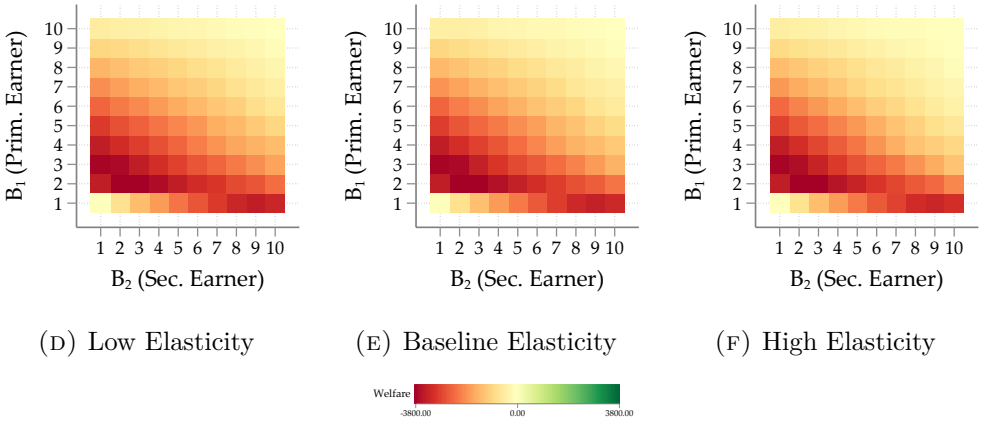
Figure B.36. – Decreasing Jointness, No Extensive Margin

Equal Weights



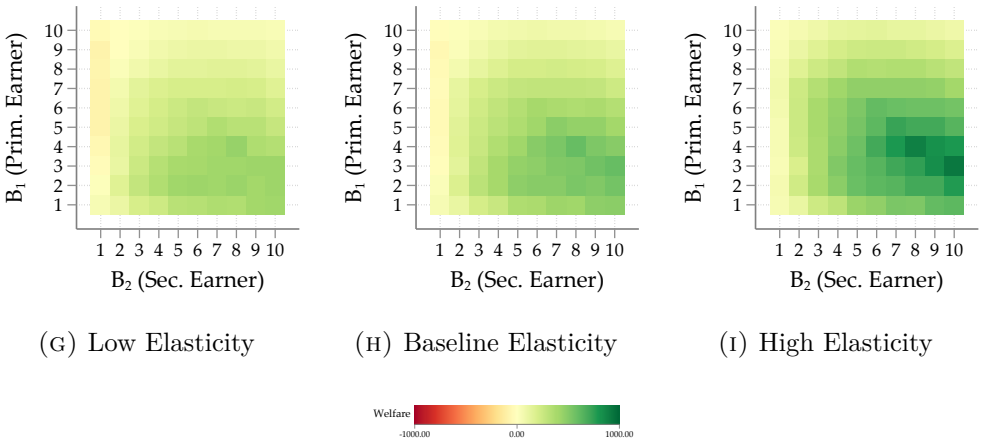
(A) Low Elasticity (B) Baseline Elasticity (C) High Elasticity

Quasi-Rawlsian Weights



(D) Low Elasticity (E) Baseline Elasticity (F) High Elasticity

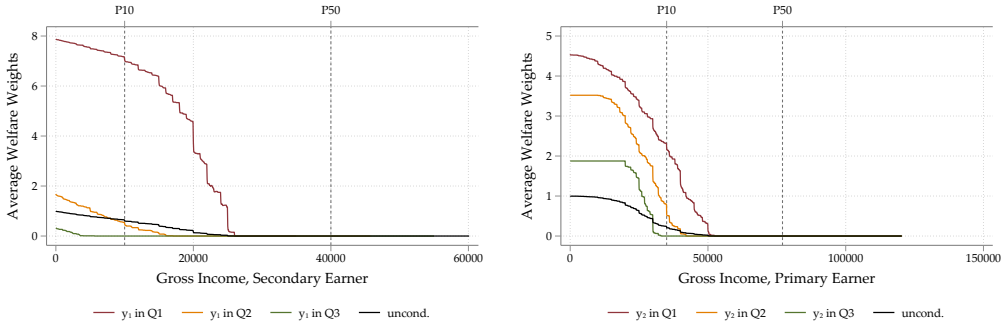
Feminist Weights



(G) Low Elasticity (H) Baseline Elasticity (I) High Elasticity

Note: This figure displays the welfare implications of decreasing jointness by decreasing secondary and primary marginal tax rates conditional on particular brackets of primary earnings and secondary earnings. Bracket thresholds are shown in Table B.6. The reform applies to a bracket of length $l(B_1) = l(B_2) = 500$ and has the magnitude $\tau = -0.01$. The figure distinguishes between three different forms of welfare weights (see Table B.5) and three elasticity scenarios (see Table B.2). All results are shown excluding extensive margin responses. Results with extensive margin responses are shown in Figure B.35.
Source: Own calculations based on CPS-ASEC (2019).

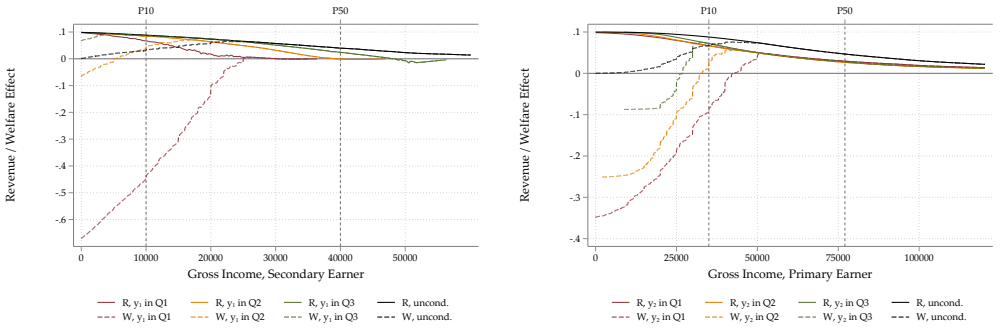
Figure B.37. – Conditional Reforms at Bottom (2019), Welfare Weights



(A) Secondary Earnings (B) Primary Earnings

Note: This figure shows average welfare weights used for the calculation of welfare effects in Figure 2.8. Weights for secondary (primary) earners above a specific value conditional on primary (secondary) earnings income deciles at the bottom are shown in Panel a (b). Weights are based on a quasi-Rawlsian welfare function (see Table B.5) which concentrates weights on couples below the 10th decile of the joint income distribution. Black lines indicate unconditional welfare weights.
Source: Own calculations based on CPS-ASEC (2019).

Figure B.38. – Conditional Reforms at Bottom (2019), No Extensive Margin

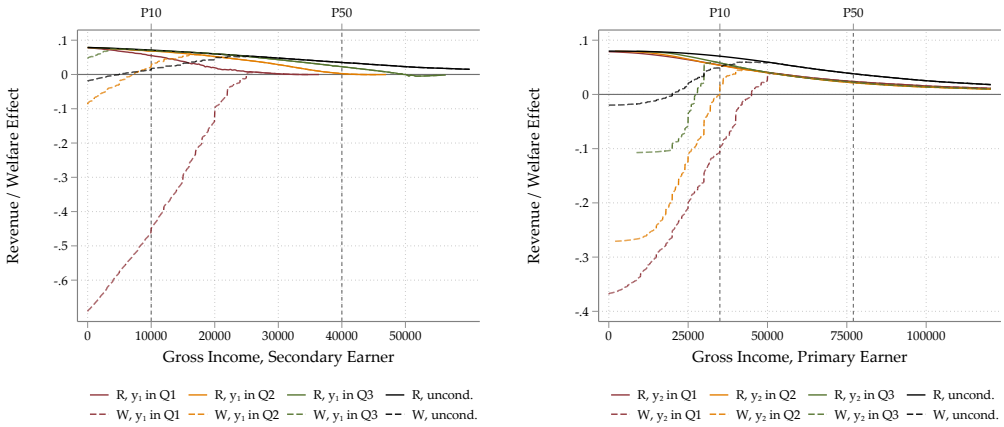


(A) Secondary Earnings (B) Primary Earnings

Note: This figure replicates Figure 2.8 without extensive margin responses. The figure shows the revenue and welfare effects for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles at the bottom of the income distribution in Panel a (b) as of 2019. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners. The figure displays the revenue effect in terms of conditional revenue functions (solid lines) and the welfare effect based on quasi-Rawlsian welfare weights (dashed lines).
Source: Own calculations based on CPS-ASEC (2019).

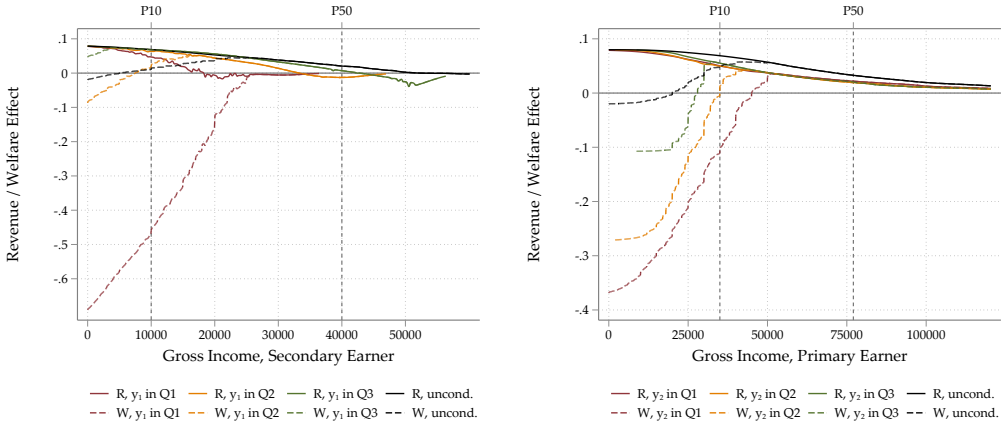
Figure B.39. – Conditional Reforms at Bottom (2019), Different Elasticities

Low Elasticity Scenario



(A) Secondary Earnings (B) Primary Earnings

High Elasticity Scenario



(C) Secondary Earnings (D) Primary Earnings

Note: This figure replicates Figure 2.8 under the low elasticity scenario from Table B.2. The figure shows the revenue and welfare effects for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles at the bottom of the income distribution in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. In the low elasticity scenario, we assume intensive margin elasticities of 0.15 (0.35) for primary (secondary) earners. In the high elasticity scenario, we assume intensive margin elasticities of 0.5 (1.5) for primary (secondary) earners. The extensive margin scaling factor is held constant at 0.9. The figure displays the revenue effect in terms of conditional revenue functions (solid lines) and the welfare effect based on quasi-Rawlsian welfare weights according to Table B.5 (dashed lines). Black lines indicate unconditional revenue and welfare functions.
Source: Own calculations based on CPS-ASEC (2019).

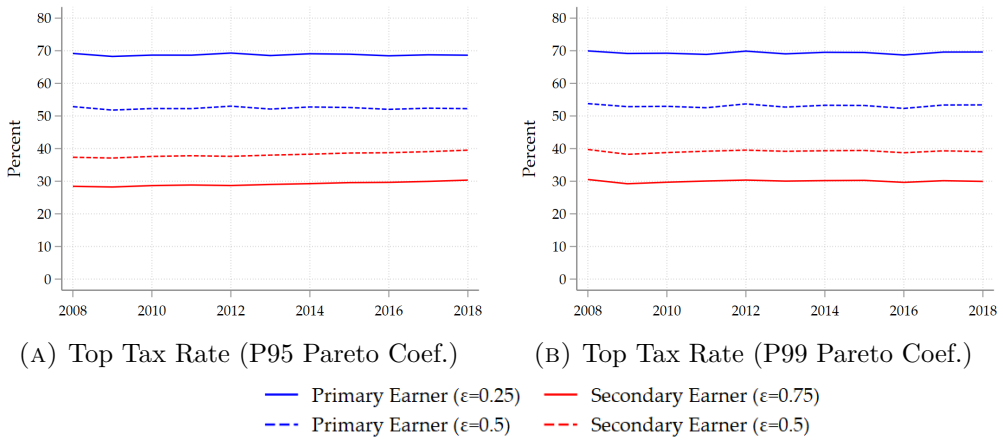
B.2.5 Optimal Top Tax Rates

In Section 2.3.5 of the main text, we use tabular income tax information on all married couples for the estimation of Pareto coefficients and associated optimal tax rates in the most recent available year. In the following, we provide insights on the stability of optimal tax rates over time (Section B.2.5), heterogeneity analyses that estimate Pareto coefficients and optimal tax rates separately for single earner and dual earner couples (Section B.2.5), and robustness tests that estimate Pareto coefficients from CPS data (Section B.2.5).

Optimal Tax Rates over Time

Figure B.40 presents optimal tax rates based on Pareto coefficients that have been calculated at the 95th and 99th percentile for all available years of the tabulated income tax return data. Since the distributional properties did not change over this relatively recent time horizon, optimal top tax rates stay rather constant.

Figure B.40. – Optimal Top Tax Rates over Time



(A) Top Tax Rate (P95 Pareto Coef.) (B) Top Tax Rate (P99 Pareto Coef.)

Note: This figure displays optimal top tax rates for different values of the Frisch elasticity based on the Pareto coefficients estimated from tabulated data on wages on W2 forms for joint return taxpayers with wage income.
Source: Own calculations based on SOI Tax Stats - Individual Information Return Form W-2 (2018).

Heterogeneity Analyses

Tables B.7 (B.8) display Pareto coefficients and respective optimal tax rates for single earner couples (dual earner couples with relatively equal income shares). The tables pro-

vide two insights. First, the Pareto coefficient for single earner couples (primary earners) is much lower than for dual earner couples (primary earner and secondary earner) pushing towards higher optimal tax rates for single earner couples compared to dual earner couples. Second, within dual earner couples, the optimal tax rate gap between primary and secondary earners almost vanishes, since Pareto coefficients are very similar.

Table B.7. – Pareto Coefficients and Optimal Tax Rates, Single Earner Couples (2018)

	P95		P99	
	Primary	Secondary	Primary	Secondary
Panel A: Pareto coefficients				
Wages on W2 Form	1.65	-	1.61	-
Panel B: Optimal tax rates				
Elasticity = .25	71%	-	71%	-
Elasticity = .5	55%	-	55%	-
Elasticity = .75	45%	-	45%	-

Note: This table shows Pareto coefficients and optimal top tax rates. Panel A displays Pareto coefficients for primary and secondary earners based on a generalized Pareto interpolation using tabulated data on wages on W2 forms for joint return single earner couples with wage income. We distinguish between interpolations at the 95th and the 99th percentile. Panel B displays optimal top tax rates associated with these Pareto coefficients. We distinguish between different elasticities.
Source: Own calculations based on SOI Tax Stats - Individual Information Return Form W-2 Statistics (2018)

Table B.8. – Pareto Coefficients and Optimal Tax Rates, Dual Earner Couples (2018)

	P95		P99	
	Primary	Secondary	Primary	Secondary
Panel A: Pareto Coefficients				
Wages on W2 Form	2.95	3.07	3.16	3.22
Panel B: Optimal Tax Rates				
Elasticity = .25	58%	57%	56%	55%
Elasticity = .5	40%	39%	39%	38%
Elasticity = .75	31%	30%	30%	29%

Note: This table shows Pareto coefficients and optimal top tax rates. Panel A displays Pareto coefficients for primary and secondary earners based on a generalized Pareto interpolation using tabulated data on wages on W2 forms for joint return dual earner couples with wage income. Dual earner couples have a secondary earner income share of at least 25 percent. We distinguish between interpolations at the 95th and the 99th percentile. Panel B displays optimal top tax rates associated with these Pareto coefficients. We distinguish between different elasticities.
Source: Own calculations based on SOI Tax Stats - Individual Information Return Form W-2 Statistics (2018)

Robustness

In the main text, we estimate Pareto coefficients with tabular income tax return data that provides high quality at the top of the distribution. However, we can also calculate Pareto coefficients directly using CPS data. For this purpose, we estimate the empirical cumulative distribution function (ECDF) and plot the log-transformed ECDF against the log-transformed income level for specific percentiles (see Figure B.41). The slope of the regression lines indicate the Pareto coefficient. In line with main results from tabulated tax return data, the Pareto coefficient for primary earners is larger than the one for secondary earners. This holds irrespective of the type of income used in the estimation and the sample restriction (see Table B.9). As shown in Table B.10, the optimal tax rate gap between primary and secondary earners is even larger than what is suggested by the tabulated data.

Table B.9. – Pareto Coefficients, CPS (2018)

Sample	Income	Top 5%		Top 1%	
		Primary	Secondary	Primary	Secondary
Baseline	Wages	1.63 (N=933)	3.35 (N=953)	1.37 (N=206)	3.27 (N=194)
	Wage, Business, Farm Inc.	1.61 (N=942)	3.33 (N=1022)	1.37 (N=194)	3.33 (N=216)
	Gross Income	1.64 (N=925)	3.28 (N=933)	1.42 (N=187)	3.28 (N=185)
Alt.	Wages	1.72 (N=1700)	3.25 (N=1743)	1.42 (N=378)	3.21 (N=358)
	Wage, Business, Farm Inc.	1.69 (N=1789)	3.13 (N=1871)	1.4 (N=350)	3.04 (N=351)
	Gross Inc.	1.74 (N=1705)	3.05 (N=1729)	1.47 (N=348)	3 (N=333)

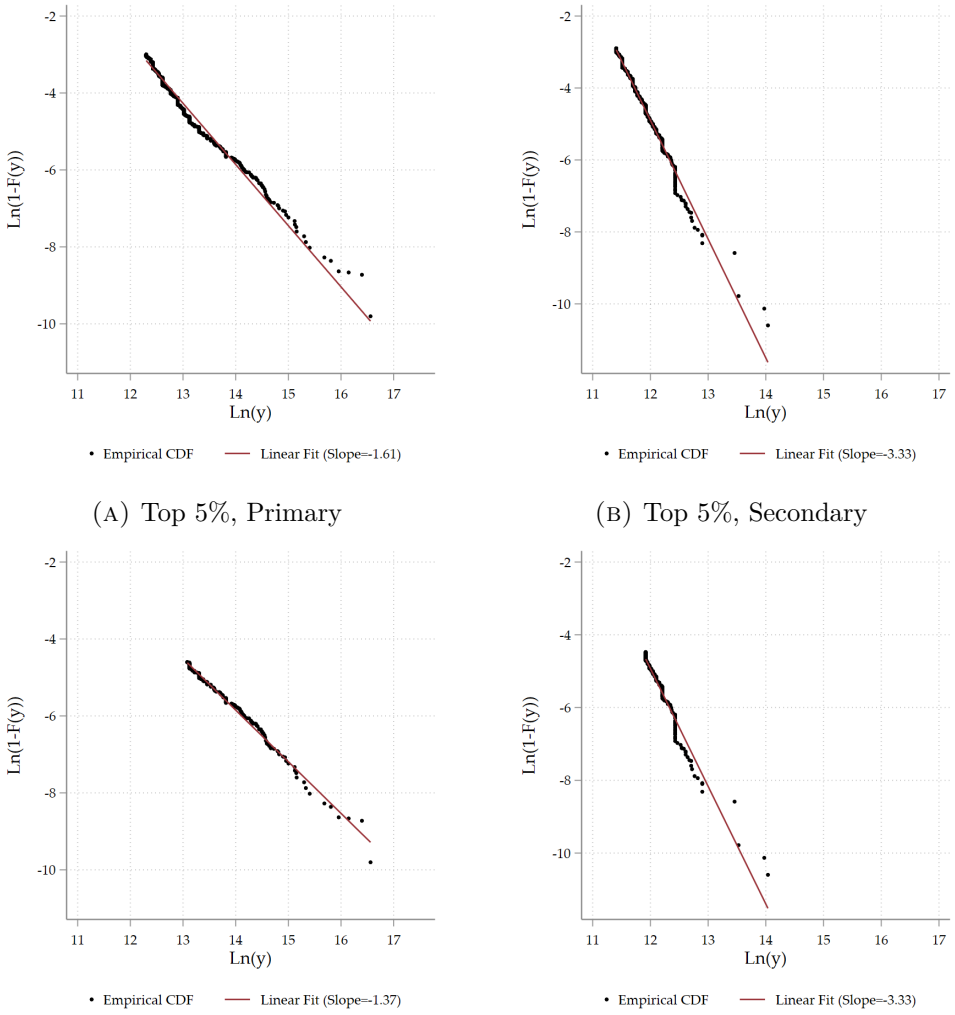
Note: This table shows the Pareto coefficients estimated for different sample restrictions, different income concepts, and at different parts of the income distribution. The baseline sample is restricted to all married couples where both spouses are between 25 and 55 years old. Under an alternative sample restriction, we include all adult married spouses.
Source: Own calculations based on CPS-ASEC (2018)

Table B.10. – Optimal Top Tax Rates, CPS (2018)

Elasticity	P95		P99	
	Primary	Secondary	Primary	Secondary
.25	71%	55%	74%	55%
.5	55%	38%	59%	38%
.75	45%	29%	49%	29%

Note: This table shows optimal top tax rates for different values of the Frisch elasticity based on the Pareto coefficients estimated from CPS data under the baseline sample restriction where we focus on all married couples where both spouses are between 25 and 55 years old. Pareto coefficient have been estimated using earnings of primary and secondary earner that include wage, business, and farm income.
Source: Own calculations based on CPS-ASEC (2018)

Figure B.41. – Pareto Coefficient, CPS Data



(A) Top 5%, Primary

(B) Top 5%, Secondary

(C) Top 1%, Primary

(D) Top 1%, Secondary

Note: This figure displays the top earnings inequality among primary and secondary earners as of 2018. The sample is restricted to all married couples where both spouses are between 25 and 55 years old. The figure shows the shape of the empirical CDF separately for primary and secondary earners in married couples, i.e. $ln(1 - F(y))$ on the vertical and $ln(y)$ on the horizontal axis. The figure also shows the linear fit as a red line, the slope of which is the estimated Pareto coefficient. Earnings of primary and secondary earner include wage, business, and farm income.
Source: Own calculations based on CPS-ASEC (2018).

Appendix to Chapter 3

C.1 APPENDIX: THEORETICAL ANALYSIS

C.1.1 Proof of Theorem 4

A reform *of* the system lowers marginal tax rates for secondary earners or increases them for primary earners. Let

$$h_m(y_1, y_2) = \tau_1 \; h_1(y_1) + \tau_2 \; h_2(y_2) \; .$$

Then, the function $y_1 \mapsto \tau_1 \; h_1'(y_1)$ gives the change of marginal tax rates on primary earnings and the function $y_2 \mapsto \tau_2 \; h_2'(y_2)$ gives the change of marginal tax rates on secondary earnings.

By Theorem 3 in Bierbrauer et al. (2024), the marginal impact on overall tax revenue that is due to secondary earners is given by

$$\tau_2 \int_{\mathcal{Y}} h_2'(y_2)\mathcal{R}_2(y_2) \; dy_2 \; , \tag{C.1}$$

where $\mathcal{R}_2(y_2)$ is the revenue impact of a small one-bracket reform that changes the tax rate for secondary earnings close to y_2. Analogously, the revenue change due to primary earnings is given by

$$\tau_1 \int_{\mathcal{Y}} h_1'(y_1)\mathcal{R}_1(y_1) \; dy_1 \; , \tag{C.2}$$

where $\mathcal{R}_1(y_1)$ gives the marginal change in tax revenue when tax rates are increased for all primary earners with an income close to y_1.

By Equations (C.1) and (C.2), revenue neutrality requires that

$$\frac{\tau^1}{\tau^2} = -\frac{\int_{\mathcal{Y}} h_2'(y_2)\mathcal{R}_2(y_2) \; dy_2}{\int_{\mathcal{Y}} h_1'(y_1)\mathcal{R}_1(y_1) \; dy_1} \; . \tag{C.3}$$

A special case of interest is that marginal tax rates are increased for all primary earners and decreased for all secondary earners. In this case $h_1(y_1) = y_1$, for all y_1 and $h_2(y_2) = y_2$, for all y_2. Such a reform is revenue neutral if

$$\frac{\tau^1}{\tau^2} \quad = \quad -\frac{\int_{\mathcal{Y}} \mathcal{R}^2(y_2) dy_2}{\int_{\mathcal{Y}} \mathcal{R}^1(y_1) dy_1} \; . \tag{C.4}$$

A married couple that has earnings of y_1^0 and y_2^0 in the status quo is made better off if

$$\tau_1 h_1(y_1^0) + \tau_2 h_2(y_2^0) \quad = \quad \tau_1 \; y_1^0 + \tau_2 \; y_2^0 \quad < \quad 0 \; ,$$

or, equivalently, if

$$y_1^0 \quad < \quad \frac{\int_{\mathcal{Y}} \mathcal{R}^1(y_1)dy_1}{\int_{\mathcal{Y}} \mathcal{R}^2(y_2)dy_2} \; y_2^0 \; .$$

From the perspective of a generic social welfare function, a revenue neutral reform with $h_m(y_1, y_2) = \tau_1 \; y_1 + \tau_2 \; y_2$ is desirable if and only if

$$\mathbf{E}_{(\theta_m,\gamma_m)} \left[\mathbf{g}_m(\gamma_m, \theta_m) \; y_1^0(\gamma_m, \theta_m) \right]$$

$$< \left(\frac{\int_{\mathcal{Y}} \mathcal{R}^1(y_1)dy_1}{\int_{\mathcal{Y}} \mathcal{R}^2(y_2)dy_2} \right) \; \mathbf{E}_{(\theta_m,\gamma_m)} \left[\mathbf{g}_m(\gamma_m, \theta_m) \; y_2^0(\gamma_m, \theta_m) \right] \; .$$

C.1.2 Revenue Functions

We derive a characterization of the revenue functions $\mathcal{R}_1$ and $\mathcal{R}_2$. The proof of Proposition 4 builds on Proposition 1 in Bierbrauer et al. (2024).

In Bierbrauer et al. (2024), we analyze simple tax reforms that affect the marginal tax rate of one spouse, conditional on the income of the other spouse being in a narrow range of incomes. This yields conditional revenue functions. The revenue functions $\mathcal{R}_1$ and $\mathcal{R}_2$ considered in this paper are obtained by integrating over the conditioning variable. We now introduce these objects more formally.

A reform that raises the marginal tax rate on the earnings of spouse 2, conditional on the earnings of spouse 1 belonging to a bracket $B_1(y_{1s}, \ell_1) = [y_{1s}, y_{1s} + \ell_1]$ and the earnings of spouse 2 belonging to a bracket $B_2(y_{2s}, \ell_2) = [y_{2s}, y_{2s} + \ell_2]$ can be represented by a scalar τ and a function $y \mapsto h(y)$ such that

$$h(y) \quad = \quad \begin{cases} y_2 - y_{2s}, & \text{if} \quad y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2) \; , \\ \ell_2, & \text{if} \quad y_1 \in B_1(y_{1s}, \ell_1) \quad \text{and} \quad y_2 \geq y_{2s} + \ell_2 \; , \\ 0, & \text{otherwise} \; . \end{cases}$$

Consequently, the marginal tax rates on the earnings of spouse 2 change when the earnings of the spouses belong to the relevant brackets and do not change otherwise,

$$T_{y_2}^1(y) - T_{y_2}^0(y) = \begin{cases} \tau, & \text{if} \quad y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2) \; , \\ 0, & \text{otherwise} \; . \end{cases}$$

Reforms so that B_1 and B_2 are brackets with positive length affect the earnings incentives of dual-earner couples. We use different notation for reforms that affect the earnings incentives of single-earner couples. For instance, a pair (τ, h^{s1}), with

$$h^{s1}(y) \quad = \quad \begin{cases} y_1 - y_{1s}, & \text{if} \quad y_1 \in B_1(y_{1s}, \ell_1) \quad \text{and} \quad y_2 = 0 \ , \\ \ell_1, & \text{if} \quad y_1 \geq y_{1s} + \ell_1 \quad \text{and} \quad y_2 = 0 \ , \\ 0, & \text{otherwise} \ , \end{cases}$$

raises marginal tax rates on the earnings of spouse 1 conditional on $y \in B_1(y_{1s}, \ell_1)$ and spouse 2 having no earnings, $y_2 = 0$.

Proposition 8. *Consider an increase of the marginal tax rates on the earnings of spouse 2 conditional on $y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2)$. For the limit case $\ell_1 \to 0$ and $\ell_2 \to 0$, the revenue effect of this reform is given by*

$$\mathcal{R}(y_{2s} \mid y_{1s}) \quad = \quad f_{y_1}(y_{1s}) \left(\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \left(1 - \mathcal{E}^d_x(y_{2s} \mid y_{1s}) \right) \right) \ ,$$

where $y_1 \mapsto f^{y_1}(y_1)$ is the density associated with the marginal distribution of y_1 and $y_2 \mapsto F^{y_2}(y_{2s} \mid y_{1s})$ is the cdf of y_2, conditional on y_1 being equal to y_{1s}.

The proof of Proposition 8 corresponds to the proof of Proposition 1 in Bierbrauer et al. (2024).

The revenue function simplifies further under an additional assumption: When the tax system is piecewise linear, effort-cost functions are iso-elastic

$$k_2(y_2, \omega_2) \quad = \quad \frac{1}{1 + \frac{1}{\varepsilon_2}} \left(\frac{y_2}{\omega_2} \right)^{1 + \frac{1}{\varepsilon_2}} \ ,$$

and earnings choices are characterized by first order conditions, then

$$\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) \quad = \quad -f_{y_2}(y_{2s} \mid y_{1s}) \, \frac{T^0_{y_2}(y_{1s}, y_{2s})}{1 - T^0_{y_2}(y_{1s}, y_{2s})} \, y_{2s} \, \varepsilon_2 \ . \tag{C.5}$$

Rewriting the equation in Proposition 8, we get

$$\mathcal{R}(y_{2s} \mid y_{1s}) \quad = \quad f_{y_1}(y_{1s}) \left(\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \right)$$

$$- f_{y_1}(y_{1s}) \, (1 - F_{y_2}(y_{2s} \mid y_{1s})) \, \mathcal{E}^d_x(y_{2s} \mid y_{1s}) \ .$$

$$\mathcal{R}(y_{2s}) \quad = \int_{y_{1s}} \mathcal{R}(y_{2s} \mid y_{1s}) dy_{1s}$$

$$= \underbrace{\int_{y_{1s}} f_{y_1}(y_{1s}) \left(\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \right) dy_{1s}}_{:=\mathbf{A}}$$

$$\underbrace{- \int_{y_{1s}} f_{y_1}(y_{1s}) \left(1 - F_{y_2}(y_{2s} \mid y_{1s})\right) \mathcal{E}^d_x(y_{2s} \mid y_{1s}) dy_{1s}}_{:=\mathbf{B}} \quad .$$

Rewriting A.

$$\mathbf{A} = \int_{y_{1s}} f_{y_1}(y_{1s}) \left(\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \right) dy_{1s}$$

$$= \mathbf{E}_{y_{1s}} \left[\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \right]$$

$$= \mathbf{E}_{y_{1s}} \left[\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) \right] + 1 - F_{y_2}(y_{2s}) \quad .$$

Using Equation (C.5), we get

$$\mathbf{A} = \mathbf{E}_{y_{1s}} \left[-f_{y_2}(y_{2s} \mid y_{1s}) \, \frac{T^0_{y_2}(y_{1s}, y_{2s})}{1 - T^0_{y_2}(y_{1s}, y_{2s})} \, y_{2s} \, \varepsilon_2 \right] + 1 - F_{y_2}(y_{2s})$$

$$= -y_{2s} \, \varepsilon_2 \mathbf{E}_{y_{1s}} \left[f_{y_2}(y_{2s}) \frac{f_{y_2}(y_{2s} \mid y_{1s})}{f_{y_2}(y_{2s})} \, \frac{T^0_{y_2}(y_{1s}, y_{2s})}{1 - T^0_{y_2}(y_{1s}, y_{2s})} \right] + 1 - F_{y_2}(y_{2s})$$

$$= -\varepsilon_2 y_{2s} \, f_{y_2}(y_{2s}) \mathbf{E}_{y_{1s}} \left[\frac{T^0_{y_2}(y_{1s}, y_{2s})}{1 - T^0_{y_2}(y_{1s}, y_{2s})} \mid y_2 = y_{2s} \right] + 1 - F_{y_2}(y_{2s}) \quad .$$

When the status quo has joint taxation, $T^0_{y_2}(y_{1s}, y_{2s}) = T'_m(y_{1s} + y_{2s})$, hence

$$\mathbf{A} = -\varepsilon_2 y_{2s} \, f_{y_2}(y_{2s}) \mathbf{E}_{y_{1s}} \left[\frac{T'_m(\cdot)}{1 - T'_m(\cdot)} \mid y_2 = y_{2s} \right] + 1 - F_{y_2}(y_{2s}) \quad .$$

Rewriting of B.

$$\mathbf{B} = - \int_{y_{1s}} f_{y_1}(y_{1s}) \left(1 - F_{y_2}(y_{2s} \mid y_{1s})\right) \mathcal{E}^d_x(y_{2s} \mid y_{1s}) dy_{1s}$$

$$-\mathbf{B} = F_{y_1}(0)\left(1 - F_{y_2}(y_{2s} \mid 0)\right)\mathcal{E}_x^d(y_{2s} \mid 0) + \mathbf{E}_{y_{1s}}\left[\left(1 - F_{y_2}(y_{2s} \mid y_{1s})\right)\mathcal{E}_x^d(y_{2s} \mid y_{1s}) \mid y_{1s} > 0\right]$$

We can rewrite the first part of this expression as

$$F_{y_1}(0)\left(1 - F_{y_2}(y_{2s} \mid 0)\right)\mathcal{E}_x^d(y_{2s} \mid 0) = F_{y_1}(0)\int_{y_{2s}}^{\bar{y_2}} f y_2(y_{2s} \mid 0)\mathcal{E}_x^d(y_{2s} \mid 0)dy_{2s}$$

where $\mathcal{E}_x^d(y_{2s} \mid 0)$ is an average elasticity capturing an extensive margin response when for all couples with $y_2 \geq y_{2s}$ and $y_1 = 0$ the tax burden increases by a marginal unit. Up writing,

$$\mathcal{E}_x^d(y_{2s} \mid 0) = \frac{T^0(0, y_{2s})}{y_{2s} - T^0(0, y_{2s})}\bar{\pi}_{sec}(y_{2s})$$

we obtain

$$F_{y_1}(0)\left(1 - F_{y_2}(y_{2s} \mid 0)\right)\mathcal{E}_x^d(y_{2s} \mid 0) = F_{y_1}(0)\int_{y_{2s}}^{\bar{y_2}} \frac{T^0(0, y_{2s})}{y_{2s} - T^0(0, y_{2s})}\bar{\pi}_{sec}(y_{2s}) f y_2(y_{2s} \mid 0)dy_{2s},$$

where $\bar{\pi}_{sec}(y_{2s})$ is now interpreted as an elasticity with respect to the participation tax.

Similarly, we rewrite the second part of expression $-\mathbf{B}$ as

$$\mathbf{E}_{y_{1s}}\left[\left(1 - F_{y_2}(y_{2s} \mid y_{1s})\right)\mathcal{E}_x^d(y_{2s} \mid y_{1s}) \mid y_{1s} > 0\right]$$

$$= \left(1 - F_{y_1}(0)\right)\mathbf{E}_{y_{1s}}\left[\int_{y_{2s}}^{\bar{y_2}} \mathcal{E}_x^d(y_{2s} \mid y_{1s}) f(y_{2s} \mid y_{1s})dy_{2s} \mid y_{1s} > 0\right]$$

$$= \left(1 - F_{y_1}(0)\right)\mathbf{E}_{y_{1s}}\left[\int_{y_{2s}}^{\bar{y_2}} \frac{T^0(y_{1s}, y_{2s})}{y_{1s} + y_{2s} - T^0(y_{1s}, y_{2s})}\bar{\pi}_{dec}(y_{2s} \mid y_{1s}) f y_2(y_{2s} \mid y_{1s})dy_{2s} \mid y_{1s} > 0\right].$$

Summing up the terms **A** and **B** gives the expression of $\mathcal{R}(y_{2s})$ in Proposition 4.

Appendix C.2.5 provides insights on how revenue functions are estimated in the data.

C.1.3 Evaluating "Large" Reforms

Recall that, for a given tax reform (τ, h), we defined $V_s(\tau, h, \rho_s, \theta_s) - V_s(0, h, \rho_s, \theta_s)$ as the reform-induced change in indirect utility for a single with characteristics θ_s. Analogously, $V_i(\tau, h, \rho_m, \theta_m, \gamma_m) - V_i(0, h, \rho_m, \theta_m, \gamma_m)$ is the reform-induced change in indirect utility for spouse i in a married couple with characteristics θ_m and intra-family bargaining weights γ_m. Equations (3.3) and (3.4) above characterize the derivatives of these expressions with

respect to the reform intensity τ and evaluate them at the status quo, i.e. at $\tau = 0$. We now generalize this and consider the effects of a change of the reform intensity also away from the status quo. Specifically, we denote the marginal effect of a further increase of the reform intensity – starting from intensity τ' – on the indirect utility of spouse i by $V_{i,\tau}(\tau', h, \rho_m, \theta_m, \gamma_m)$. We define $V_{s,\tau}(\tau', h, \rho_s, \theta_s)$ analogously. Obviously,

$$V_s(\tau, h, \rho_s, \theta_s) - V_s(0, h, \rho_s, \theta_s) = \int_0^\tau V_{s,\tau}(\tau', h, \rho_s, \theta_s)\, d\tau' \,, \tag{C.6}$$

and

$$V_i(\tau, h, \rho_m, \theta_m, \gamma_m) - V_i(0, h, \rho_m, \theta_m, \gamma_m) = \int_0^\tau V_{i,\tau}(\tau', h, \rho_m, \theta_m, \gamma_m)\, d\tau' \,. \tag{C.7}$$

By the envelope theorem,

$$V_{s,\tau}(\tau', h, \rho_s, \theta_s) = u_{s1}(\tau', \theta_s)\Big[\rho_s R_\tau(\tau', h) - h_s(y_s^*(\tau', \theta_s))\Big] \,, \tag{C.8}$$

where $u_{s1}(\tau', \theta_s)$ is the marginal utility of consumption evaluated at reform intensity τ', $R_\tau(\tau', h)$ is the derivative of aggregate tax revenue with respect to further increases of the reform intensity at τ', and finally, $y_s^*(\tau', \theta_s)$ is the utility maximizing earnings level of a type θ_s single when the reform intensity equals τ'. Analogously, we obtain

$$\begin{aligned} \tfrac{\partial}{\partial \tau} V_i(\tau', h, \rho_m, \theta_m, \gamma_m) \;&=\; u_{i1}^0(\tau', \theta_m, \gamma_m)\alpha_{i1}^0(\tau', \theta_m, \gamma_m) \times \\ &\quad \Big[\rho_m R_\tau(\tau', h) - h_m(y_m^*(\tau', \theta_m, \gamma_m))\Big] \,. \end{aligned} \tag{C.9}$$

We now impose the simplifying assumptions that preferences are quasi-linear in consumption and that household consumption is a public good. Equations (C.8) and (C.9) then become

$$\frac{\partial}{\partial \tau} V_s(0, h, \rho_s, \theta_s) = \rho_s R_\tau(\tau', h) - h_s(y_s^*(\tau', \theta_s)) \,, \tag{C.10}$$

and

$$\frac{\partial}{\partial \tau} V_i(0, h, \rho_m, \theta_m, \gamma_m) = \rho_m R_\tau(\tau', h) - h_m(y_m^*(\tau', \theta_m, \gamma_m)) \,. \tag{C.11}$$

We impose a further assumption, namely that tax revenue is rebated lump-sum at the tax unit level. This implies that $\rho_s = \rho_m = \frac{1}{\nu_s + \nu_m} = 1$. Thus,

$$\rho_s R_\tau(\tau', h) = \rho_m R_\tau(\tau', h) = R_\tau(\tau', h) \,. \tag{C.12}$$

With this assumption, heterogeneity in preferences over tax reforms is then entirely due to heterogeneity in the change of individual tax burdens. We view this as a natural benchmark.[1] Together Equations (C.6), (C.7), (C.10), (C.11) and (C.12) imply that

$$V_s(\tau, h, \rho_s, \theta_s) - V_s(0, h, \rho_s, \theta_s) = \Delta R - \int_0^\tau h_s(y_s^*(\tau', \theta_s))\, d\tau' \,, \tag{C.13}$$

and

$$V_i(\tau, h, \rho_m, \theta_m, \gamma_m) - V_i(0, h, \rho_m, \theta_m, \gamma_m) = \Delta R - \int_0^\tau h_m(y_m^*(\tau', \theta_m, \gamma_m))\, d\tau' \,, \tag{C.14}$$

where $\Delta R := R(\tau, h) - R(0, h)$. Finally, to estimate

$$\int_0^\tau h_s(y_s^*(\tau', \theta_s))\, d\tau' \quad \text{and} \quad \int_0^\tau h_m(y_m^*(\tau', \theta_m, \gamma_m))\, d\tau'$$

we impose following assumptions:

Assumption 3. *The functions h_s and h_m are monotonic functions of income.*

Assumption 4. *The functions y_s^* and y_m^* are monotonic functions of τ.*

Assumption 3 holds provided that tax reforms are monotonic in the sense that the changes of the tax burdens of singles and couples,

$$\tau_s h_s(y_s) = T_{s1}(y_s) - T_{s0}(y_s) \quad \text{and} \quad \tau_m h_m(y_m) = T_{s1}(y_m) - T_{s0}(y_m)\,,$$

are monotonic functions of y_s and y_m, respectively. This property is satisfied by most tax reforms (see Figure C.23 in the Appendix and the discussion in Bierbrauer et al. (2021)). Assumption 4 postulates that behavioral responses are monotonic in the intensity of reforms. Intuitively, if the gap between the new and the old schedule becomes larger, the behavioral adjustment does not become smaller. Under these assumptions, one can show that

$$\int_0^\tau h_s(y_s^*(\tau', \theta_s)) \quad \in \quad [\Delta T_s(y_s^1(\theta_s)), \Delta T_s(y_s^0(\theta_s))]\,, \tag{C.15}$$

[1] A conceivable alternative would be to assume that tax revenue is rebated lump-sum at the individual level, so that

$$\rho_s = \frac{\nu_s}{\nu_s + 2\nu_m} \quad \text{and} \quad \rho_m = \frac{2\nu_m}{\nu_s + 2\nu_m}\,.$$

When household consumption is treated as a public good, this amounts to the assumption that a married individual benefits from an increase of tax revenue twice as much as a single, i.e., both the "own" transfer and the spouse's transfer are sources of utility. When we assume that tax revenue is rebated lump sum at the tax unit level, this effectively amounts to the assumption that all individuals value additional tax revenue in the same way; that is, we suppress heterogeneity in preferences for the level and the composition of public expenditures.

where $y_s^1(\theta_s)$ is the post-reform income of type θ_s, $y_s^0(\theta_s)$ is the pre-reform income and, for any y_s,

$$\Delta T_s(y_s) = T_1(y_s) - T_0(y_s)$$

is the mechanical change in the tax burden. Likewise,

$$\int_0^\tau h_m(y_m^*(\tau', \theta_m, \gamma_m))\, d\tau' \quad \in \quad [\Delta T_m(y_m^1(\theta_m, \gamma_m)), \Delta T_m(y_m^0(\theta_m, \gamma_m))] \,. \tag{C.16}$$

Thus, the impact of the tax reform on individual welfare has an upper bound and a lower bound, with one bound being the mechanical change of the tax burden holding income fixed at the pre-reform level and one bound being the mechanical effect holding income fixed at the post-reform level.

Using (C.13) and (C.15) we say that a type θ_s single benefits from a tax reform if

$$\Delta R \;-\; \max\{\Delta T_s(y_s^1(\theta_s)), \Delta T_s(y_s^0(\theta_s))\} \quad \geq \quad 0 \,, \tag{C.17}$$

and loses if

$$\Delta R \;-\; \min\{\Delta T_s(y_s^1(\theta_s)), \Delta T_s(y_s^0(\theta_s))\} \quad \leq \quad 0 \,, \tag{C.18}$$

According to (C.17), a single is identified as a beneficiary of a tax reform if the change in tax revenue outweighs two measures of how the reform affects the single's tax burden: one is the change of the tax burden holding income fixed at the pre-reform level, the other is the change of the tax burden holding income fixed at the pre-reform level. Note that the two measures coincide when there are no behavioral responses to the tax reform. We identify a single as a loser of a tax reform is both these measures exceed the reform's revenue implications. If neither (C.17) nor (C.18) holds, our approach leaves open whether a type θ_s single is a reform beneficiary or a reform loser.

Analogously, we say that the spouses in a couple are reform beneficiaries if

$$\Delta R \;-\; \max\{\Delta T_m(y_m^1(\theta_m, \gamma_m)), \Delta T_m(y_m^0(\theta_m, \gamma_m))\} \quad \geq \quad 0 \,, \tag{C.19}$$

and reform losers if

$$\Delta R \;-\; \min\{\Delta T_m(y_m^1(\theta_m, \gamma_m)), \Delta T_m(y_m^0(\theta_m, \gamma_m))\} \quad \leq \quad 0 \,. \tag{C.20}$$

In Appendix C.2.4, we explain of how we make use of the TAXSIM microsimulation model to obtain, for every tax unit, an estimate of all the terms that enter in the left-hand sides

of inequalities (C.17) - (C.20). This enables us to tell for every such tax unit whether or not its members benefited from the reform. This is used for our analysis of political feasibility in Section 3.4 which rests on a comparison of the number of individuals that benefited from a reform to the number of individuals that were made worse off.

For the welfare analysis in Section C.3.3 we aggregate the gains of reform winners and the losses of reform losers using various social welfare functions. We use these social welfare functions as descriptive tools. For instance, an evaluation with a Rawlsian social welfare function will tell us whether or not "the poor" benefited from a reform. An evaluation with an "Affirmative Feminist" welfare function will tell us whether or not working women benefited from a reform. Table B.5 contains a description of all the welfare functions that we use in our analysis. Any such social welfare function involves the computation of a weighed average of individual welfare gains and losses. In the main text, we use

$$\Delta R \quad - \quad \max\{\Delta T_s(y_s^1(\theta_s)), \Delta T_s(y_s^0(\theta_s))\}$$

and

$$\Delta R \quad - \quad \max\{\Delta T_m(y_m^1(\theta_m, \gamma_m)), \Delta T_m(y_m^0(\theta_m, \gamma_m))\} \ .$$

as measures of individual welfare changes. Thus, we are using money-metric welfare functions with a "conservative" estimate of welfare gains, as our measure is lower bound on the welfare gains that reform beneficiaries actually realize.

A conceivable alternative is to have a "conservative" estimate of welfare losses based on

$$\Delta R \quad - \quad \min\{\Delta T_s(y_s^1(\theta_s)), \Delta T_s(y_s^0(\theta_s))\}$$

and

$$\Delta R \quad - \quad \min\{\Delta T_m(y_m^1(\theta_m, \gamma_m)), \Delta T_m(y_m^0(\theta_m, \gamma_m))\} \ .$$

As we show in the Appendix, which of these two approaches is taken is without consequence for our conclusions (see in particular Figure C.30).

C.2 APPENDIX: FROM THEORY TO EMPIRICS

C.2.1 Data

The Current Population Survey (CPS) is conducted by the US Census Bureau and the Bureau of Labor Statistics and contains nationally representative cross-sectional survey data from 1962 onward. We use data from the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC).[2] The sample size of CPS-ASEC increased from around 30,000 households in 1962 to more than 90,000 in the most recent wave. In contrast to tax return micro data such as the public use files (IRS-SOI PUF) from the Statistics of Income (SOI) division of the Internal Revenue Service (IRS), as, e.g., used by Bargain et al. (2015) or Bierbrauer et al. (2021), the CPS data contain exact information about the incomes of primary and secondary earners of the tax unit.[3]

To adapt the CPS to the input requirements of the microsimulation model, we transform the CPS from a household-level data set to a tax unit level data set. For this purpose, we form tax units by joining all married spouses with their dependent children. Single individuals and unmarried spouses form separate tax units. Children of single individuals are in most cases allocated to the household head. Adult individuals with a total income below the year-specific personal exemption threshold are assumed to reflect dependents of the household head. Table C.1 illustrates in detail the correspondence between variables utilized in NBER TAXSIM and variables in the CPS data.

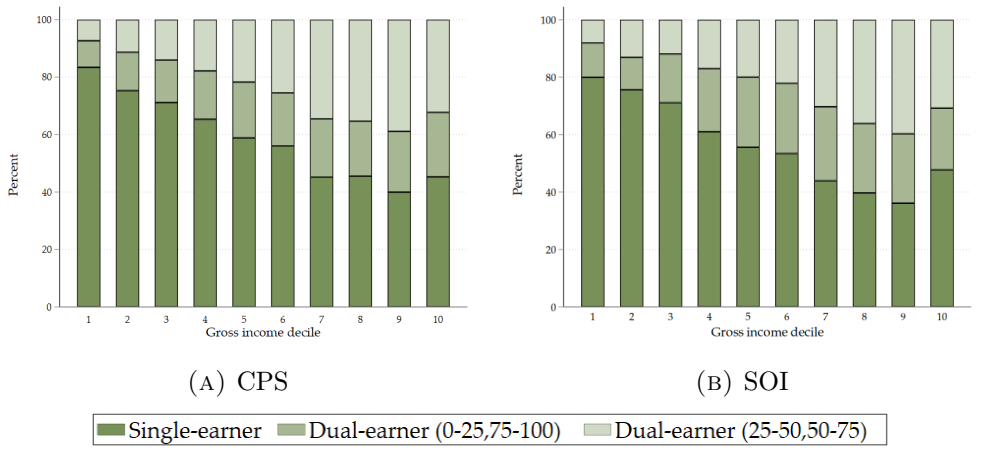
Treatment of Top Incomes In the CPS data, information on top incomes is limited by (i) public topcoding, and (ii) internal censoring. We address both limitations by harmonizing the treatment of top incomes across the different survey years and by following Piketty and Saez (2003) and Piketty (2003) in assuming that top incomes are well represented by a Pareto distribution.

In a first step, we address the challenge that public topcoding methods vary over time. In most recent years (since 2011), the Census Bureau uses a rank proximity swapping procedure to preserve the privacy for top income earners while maintaining the internal

[2]See Flood et al. (2021) and https://cps.ipums.org for a detailed description of CPS data.

[3]In the IRS-SOI PUF, the relevant information on salaries and wages from the W2-form of the primary and secondary earner is only available for the year 1974 and imputed for all other years using an undocumented procedure. For 1974, in which reliable information is available, the distribution of different couple types across per capita income distribution is very similar to the CPS data (see Figure C.1). Moreover, Bargain et al. (2015) compare inequality measures as well as the direct policy effect, ΔT, based on CPS and SOI-PUF data and show that results are very similar (except for the very top of the distribution).

Figure C.1. – Comparison of CPS and SOI Data (1974), Couple Types



Note: This figure displays for the tax year of 1974 the distribution of married couple types across deciles of the per capita income distribution. The figure compares the distribution based on the CPS data (Figure C.1a) to the distribution based on the IRS-SOI PUF tax return micro data (Figure C.1b). All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on CPS-ASEC and SOI PUF.

distribution of top incomes. In this procedure, values at or above a specific swap threshold are switched against other top income values within a bounded interval. For previous years, however, the CPS data originally contains top income values that are based on different procedures, in particular traditional topcoding (1962-1995), and a replacement value system procedure (1996-2010). To be able to consistently analyze the effect of tax reforms over the full time horizon, we apply the most recent method of rank proximity swapping also to previous years using supplementary files provided by IPUMS.[4] Thereby, we preserve the internally used distribution of top incomes whenever possible.

In a second step, we address the challenge that top incomes are also internally censored based on the value range limits of the income variables. As shown by Larrimore et al. (2008), since these censoring thresholds have changed discretely at specific points in time, the share of individuals affected by censoring varies and can reach up to one percent in specific years. To address the unequal representation of censored incomes, we replace censored incomes by random draws from a Pareto distribution. In particular, we first identify for every year and every income type the highest possible income T assigned in a given year. Based on this censoring threshold, we generate for every year and every income type the parameter α of a Pareto distribution with density $f(Y) = \alpha * T^{\alpha} * Y^{-\alpha-1}$.

[4]For details on the treatment of top incomes in general and the data used for rank proximity swapping, see https://cps.ipums.org/cps/topcodes_tables.shtml and https://cps.ipums.org/cps/income_cell_means.shtml.

We thereby assume that incomes above the 99th percentile follow a Pareto distribution and thus estimate the shape parameter α as

$$\alpha = \frac{ln\left(\frac{N_{Y \geq p99}}{N_{Y=T}}\right)}{ln\left(\frac{Y_T}{Y_{p99}}\right)}$$

where $N_{Y \geq p99}$ is the number of individuals with an income above the 99th percentile of the income distribution, $N_{Y=T}$ is the number of individuals at the highest income, and Y_T and Y_{p99} are the top income and the income at the 99th percentile respectively.[5] Finally, we use the distribution to replace the top incomes T by random draws from this calibrated distribution.[6]

Sample Restrictions We are mainly interested in the differences between married couples and single individuals. We thereby assume that married couples always file jointly. While married couples can also file separately, this filing status is usually not beneficial (see Figure C.3) and is chosen by less than two percent of all tax units (see Figure C.2).[7] Similarly, we abstract from the qualifying widow(er) filing status that gives widowed individuals a preferential tax treatment in the two years following the spouses' death. Given our sample restriction, the occurrence of widow(er)s is negligible (see also Figure C.2). If not indicated otherwise, we restrict the sample to tax units in which primary and secondary taxpayer are between 25 and 55 years old and have non-negative gross income. This sample restriction is guided by (i) our model that considers neither education nor retirement decisions, and (ii) the assumptions on labor supply responses to taxation that are not valid for young and old people with weak labor force attachment. In Section C.5 we replicate main results for an alternative sample, i.e., the full adult population.

Throughout the analysis, we calculate tax payments as well as average and marginal tax rates based on the federal income tax and abstract from state income tax and social security payroll taxes. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income.

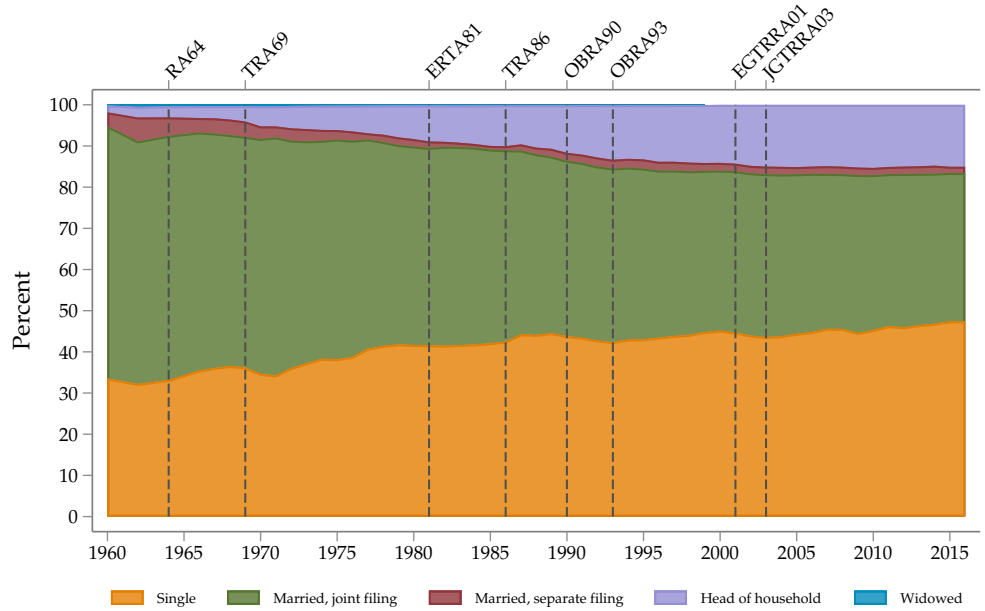
---

[5]Discussions of different estimation methods for the shape parameter of the Pareto distribution can be found in Armour et al. (2016) and Blanchet et al. (2018).

[6]To reduce the impact of random sampling on our results, we use quantiles of the distribution. The number of quantiles utilized depends on the number of individuals at the top income. For instance, if we observe 25 individuals at the top income, we assign these individuals income levels that correspond to the 25 quantiles of the randomly drawn values from the calibrated Pareto distribution. Thereby, we preserve the information of the distribution while limiting the influence of random draws.

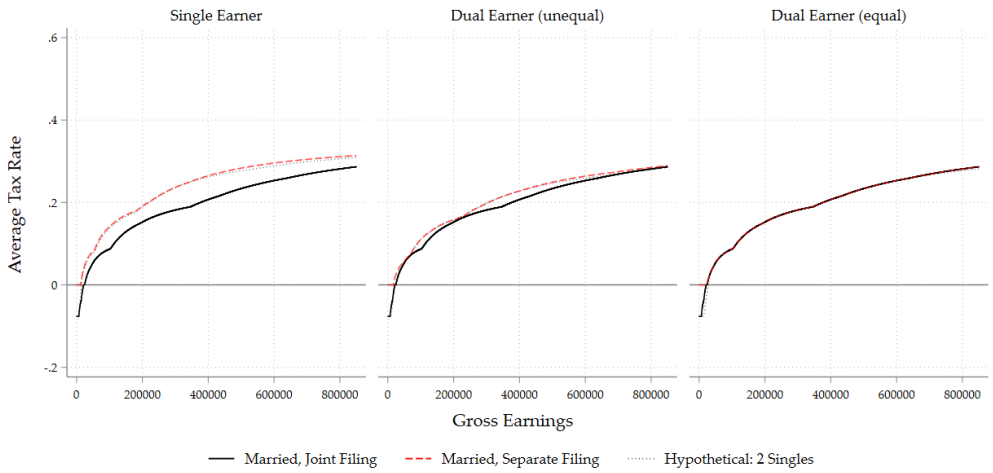
[7]Filing separately can be beneficial in very particular circumstances that we do not observe, i.e., in the case of substantial itemizable deductions (e.g. high medical expenses or student loan repayments).

Figure C.2. – Filing Status According to SOI Data



Note: This figure shows the distribution of filing status from 1960 to 2016. Filing statuses are based on the IRS-SOI PUF administrative tax return micro data.
Source: Own calculations based on SOI PUF.

Figure C.3. – Married Couples Filing Jointly and Separately (2019)



Note: This figure shows how the average tax rate of a couple with specific gross earnings differs between whether this couple files separately or jointly. In addition, the figure also shows the average tax rate of two singles with the same joint income. The figure differentiates further by the type of couple: single earner couples (95% / 5%), unequal dual earner couples (75% / 25%) and dual earner couples with equal incomes (50% / 50%).
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Table C.1. – TAXSIM Variables and CPS Application

TAXSIM	Explanation	CPS Application
taxsimid	Case ID	N/A
year	Tax year	ASEC income reference year
state	State	State of residence
mstat	Marital Status	Marital status (married vs. unmarried)
page	Age of primary taxpayer	Age of husband
sage	Age of spouse	Age of spouse
depx	Number of dependents	Number of children below and of age 18 + additional dependents
dep13	Number of children under 13	Number of children under 13
dep17	Number of children under 17	Number of children under 17
dep18	Number of qualifying children for EITC.	Number of children below and of age 18
pwages	Wage and salary income of Primary Taxpayer	Wage income + business income + farm income of husband
swages	Wage and salary income of Spouse	Wage income + business income + farm income of spouse
dividends	Dividend income	Income from dividends
intrec	Interest Received	Income from interest
stcg	Short Term Capital Gains or losses	N/A
ltcg	Long Term Capital Gains or losses.	Capital gains - capital losses
otherprop	Other property income	Income from rent
nonprop	Other non-property income	Income from other Source not specified + income from alimony
pensions	Taxable Pensions and IRA distributions	Retirement income
gssi	Gross Social Security Benefits	Social Security income
ui	Unemployment compensation received	Income from unemployment benefits
transfers	Other non-taxable transfer Income	Welfare (public assistance) income + income from worker's compensation + income from veteran benefits + income from survivor benefits + income from disability benefits + income from child support + income from educational assistance + income from SSI + income from assistance
rentpaid	Rent Paid	N/A
proptax	Real Estate taxes paid	Annual property taxes
otheritem	Other Itemized deductions	Indirect calculation via difference between adjusted gross income and taxable income calculated by the Census Bureau's tax model.
childcare	Child care expenses	N/A
mortgage	Deductions not included in otheritem	N/A

Note: This table displays the variables utilized as part of the tax calculation via the NBER TAXSIM (v32) microsimulation model and the corresponding information from the CPS used for the respective variables. For details on TAXSIM (v32) see Feenberg and Coutts (1993) and https://users.nber.org/~taxsim/.
Source: NBER TAXSIM and CPS-ASEC.

C.2.2 Tax Systems and Tax Reforms

In this section, we provide (i) a brief overview of the tax treatment of couples around the world, (ii) outline the broad changes in the tax treatment of couples and singles in the US federal income tax system (see also the review in Borella et al. (2023b)), and (iii) describe the main aspects of specific US tax reforms that we analyze. In the Appendix C.6 we also discuss narratives about a selection of tax reforms using textual analyses.

Tax Treatment of Couples around the World. The tax treatment of couples and singles around the world can be mainly differentiated by the tax unit, to which the tax code in the respective countries applies to. As shown in Table C.2, a large majority of countries treats the individual as the relevant tax unit, while only few nowadays either treat the household as the tax unit or allow for a choice between individual and household level taxation. This has not always been the case, as many countries that feature nowadays individual-level or optional tax treatment previously had household level taxation. Some of these changes have been analyzed, among others, for Canada, the Czech Republic, Spain, Sweden, and the United Kingdom (see Crossley and Jeon (2007), Kalíšková (2014), Fuenmayor et al. (2018), Selin (2014), and Stephens and Ward-Batts (2004)). For the United States, LaLumia (2008) analyzes a reform in 1948 that introduced joint taxation across the United States, but affected only a subset of states (see also discussions below).

It is also important to note that the tax unit type is not sufficient to evaluate the tax treatment of couples, because even though the tax unit might be the individual, there are often particular rules in place, that account for the presence of spouses in the household like spousal allowances or spousal tax credits.

Tax Treatment of Singles and Couples in the US. The federal income tax code in the United States consists of a tax schedule which is differentiated by four filing statuses referring to (i) married individuals filing jointly,[8] (ii) heads of households,[9] (iii) unmarried individuals, and (iv) married individuals filing separately. While the objective of this differentiation is to balance out conflicting goals (tax progressivity, equal treatment of married couples, equal treatment of married and unmarried couples), it results in a complex system of marriage bonuses and penalties across the income distribution.[10]

[8]This filing status also refers to qualifying widow(er)s, i.e. taxpayers whose spouse died during the last two years, who maintains a household with dependent children and who has not remarried.

[9]Unmarried taxpayers who are not a surviving spouse and who maintain a household with dependent persons (e.g. children, father/mother), if a deduction for these persons is possible.

[10]See, for instance, the arguments discussed in a study by the Congressional Budget Office in 1997.

Table C.2. – Tax Treatment of Couples around the World

Country	Tax unit	Particularities		Exemplary reforms
Argentina	individual			
Australia	individual			
Austria	individual	tax credit for spouse	1973	Introduction of individual tax treatment.
Belgium	household			
Brazil	optional			
Canada	individual	tax credit for spouse	1988	Reduction of the "jointness" of the income tax system.
Costa Rica	individual			
Croatia	individual			
Czech Republic	individual		2008	Introduction of individual tax treatment.
Denmark	individual		1970	Introduction of individual tax treatment.
Estonia	individual	flat rate, allowance for spouse		
Finland	individual		1970	Introduction of individual tax treatment.
France	household			
Germany	optional		1958	Introduction of optional individual tax treatment.
Greece	individual			
Hungary	individual	flat rate		
Iceland	household			
Indonesia	household			
Ireland	optional	allowance / tax credit for spouse	2000	Introduction of individual tax treatment.
Israel	individual			
Italy	individual	allowance / tax credit for spouse	1973	Introduction of individual tax treatment.
Japan	individual	allowance for spouse	1950	Introduction of individual tax treatment.
Kenya	individual			
Latvia	individual	allowance for spouse		
Luxembourg	optional	allowance for spouse	2018	Introduction of optional tax treatment.
Mexico	individual			
Malaysia	optional		1977 (1990)	Introduction of (automatic) individual tax treatment.
Montenegro	individual			
Netherlands	optional		1970	Introduction of individual tax treatment.
New Zealand	individual		1973	Introduction of individual tax treatment.
Norway	individual		1970	Introduction of individual tax treatment.
Peru	individual			
Portugal	optional		2015	Introduction of optional tax treatment.
Romania	individual	flat rate		
San Marino	individual			
Slovakia	individual	allowance for spouse		
Slovenia	individual			
South Africa	individual			
South Korea	individual	allowance for spouse	1954	Introduction of individual tax treatment
Spain	optional		1988	Introduction of optional tax treatment.
Sweden	individual		1970	Introduction of individual tax treatment.
Switzerland	household	allowance for spouse	-	
Tunisia	individual			
Turkey	individual	tax credit for spouse		
Ukraine	optional	flat rate		
United Kingdom	individual	allowance for spouse	1990	Introduction of individual tax treatment.
United States	household		see detailed analysis below	

Note: This table provides an overview on the tax treatment of couples in selected countries around the world by displaying information on the relevant tax unit, the progressivity of the tax system, and particularities associated with the tax treatment of couples. In addition, if available, the table displays information about exemplary reforms of the tax treatment of couples in the respective country.
Source: OECD, 2022, PWC Tax Summaries, 2022.

The history of joint taxation in the US can be broadly separated into four periods (see Table C.3). Between 1913 and 1948, the US *formally* had a federal income tax system based on individual income taxation.[11]

With the Revenue Act of 1948, the United States introduced a system of joint taxation, in which the couples' tax liability was calculated by applying the tax schedule to the average income of the couple and by multiplying the resulting tax liability by two. The resulting system resembles very closely the current system of joint taxation in Germany. The system of joint taxation with income splitting was replaced in 1954 with the introduction of two separate tax schedules for couples filing jointly and couples filing separately (also applied to single filers). However, the de-facto treatment of couples stayed the same, because marginal tax rates were not differentiated and all tax brackets for joint filers were set to be twice as large compared to those of separate filers (see Figure C.4).[12]

The de-facto splitting system led to tax liabilities for singles which were up to 42 percent higher compared to couples with the same income level. While some of this marriage bonus was considered to be justified on the basis of different living expenses, the size of the penalty for singles was considered to be too high. Therefore, the Tax Reform Act of 1969 (TRA69) installed a new tax schedule for unmarried persons (not falling under the head of household filing status), which had both lower marginal tax rates and different tax brackets. It was designed specifically to reduce the difference in tax liability between singles and couples with the same income. Both at very low and high incomes, the marriage bonus gradually decreased. Since the tax schedule for couples filing separately was still in place, married couples now faced a higher tax liability than two singles with the same joint income. This was justified on the grounds that even though a married couple plausibly has higher living expenses than a single with the same income and should therefore pay less taxes, the couple might well have less living expenses than *two* singles (unmarried couple) with the same joint income. In this third period, the fixed relationship between bracket lengths across the income distribution was broken (see Figure C.4).

[11]For some states - mostly community property law states - couples' income was assigned in equal terms to both spouses already before 1948. For details, see LaLumia (2008).

[12]In order to give some of the splitting benefits of joint taxation to widows, widowers, and single persons with dependents in their households, the Revenue Act of 1951 introduced a separate tax schedule for heads of households. While this was implemented using a separate tax schedule with both different marginal tax rates and different tax brackets, it was designed to result at any given income level in a tax liability which lies halfway between the tax paid by couples and singles. For details, see General Explanation of the Tax Reform Act of 1969.

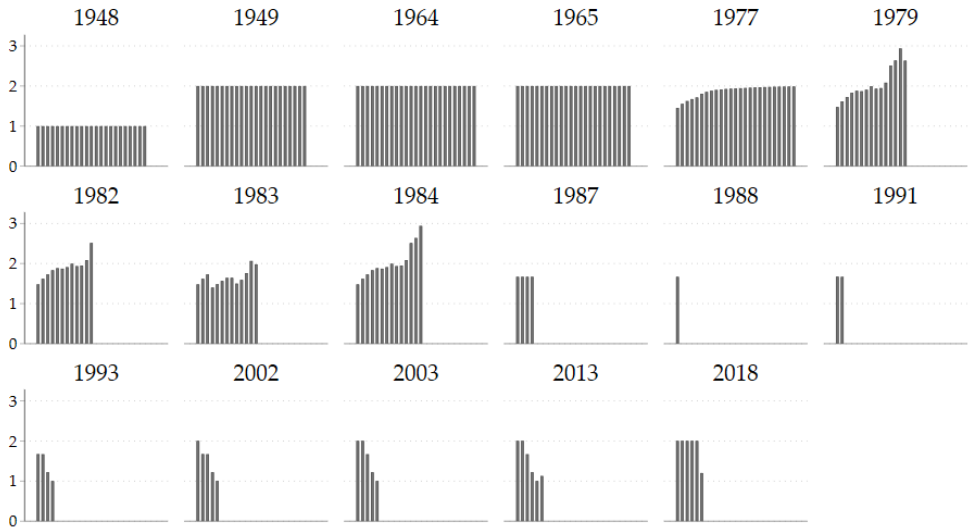
Table C.3. – US Tax Treatment of Singles and Couples

Tax Year	Difference
1913-1948	Income splitting in community law states (Washington, Idaho, Nevada, California, Arizona, New Mexico, Texas, Louisiana), individual taxation in common law states
1949-1970	Income splitting
1971-1986	Difference in tax brackets and differences in marginal tax rates
1987-2020	Only difference in tax brackets, same marginal tax rates

The fourth period started with the Tax Reform Act of 1986 (TRA86) which harmonized marginal tax rates for all filing statuses and only treated them differently with respect to the length of tax brackets. This relationship between brackets across singles and couples has changed significantly over time. From 1987 to 1992, although the number of tax brackets varied, the relationship between upper bracket limits for couples and singles was constant. The Omnibus Budget Reconciliation Act of 1993 (OBRA93) differentiated this relationship with the newly introduced tax brackets and thereby increased the potential for marriage penalties at higher incomes. The Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA01) eliminated the marriage penalty for the lowest bracket but strengthened the potential of marriage penalties at the upper tail of the distribution by the newly introduced upper bracket. Starting with the tax year 2003, the marriage penalty in the second tax bracket (15 percent marginal tax rate) was eliminated. Furthermore, the Tax Cuts and Jobs Act of 2017 (TCJA17) eliminated marriage penalties for brackets 3 to 5.

Major US tax Reforms. We analyze all major changes in the US personal income tax system from 1964 until 2017. Table C.4 provides an overview of the 11 reforms that we identified and analyze. We concentrate on large legislative changes which drive the tax policy effect. These reforms are the Revenue Act of 1964 (RA64), the Tax Reform Act of 1969 (TRA69), the Revenue Act of 1978 (RA78), the Economic Recovery Tax Act of 1981 (ERTA81), the Tax Reform Act of 1986 (TRA86), the Omnibus Budget Reconciliation Act of 1990 and 1993 (OBRA90 and OBRA93), the Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA01), the Jobs and Growth Tax Relief Reconciliation Act of 2003 (JGTRRA03), the American Taxpayer Relief Act of 2012 (ATRA12) and the Tax Cuts and Jobs Act of 2017 (TCJA17).

Figure C.4. – Upper Limit of Tax Bracket, Single/Couple



Note: This figure shows the relation between the upper tax bracket limit of the tax schedule of couples and singles for all tax brackets in the respective tax year. We display those years, in which the number or relation of brackets changed. In all years, the highest bracket is excluded since it has no upper limit.
Source: Historical U.S. Federal Individual Income Tax Rates and Brackets.

Table C.4. – Overview of US Reforms

Tax reform	pre	post	key features of the reform
RA64	1963	1966	Tax cut (top rate from 91% to 70%)
TRA69	1968	1971	Introduction of Alternative Minimum Tax and new tax schedule for single taxpayers
RA78	1978	1979	Widening of tax brackets (and reducing their number)
ERTA81	1980	1984	Tax cut (top rate from 70% to 50%)
TRA86	1985	1988	Broadening of tax base and reductions in MTRs (top rate from 50% to 28%)
OBRA90	1990	1991	Increase of top tax rate from 28% to 31%
OBRA93	1992	1993	Expansion of EITC and increase of top tax rate from 31% to 39.6%
EGTRRA01	2000	2002	Reductions in marginal tax rates
JGTRRA03	2002	2003	Reductions in marginal tax rates
ATRA12	2012	2013	Increase of tax rates for high income earners
TCJA17	2016	2018	Tax cuts (top rate from 39.6% to 37%)

Note: Table C.4 lists the major reforms of the federal income tax in the US after WWII: the Revenue Act of 1964 (RA64), the Tax Reform Act of 1969 (TRA69), the Revenue Act of 1978 (RA78), the Economic Recovery Tax Act of 1981 (ERTA81), the Tax Reform Act of 1986 (TRA86), the Omnibus Budget Reconciliation Act of 1990 and 1993 (OBRA90 and OBRA93), the Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA01), the Jobs and Growth Tax Relief Reconciliation Act of 2003 (JGTRRA03), the American Taxpayer Relief Act of 2012 (ATRA12) and the Tax Cuts and Jobs Act of 2017 (TCJA17). The pre reform year is always the last year before any change was implemented while the post reform year is the one after all changes are phased in. See Appendix H of Bierbrauer et al. (2021) for more details and the distributional implications of these reforms.

C.2.3 Descriptives

Section 3.4 of the main text provides information about demographic trends, the evolution of marriage penalties and bonuses, and shifts in the splitting function. In the following, we provide supplementary results and explanations.

Demographic Trends. Figure 3.1 in the main text shows that there was an expansion of singles and dual-earner couples since the 1960s. This demographic change holds broadly across different US states (see Figure C.5). The expansion of singles has been particularly prominent in the lower part of the income distribution while dual earner couples are particularly relevant at the top of the income distribution (see Figure C.6).

Marriage Penalties and Bonuses

Figure 3.3 in the main text illustrates that since the 1960s, the share of couples experiencing a marriage bonus decreased while the share of couples facing a marriage penalty increased. Figure C.7 shows how that the magnitude of bonuses and penalties has increased over time. Here we briefly explain, how we construct these measures based on information from the TAXSIM microsimulation model. The microsimulation model allows us to compare for every couple in the data its actual tax liability ($T^{act} = T_m(y_1 + y_2)$) in the status quo tax system with a hypothetical tax liability in a situation in which the couple would not be married and thus file as two singles ($T^{hyp} = T_s(y_1) + T_s(y_2)$). In the case of dependent children, we must allocate dependents to either one of the two spouses. The counterfactual tax burden is an average over two hypothetical tax burdens in which dependents are allocated to either one of the spouses. Based on actual and hypothetical tax payments, we can construct

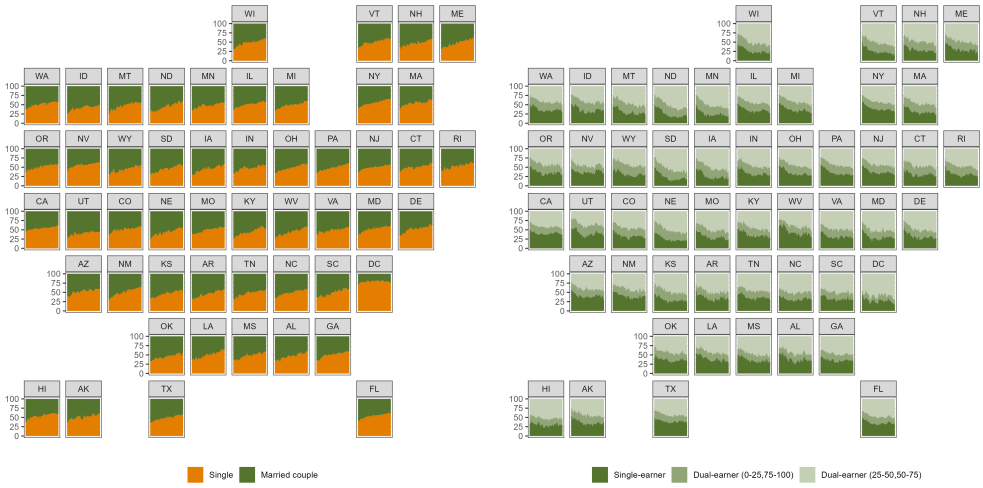
Absolute marriage bonus: $B^{abs} = T^{hyp} - T^{act}$,

Relative marriage bonus: $B^{rel} = \frac{B^{abs}}{y_1+y_2}$.

Empirical Splitting Function σ

The evolution of marriage penalties and bonuses is driven by (i) the change in demographics towards a larger share of dual earner couples and (ii) changes in the relative treatment

Figure C.5. – Demographic Change, by State



(A) Tax Unit Types (B) Couple Types

Note: This figure displays the distribution of tax unit types over time for all states starting in 1976. Figure C.5a displays the share of single tax units (orange area) and the share of couple tax units (green area). Figure C.5b displays the share of single-earner and dual-earner couples. A single-earner couple refers to a married couple, in which one spouse is not employed (dark green area). The figure further displays the share of dual-earner couples in which both spouses are employed and (i) one spouse earns between 0 and 25 percent (mid green area) and (ii) between 25 and 50 percent of total earnings (light green area). Earnings shares are computed on the basis of wage, business and farm income. Reforms of the federal income tax code as described in Table C.4 are displayed as vertical lines. All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on CPS-ASEC.

of couples and singles in the tax system. As discussed in the main text, σ allows us to describe changes in this relative treatment in a constructive manner. We can reformulate Equation (3.1) as an implicit relationship between the average tax rates of couples and singles

$$\bar{\tau}_m(y_m) = \bar{\tau}_s\left(\frac{y_m}{\sigma(y_m)}\right) , \tag{C.21}$$

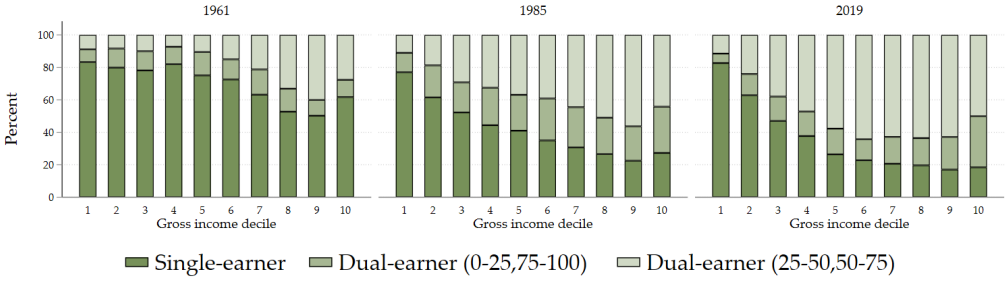
where

$$\bar{\tau}_m(y_m) := \frac{T_m(y_m)}{y_m} \quad \text{and} \quad \bar{\tau}_s\left(\frac{y_m}{\sigma(y_m)}\right) := \frac{\sigma(y_m)}{y_m} T_s\left(\frac{y_m}{\sigma(y_m)}\right) .$$

Figure C.8 illustrates that the expression in terms of average tax rates is instrumental to the empirical estimation of the function $\sigma : y_m \mapsto \sigma(y_m)$. For the empirical estimation of σ, we proceed in two steps. First, since in the baseline we abstract from heterogeneity beyond income and filing status, we estimate the mean average tax rate for couples and singles at every income level z:

Figure C.6. – Demographic Change, by Income Decile

(A) Tax Unit Types



(B) Couple Types

Note: This figure displays the distribution of tax unit types by deciles of the per-capita gross income distribution. Figure C.6a displays the share of single tax units and the share of couple tax units. Figure C.6b displays the share of single-earner and dual-earner couples. A single-earner couple refers to a married couple, in which earnings of one spouse are zero. The figure further displays the share of dual-earner couples, in which one spouse earns between 0 and 25 percent (25 and 50 percent) of total earnings. Earnings shares are computed on the basis of wage, business and farm income. Reforms of the federal income tax code are displayed as vertical lines. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on CPS-ASEC.

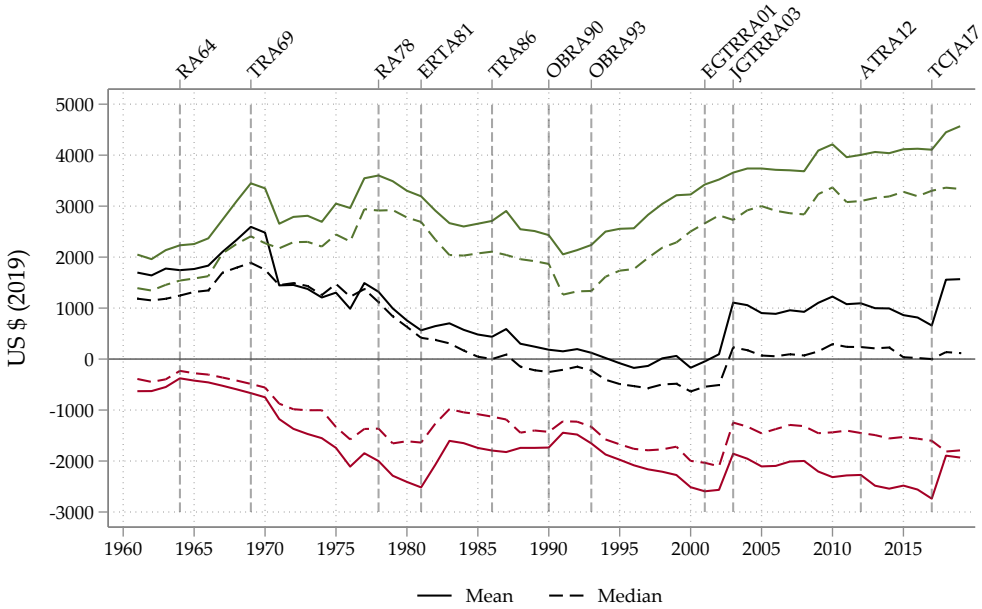
$$\bar{\tau}_i^*(z) = \mathbb{E}\left[\bar{\tau}_i(z,.)\right] \quad , \quad i \in \{s,m\} \; . \tag{C.22}$$

This is reported in Figure C.9. Second, we solve the following equation for $\sigma(z)$:

$$\bar{\tau}_m^*(z) = \bar{\tau}_s^*\left(\frac{z}{\sigma(z)}\right) \; . \tag{C.23}$$

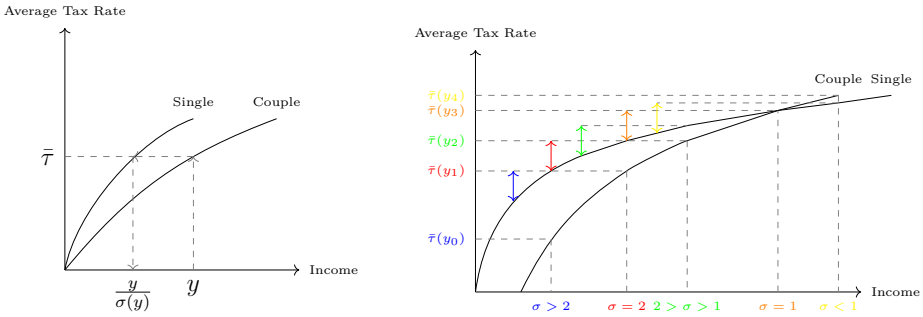
Figure C.10 provides evidence on the heterogeneity of σ when we do not consider the mean average tax rate of all singles and all couples jointly but make separate comparisons of σ by conditioning on the number of children in a household.

Figure C.7. – Absolute Marriage Bonuses and Penalties over Time



Note: This figure shows how the magnitude of the absolute marriage bonus (penalty) B^{abs} changed over time. Mean bonuses / penalties are CPI-adjusted. Marriage bonuses and penalties have been constructed by estimating for every married couple a counterfactual tax burden of two singles with their respective individual incomes. The counterfactual tax burden is an average over two tax burdens that allocate dependent to either spouse. Absolute marriage bonuses are CPI-adjusted and measured in 2019 US dollars. Reforms of the federal income tax code as described in Table C.4 are displayed as vertical lines. All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

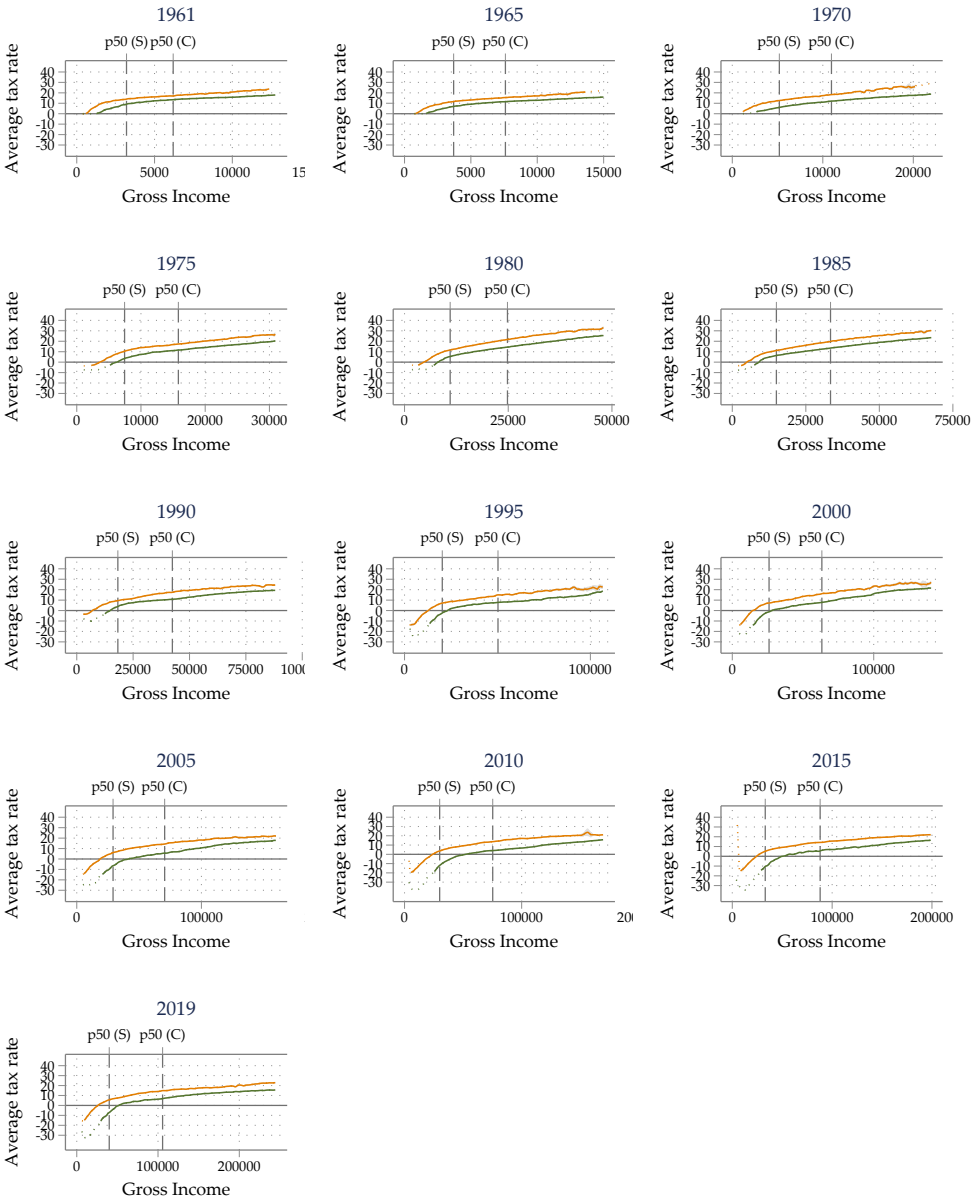
Figure C.8. – Illustration of σ-Function



(A) σ-function (B) Marriage Penalties and Bonuses

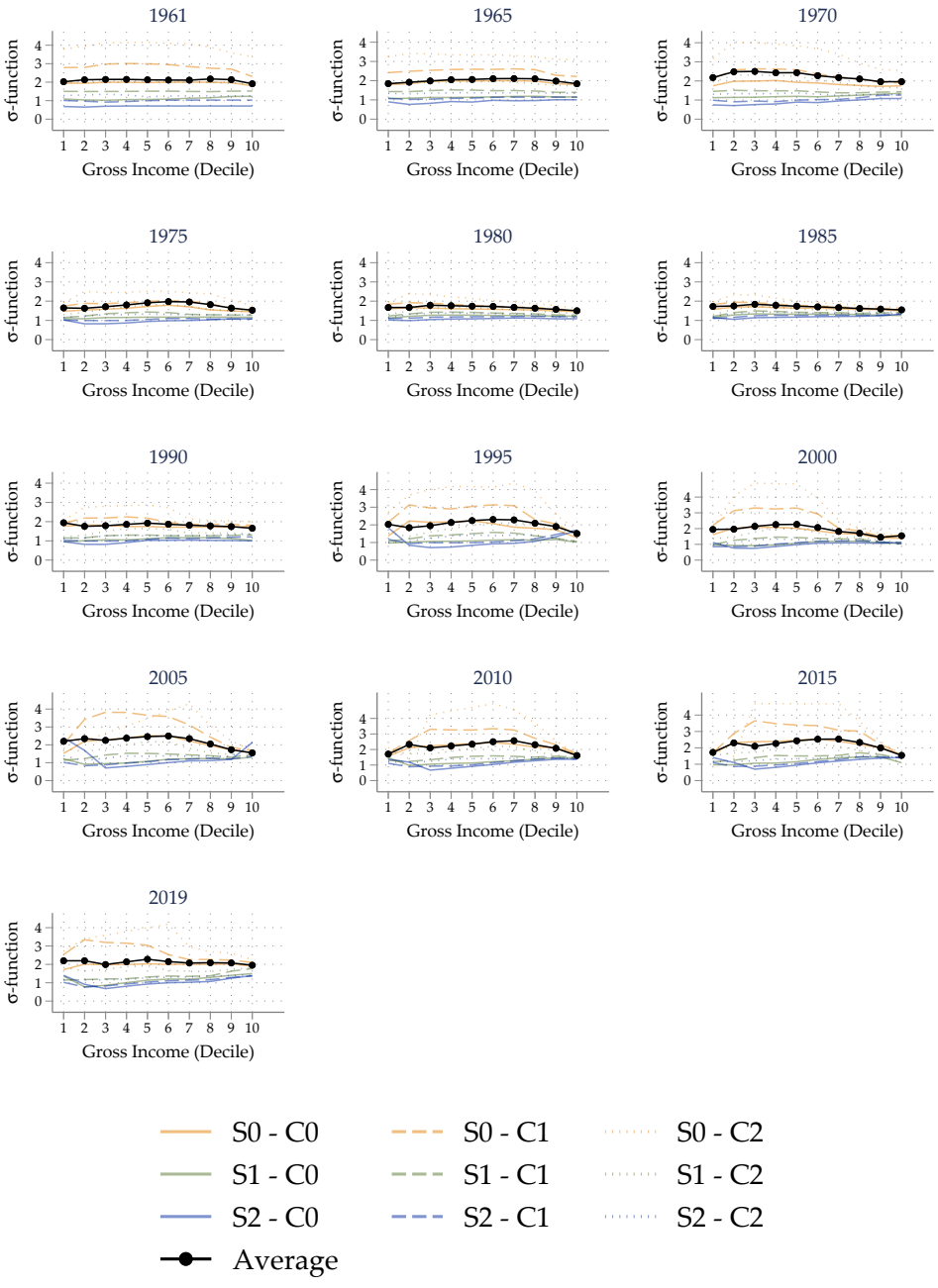
Note: This figure illustrates the reasoning behind the estimation of the splitting function σ (Figure C.8a) and relates it to the possibilities of marriage bonuses and penalties (Figure C.8b). Figure C.8b shows average tax rates of couples filing jointly at different income levels, where the splitting function is greater than 2 (blue), equal to 2 (red), between 1 and 2 (green), equal to 1 (orange) and smaller than one (yellow). The vertical colored arrows indicate the corresponding range of average tax rates for two singles with the same joint income like the couple filing jointly.

Figure C.9. – Average Tax Rates



Note: This figure displays average tax rates for married couples and single individuals for selected years. Average tax rates have been estimated using a kernel weighted local mean estimation. The solid part of the estimated average tax rate function satisfies the conditions for which σ-functions can be estimated uniquely (see Figure 3.2 and Figure C.10). All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.10. – Heterogeneity Analysis of the Empirical Splitting Function σ



Note: This figure explores the heterogeneity behind the empirical splitting function σ (see Figure 3.2). The baseline (comparison) tax unit is always a single tax unit (S) to a couple tax unit (C) accounting for different number of children (0 - no children, 1 - one child, 2 - more than one child) in both the baseline and the comparison tax unit. The σ-function is calculated for tax units by estimating mean average tax rates of couples and singles with different number of children. Deciles refer to the gross income distribution of couples in the respective year. All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

C.2.4 Political Feasibility and Welfare

In Section 3.4 of the main text, we summarize our analysis of past reforms from a political economy and a welfare perspective. Here we explain in detail, how we evaluate large reforms like the reforms in the past empirically.

For each tax unit j, we observe gross income y_0^j prior to the reform. Moreover, we observe whether a couple is a single earner couple or a dual earner couple. Based on the information from the CPS, we use the TAXSIM microsimulation model to calculate the person's tax payment $T_0(y_0^j)$, the average tax rate t_0^j and the marginal tax rate τ_0^j that are relevant for this tax unit prior to the reform. Finally, we observe the post-reform counterparts t_1^j and τ_1^j. We do not use the post-reform income y_1^j. Instead, we construct a (counterfactual) measure of the change in a taxpayer's tax burden that is only due to the reform, holding all individual characteristics, including the person's income, fixed, i.e. we compute the so-called "tax policy effect" (see, e.g., Eissa et al. (2008), Bargain et al. (2015), or Bierbrauer et al. (2021)).

We first explain how we would estimate the quantities of interest on the assumption that a tax reform takes place in an instant. In a second step, we take account of the complication that arises when a reform, such as the Reagan tax cuts, is phased in over several years.

Approximating the Revenue Implications of Tax Reforms. For each tax unit, we compute the change in taxes paid due to the reform. The change in taxes paid for a tax unit j is given by

$$\Delta T^j = t_1^j \, y_1^j - t_0^j \, y_0^j \, . \tag{C.24}$$

Behavioral responses to taxation imply that y_1^j will in general be different from y_0^j.

We take account of behavioral responses both at the extensive and the intensive margin. To do so, we think of every tax unit as being representative of a group of tax units with similar characteristics. Tax units similar to j are split into two groups. One group opts out and has $y_1^j = 0$ after the reform. The other groups stays in, $y_1^j > 0$. Remember that the extensive margin elasticity at an income of y measures the percentage share of tax units with an initial income close to y that choose zero earnings in response to a one percent decrease of their disposable income. Possibly, these elasticities differ not only by income, but also depend on marital status – i.e., on whether the tax unit is a single, a single earner couple, or a dual earner couple. Let π^j be the extensive margin elasticity

that applies to tax units similar to j. The reform induced percentage change in disposable income for tax units j is given by

$$\frac{y_0^j - T_1(y_0^j) - (y_0^j - T_0(y_0^j))}{y_0^j - T_0(y_0^j)} = \frac{T_1(y_0^j) - T_0(y_0^j)}{y_0^j - T_0(y_0^j)} = \frac{(t_j^1 - t_j^0) y_0^j}{y_0^j - t_0^j\, y_0^j} = \frac{t_j^1 - t_j^0}{1 - t_0^j} \,. \tag{C.25}$$

Thus, the fraction dropping out of the labor market is given by $\pi^j \, \frac{t_j^1 - t_j^0}{1 - t_0^j}$. The complementary fraction is staying in. For those who stay in, there are behavioral responses at the intensive margin. Our assumptions on preferences imply that such behavioral responses are driven entirely by changes of the marginal tax rates that tax units face. Thus, using a first order Taylor approximation,

$$y_1^j = y_0^j + (\tau_1^j - \tau_0^j)\; y_\tau^j \,,$$

where y_τ^j is the marginal effect that an infinitesimal change of the marginal tax rate has on j's taxable income (in the status quo). Using that $y_\tau^j = -y_{1-\tau}^j$, we can express this also via the marginal effect associated with a change of the net of tax rate $1 - \tau$. Hence,

$$y_1^j = y_0^j - (\tau_1^j - \tau_0^j)\; y_{1-\tau}^j \,,$$

Using the definition of the elasticity of taxable income (ETI), $\varepsilon^j := y_{1-\tau}^j \, \frac{1-\tau_0^j}{y_0^j}$, we can rewrite this as well as

$$y_1^j = \left(1 - \frac{\tau_1^j - \tau_0^j}{1 - \tau_0^j} \, \varepsilon^j\right) y_0^j \,.$$

Thus, for tax units that stay in, we have that

$$t_1^j \; y_1^j = t_1^j \left(1 - \frac{\tau_1^j - \tau_0^j}{1 - \tau_0^j} \, \varepsilon^j\right) y_0^j \,.$$

Collecting terms, overall we have that

$$\Delta T^j = \left(1 - \pi^j \, \frac{t_j^1 - t_j^0}{1 - t_0^j}\right) t_1^j \left(1 - \frac{\tau_1^j - \tau_0^j}{1 - \tau_0^j} \, \varepsilon^j\right) y_0^j - t_0^j \; y_0^j \,. \tag{C.26}$$

By summing across all tax units, we obtain an estimate for the aggregate change of tax revenue $\sum_j \Delta T^j$. Dividing by the number of tax units J yields an estimate for the revenue change per tax unit

$$\Delta R = \frac{1}{J} \sum_j \Delta T^j \,. \tag{C.27}$$

Implications for Individual Welfare: Number of Winners and Losers. The analysis of political feasibility in Section 3.4 rests on a comparison of the number of

individuals that benefit from a reform to the number of individuals that are made worse off. We now explain how we get to these number. We use conditions (C.17)-(C.20) to determine whether or not tax units benefit from a reform. When we bring these conditions to the data, we say that an individual tax unit j is a reform beneficiary if

$$\Delta R \;-\; (t_1^j - t_0^j)\ \max\{y_1^j, y_0^j\} \quad \geq \quad 0\,, \tag{C.28}$$

and loses if

$$\Delta R \;-\; (t_1^j - t_0^j)\ \min\{y_1^j, y_0^j\} \quad \leq \quad 0\,. \tag{C.29}$$

However, and as explained above, we think of an individual tax unit j as being representative of a group of tax units with similar characteristics. Thus, when average tax rates go up, $t_1^j - t_0^j > 0$, a fraction

$$\pi^j\ \frac{t_j^1 - t_j^0}{1 - t_0^j}$$

of this group has $y_1^j = 0$ and the complementary fraction with mass

$$1 - \pi^j\ \frac{t_j^1 - t_j^0}{1 - t_0^j}$$

has

$$y_1^j = \left(1 - \frac{\tau_1^j - \tau_0^j}{1 - \tau_0^j}\ \varepsilon^j\right) y_0^j\,.$$

By contrast, when average tax rates go down $t_1^j - t_0^j > 0$, we have tax units with $y_0^j = 0$ who now opt in at an income level of

$$y_1^j = \left(1 - \frac{\tau_1^j - \tau_0^j}{1 - \tau_0^j}\ \varepsilon^j\right) y_0^j\,.$$

and tax units with $y_0^j > 0$ who also choose this income level after the reform. The mass of tax units opting in equals

$$-\pi^j\ \frac{t_j^1 - t_j^0}{1 - t_0^j}$$

and the mass of tax units with a post-reform income of y_j^1 is then equal to

$$1 - \pi^j\ \frac{t_j^1 - t_j^0}{1 - t_0^j}\,.$$

Implications for Individual Welfare: Gains and Losses. For the welfare analysis in Section C.3.3 we aggregate the gains of reform winners and the losses of reform losers

using various social welfare functions. The expressions on the left-hand sides of (C.28) and (C.29), respectively, are alternative measures of by how much individuals are affected. In the main text we present an analysis using the left-hand sides of (C.28). This makes it demanding to find welfare gains. In Appendix C.3 we compare this baseline welfare measure against the alternative measure that uses the left-hand sides of (C.29), which makes it demanding to find welfare losses.

Adjusting for the Time Gap between Pre- and Post-Reform Years. The need of adjustment comes from the fact that the pre- and the post-reform tax systems apply in different calendar years. In case of the Reagan tax cuts, the reform was gradually implemented over several years, and we take 1985 as the last year with pre-reform schedule and 1988 as the first year with the post-reform schedule. We want an answer to a ceteris paribus question: All else equal, what is the effect of replacing the 1985-schedule by the 1988-schedule? To answer this question, we will have to compute an inflation adjusted version of y_0^j that we will denote by $\hat{y}_0^j$. If y_0^j is pre-reform income in 1985 USD, we think of $\hat{y}_0^j$ as the same pre-reform income, but expressed in 1988 USD.[13] Put differently, in moving from y_0^j to $\hat{y}_0^j$ we keep real income constant. We now explain how this adjustment modifies the above formulas.

First, note that we can express τ_0^j and τ_1^j also as

$$\tau_0^j = T_0'(y_0^j) \quad \text{and} \quad \tau_1^j = T_1'(\hat{y}_0^j) \,,$$

and t_0^j and t_1^j also as

$$t_0^j = \frac{T_0(y_0^j)}{y_j^0} \quad \text{and} \quad t_1^j = \frac{T_1(\hat{y}_0^j)}{\hat{y}_0^j} \,.$$

The *direct policy effect* of the reform is then given by

$$\Delta T^j = T_1(\hat{y}_0^j) - T_0(y_0^j). \tag{C.30}$$

The reform induced percentage change in disposable income for tax units j – see Equation (C.25) – is now given by

$$\frac{\hat{y}_0^j - T_1(\hat{y}_0^j) - (y_0^j - T_0(y_0^j))}{y_0^j - T_0(y^j)} = \frac{\hat{y}_0^j - T_1(\hat{y}_0^j)}{y_0^j - T_0(y^j)} - 1 := r^j \,. \tag{C.31}$$

[13] As Bierbrauer et al. (2021), we use the Consumer Price Index research series using current methods (CPI-U-RS) from the Bureau of Labor Statistics as an uprating factor to inflate/deflate incomes.

The fraction opting out when this expression is positive is given by $\pi^j \; r^j$. For those who stay in,

$$y_1^j = \hat{y}_0^j + (\tau_1^j - \tau_0^j) \; y_\tau^j \; .$$

Again, using the definition of the elasticity, $\varepsilon^j := y_{1-\tau}^j \; \frac{1-\tau_0^j}{y_0^j}$, we can rewrite this as well as

$$y_1^j = \left(1 - \frac{\tau_1^j - \tau_0^j}{1 - \tau_0^j} \; \varepsilon^j\right) \hat{y}_0^j \; .$$

Thus, for tax units that stay in, we have that

$$t_1^j \; y_1^j = t_1^j \left(1 - \frac{\tau_1^j - \tau_0^j}{1 - \tau_0^j} \; \varepsilon^j\right) \hat{y}_0^j \; .$$

Collecting terms, overall we have that

$$\Delta T^j = \left(1 - \pi^j \; \frac{t_j^1 - t_j^0}{1 - t_0^j}\right) t_1^j \left(1 - \frac{\tau_1^j - \tau_0^j}{1 - \tau_0^j} \; \varepsilon^j\right) \hat{y}_0^j - t_0^j \; y_0^j \; . \tag{C.32}$$

By summing across all tax units we obtain an estimate for the aggregate change of tax revenue $\sum_j \Delta T^j$. Dividing by the number of tax units J yields an estimate for the revenue change per tax unit

$$\Delta R = \frac{1}{J} \sum_j \Delta T^j \; . \tag{C.33}$$

We say that an individual tax unit j is a reform beneficiary if

$$\Delta R \;\; - \;\; (t_1^j - t_0^j) \; \max\{y_1^j, \hat{y}_0^j\} \quad \geq \quad 0 \; , \tag{C.34}$$

and loses if

$$\Delta R \;\; - \;\; (t_1^j - t_0^j) \; \min\{y_1^j, \hat{y}_0^j\} \quad \leq \quad 0 \; . \tag{C.35}$$

and we use the left-hand sides of (C.34) and (C.35) as measures of how much individuals gain or lose due to a tax reform.

Hypothetical Reforms. Our methodology is not only valid for reforms that were implemented in the past. We can also apply it to the analysis of hypothetical tax reforms that did not take place. Given our focus on the tax treatment of couples and singles, one such hypothetical reform type is one in which we take the change in taxes for singles through a particular historical reform as given but translate that observed tax change for singles according to the pre-reform σ-function to couples.

In particular, we replace the post-reform tax function for couples by a hypothetical one that is linked via the pre-reform σ-function to the post-reform tax schedule for singles. For instance, the hypothetical mechanical change in tax payments for couples is then given by

$$\Delta T^j = T_{m1}^{hyp}(\hat{y}_{m0}) - T_{m0}(y_{m0}), \tag{C.36}$$

$$\Delta T^j = \sigma_0 \, T_{s1}\left(\frac{\hat{y}_{m0}}{\sigma_0}\right) - T_{m0}(y_{m0}). \tag{C.37}$$

C.2.5 Calibration of Revenue Functions

We explain, how we calibrate the revenue functions characterized in Proposition 1.

The main ingredients of these equations are (i) estimates of the gross income distribution through the cumulative distribution function and the probability density function, (ii) an approximation of marginal and participation tax rates, and (iii) assumptions about intensive margin behavioral responses.

We estimate gross income distributions for couples and singles from the CPS data using an adaptive kernel density estimator with a Gaussian kernel on an equally spaced grid between the first percentile and a value equal to the 99.9th percentile of the gross income distribution. Subsequently, we adjust the estimated distribution to the share of tax units without any income. Figures C.11 and C.12 show the resulting cumulative distribution functions (CDF) and probability density functions (PDF).

We estimate effective marginal tax rates based on the TAXSIM microsimulation model for every tax unit in the data. To approximate effective marginal tax rates at a given income level, we estimate a kernel-weighted local polynomial using the same grid and bandwidth as for the estimation of the income distributions. Figure C.13 shows the estimated marginal tax rates.

Based on the assumptions about behavioral responses at the intensive margin illustrated in Table 3.1, we assign every single tax unit the respective intensive margin elasticity and every couple a weighted average based on the income shares of the primary and secondary earner. In line with the estimation of average effective marginal tax rates, we approximate the intensive margin elasticity at a given income level using a kernel weighted local

polynomial. Note that even though elasticities are constant for primary and secondary earners, the average elasticity for couples can vary across the income distribution and across years due to the change in the earnings share of primary and secondary earners. Figure C.14 shows for the baseline assumptions about the elasticity of taxable income, how the average elasticities assigned to couples varies across the income distribution.

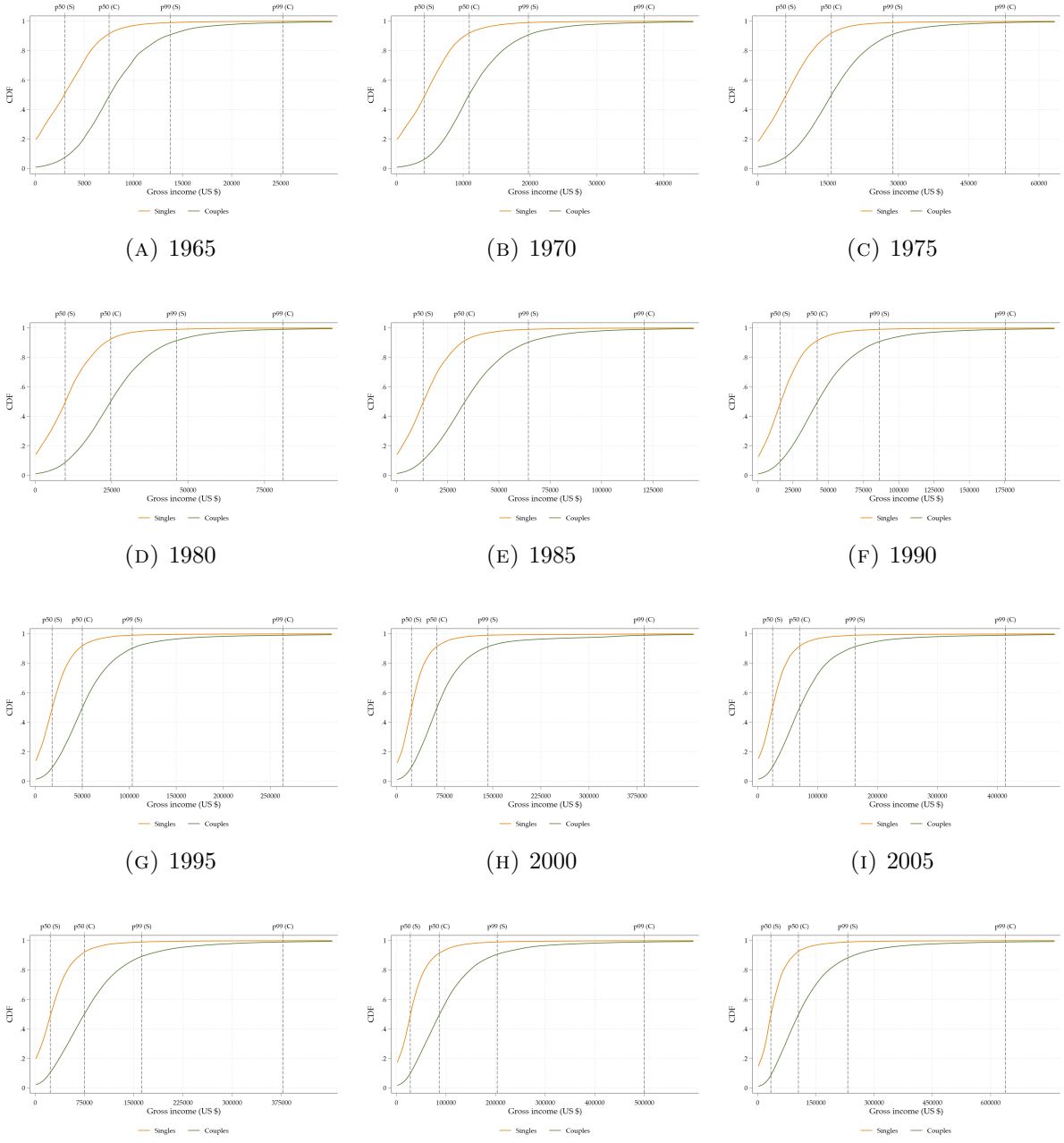
The additional ingredients with respect to the ones used above are (i) the share of dual and single earner couples, (ii) separate income distributions for dual and single earner couples - see Figures C.15 and C.16, (iii) estimates of the participation tax rate - see Figure C.17, and (iv) assumptions about the participation elasticity. For the latter, we assume that the participation elasticity does not vary across tax unit types, but vary across the income distribution, i.e. it decreases from 0.65 to 0.25 between a gross income of zero and the 90th percentile of the gross income distribution (see Figure C.18).

Beyond the elasticities for primary and secondary earners, the estimation of these revenue functions requires (i) separate income distributions for the primary and the secondary earner - see Figures C.19 and C.20, and (ii) an estimate of the couples' marginal tax rate at a given primary and secondary earnings level (see Figures C.21 and C.22).

For the consideration of extensive margin responses, we assume that the extensive margin reaction of dual earner couples does not differ of whether the tax treatment of primary or secondary earnings are modified, or whether they are single earner couples or dual earner couples, i.e. $\pi_{dec,1} = \pi_{dec,2} = \pi_{dec} = \pi_{sec}$.

Again, we assume that participation responses are larger at the bottom of the income distribution, i.e. the participation elasticities decrease from 0.65 to 0.25 between a gross income of zero and the 90th percentile of the gross income distribution (see Figure C.18). Note that in contrast to intensive margin responses, we cannot put the participation elasticity in front of the expectation operator, because the participation elasticity is assumed to be income dependent. Therefore, we first compute the term inside the expectation operator at the tax unit level, and estimate the average of this term across varying levels of primary and secondary earnings.

Figure C.11. – Cumulative Distribution Function



(A) 1965 (B) 1970 (C) 1975

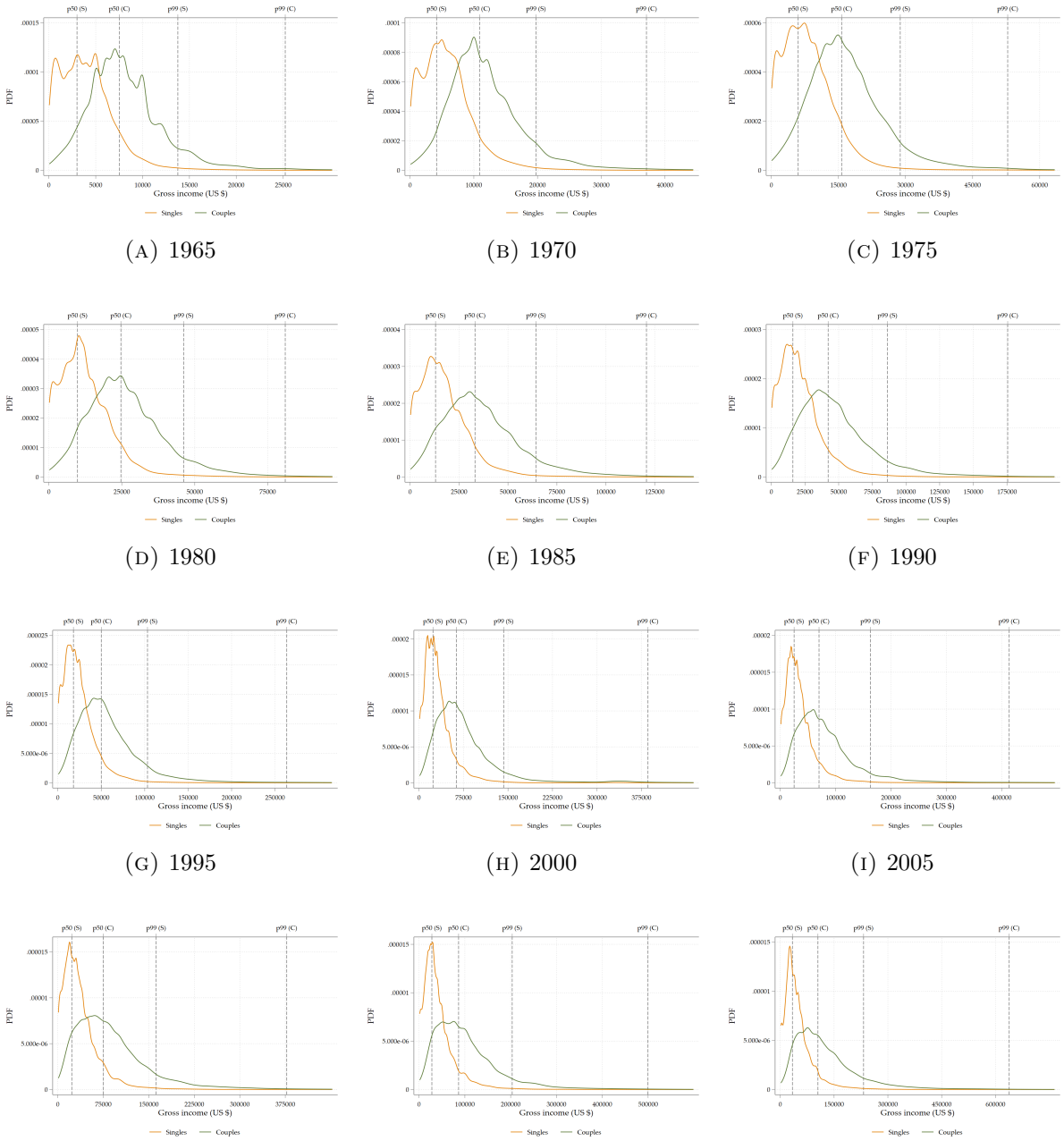
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure displays estimates of the cumulative distribution function of gross income for singles (orange line) and couples (green line) in the respective year. Distributions are estimated using an adaptive kernel density estimator with a Gaussian kernel on an equally spaced grid between the first percentile and a value equal to the 99.9th percentile of the gross income distribution.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.12. – Probability Density Function



(A) 1965 (B) 1970 (C) 1975

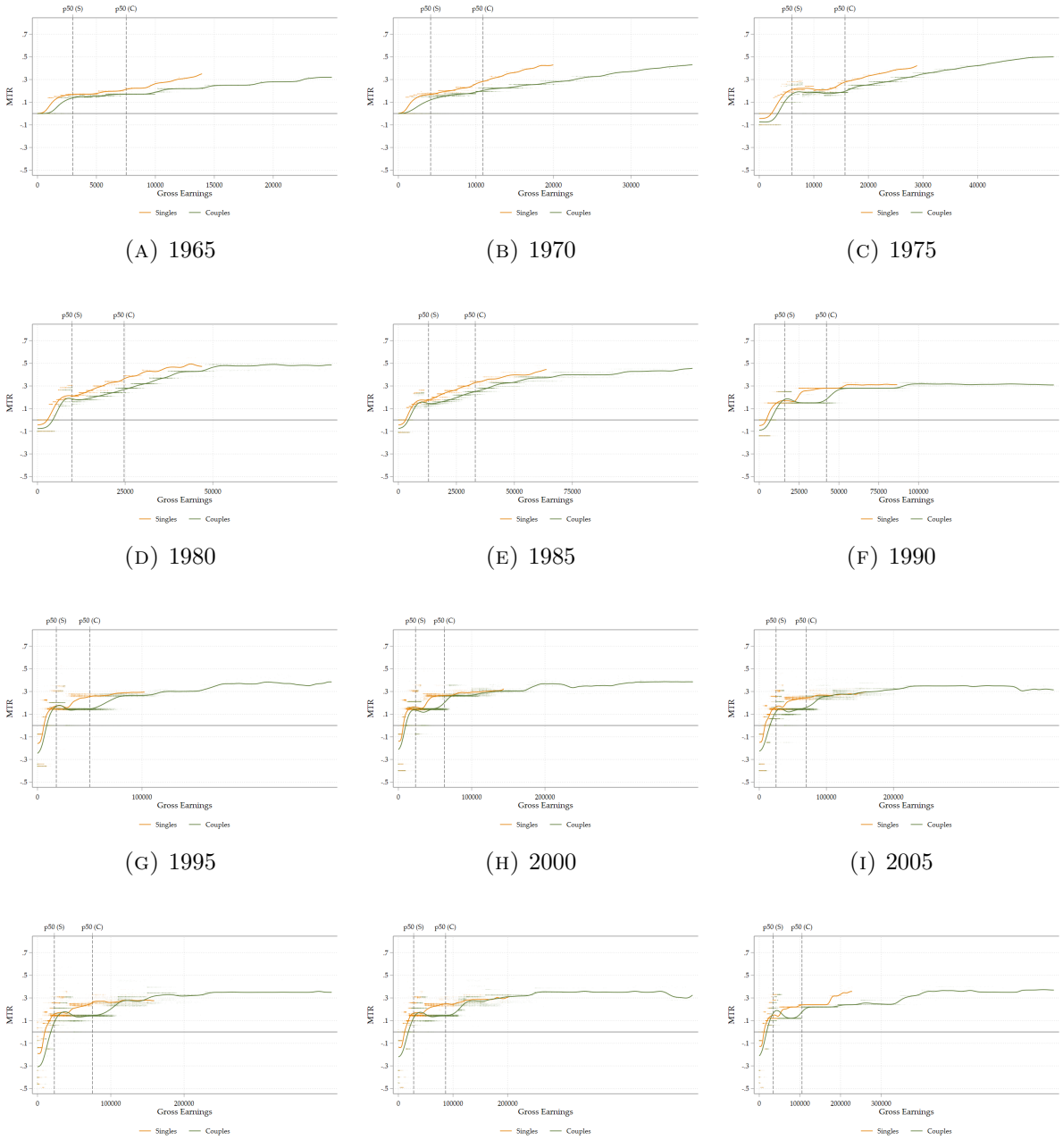
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure displays estimates of the probability density function of gross income for singles (orange line) and couples (green line) in the respective year. Distributions are estimated using an adaptive kernel density estimator with a Gaussian kernel on an equally spaced grid between the first percentile and a value equal to the 99.9th percentile of the gross income distribution.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.13. – Effective Marginal Tax Rates



(A) 1965 (B) 1970 (C) 1975

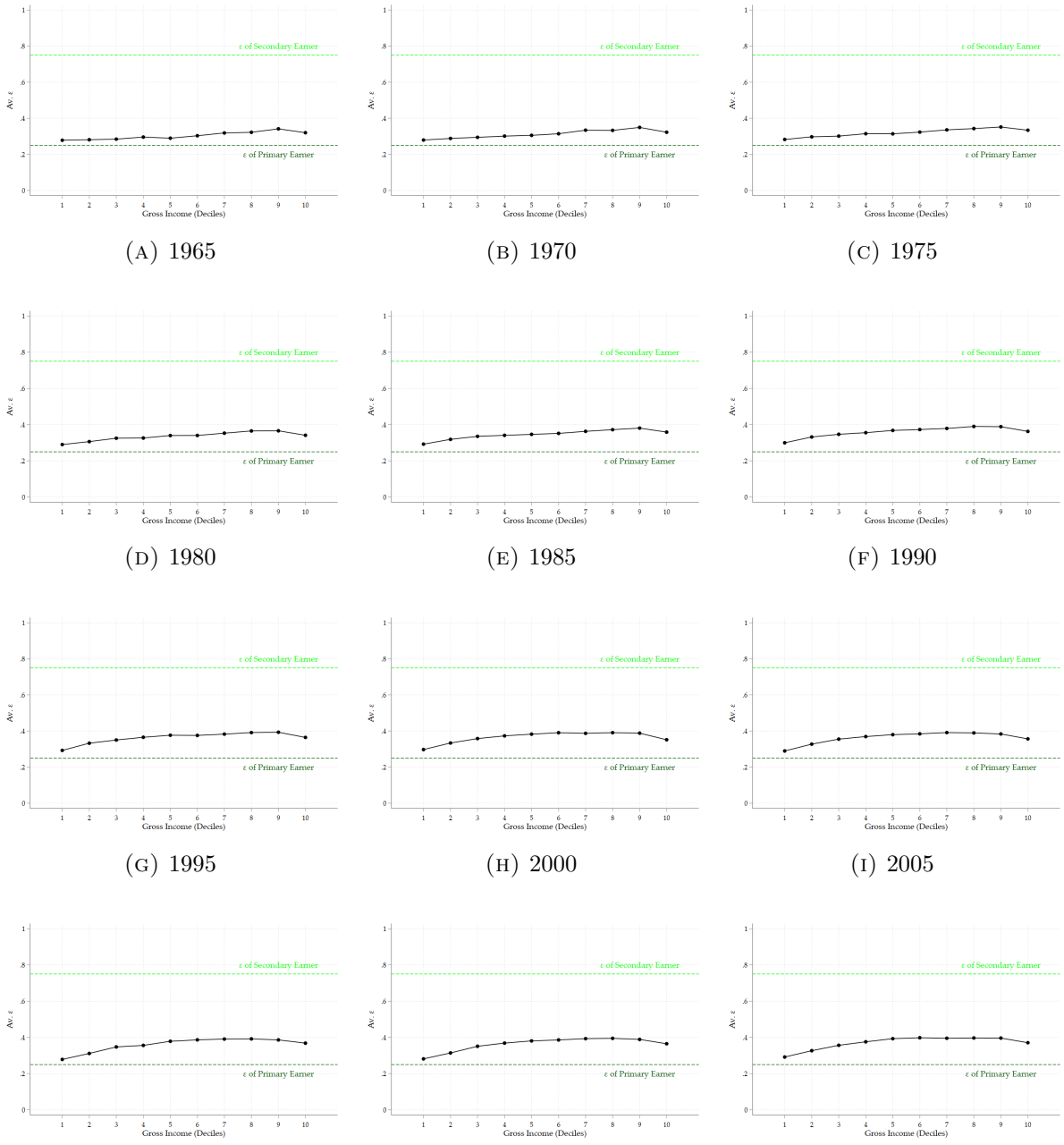
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure displays average effective marginal tax rates for singles (orange lines) and couples (green lines) before the reform (solid lines) and after the reform (dashed lines). Average marginal tax rates at a given gross income level are estimated with a kernel-weighted local polynomial using the same grid and bandwidth as for the estimation of the income distributions (see Figure C.11 and C.12).
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.14. – Average Elasticities of Couples



(A) 1965 (B) 1970 (C) 1975

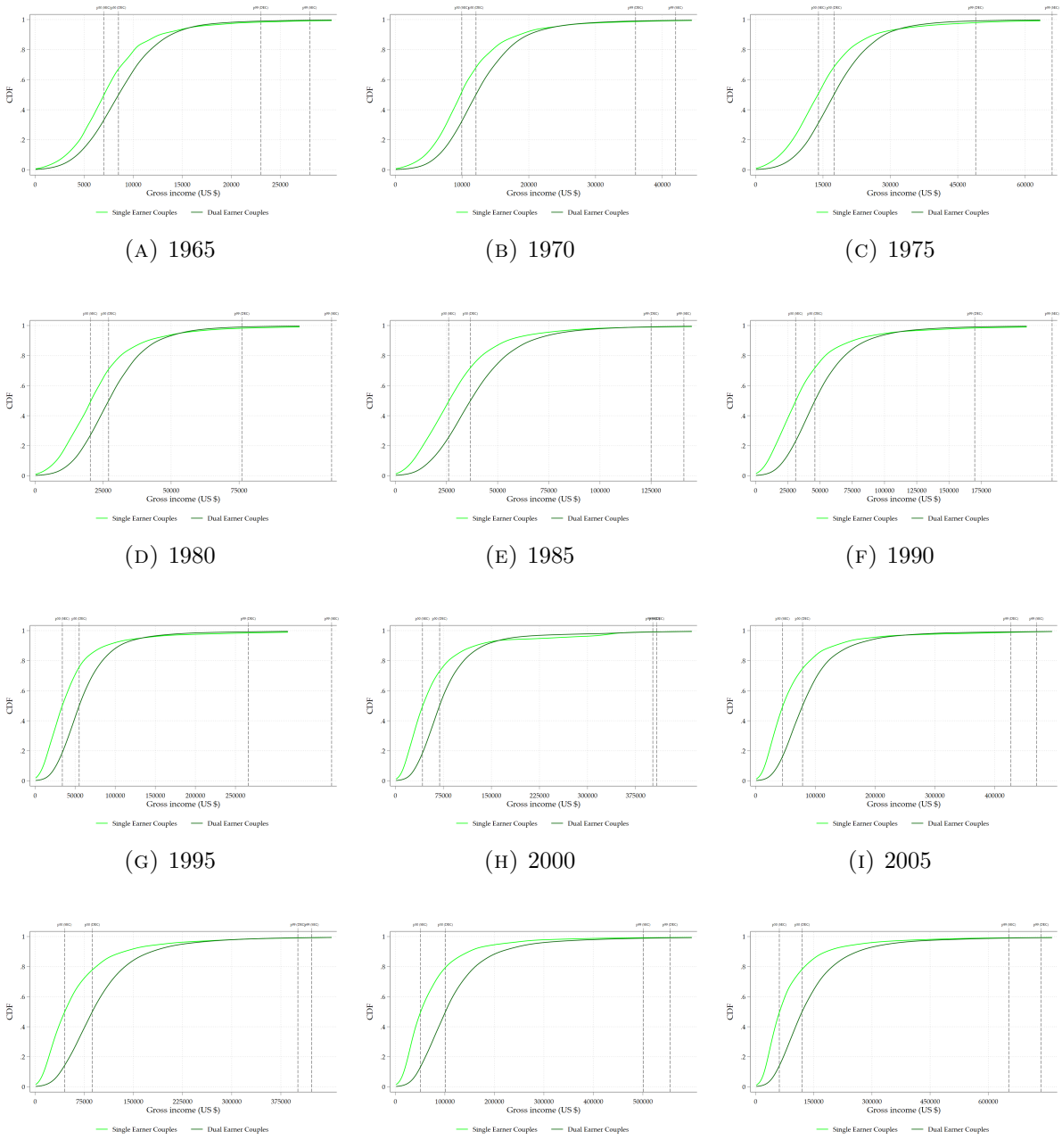
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure displays the average intensive margin elasticity of taxable income for couples across gross income deciles in the respective year. Elasticities are calculated for every couple based on an income-share weighted elasticity of 0.25 for the primary earner and 0.75 for the secondary earner (see Table 3.1). Deciles are computed based on the gross income distribution of couples. Earnings shares are based on wage, business and farm income.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.15. – CDF, Single Earner and Dual Earner Couples



(A) 1965 (B) 1970 (C) 1975

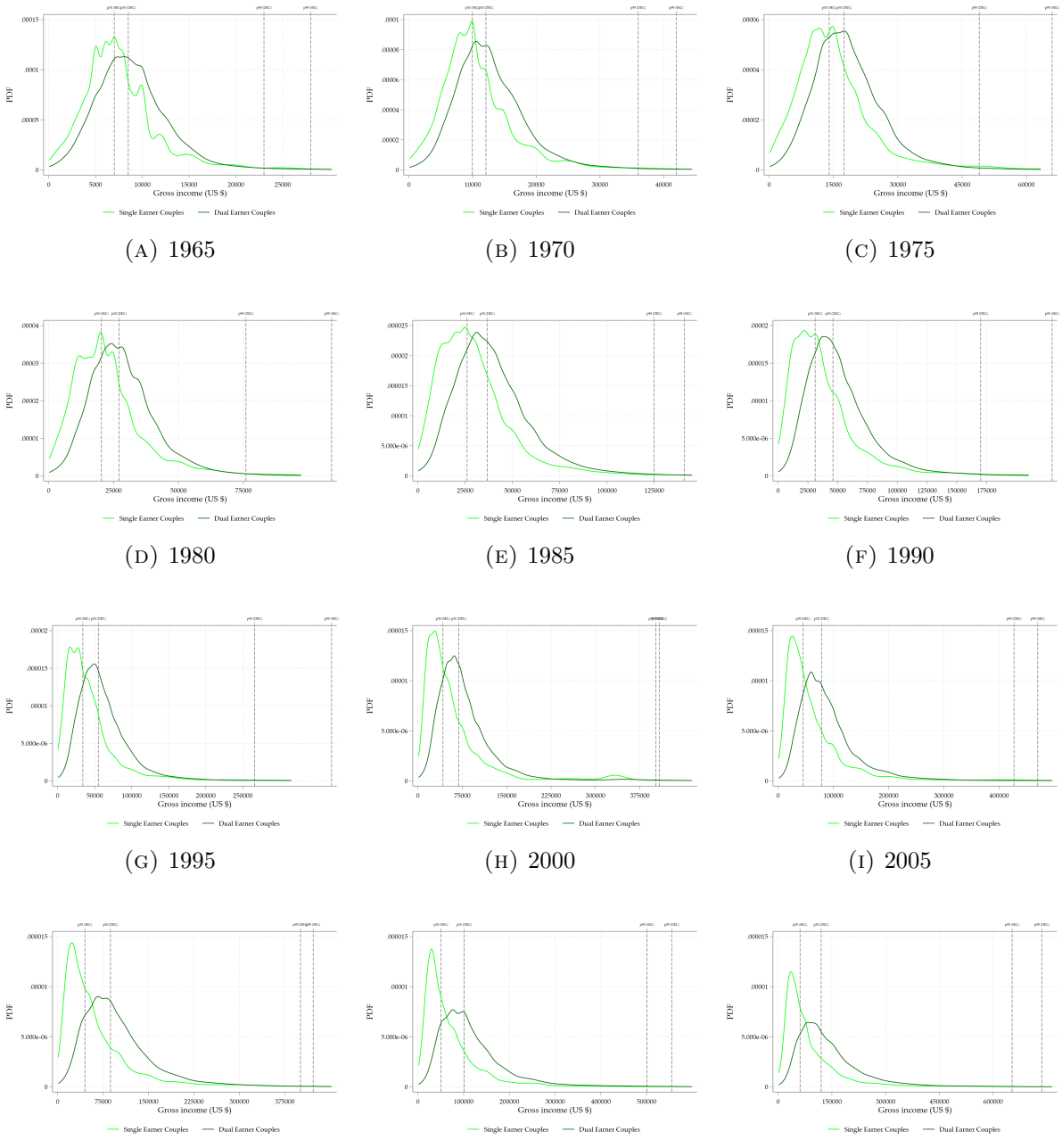
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure displays estimates of the cumulative distribution function of gross income for single earner (light green line) and dual earner couples (dark green line) in the respective year. Distributions are estimated using an adaptive kernel density estimator with a Gaussian kernel on an equally spaced grid between the first percentile and a value equal to the 99.9th percentile of the gross income distribution.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.16. – PDF, Single Earner and Dual Earner Couples



(A) 1965 (B) 1970 (C) 1975

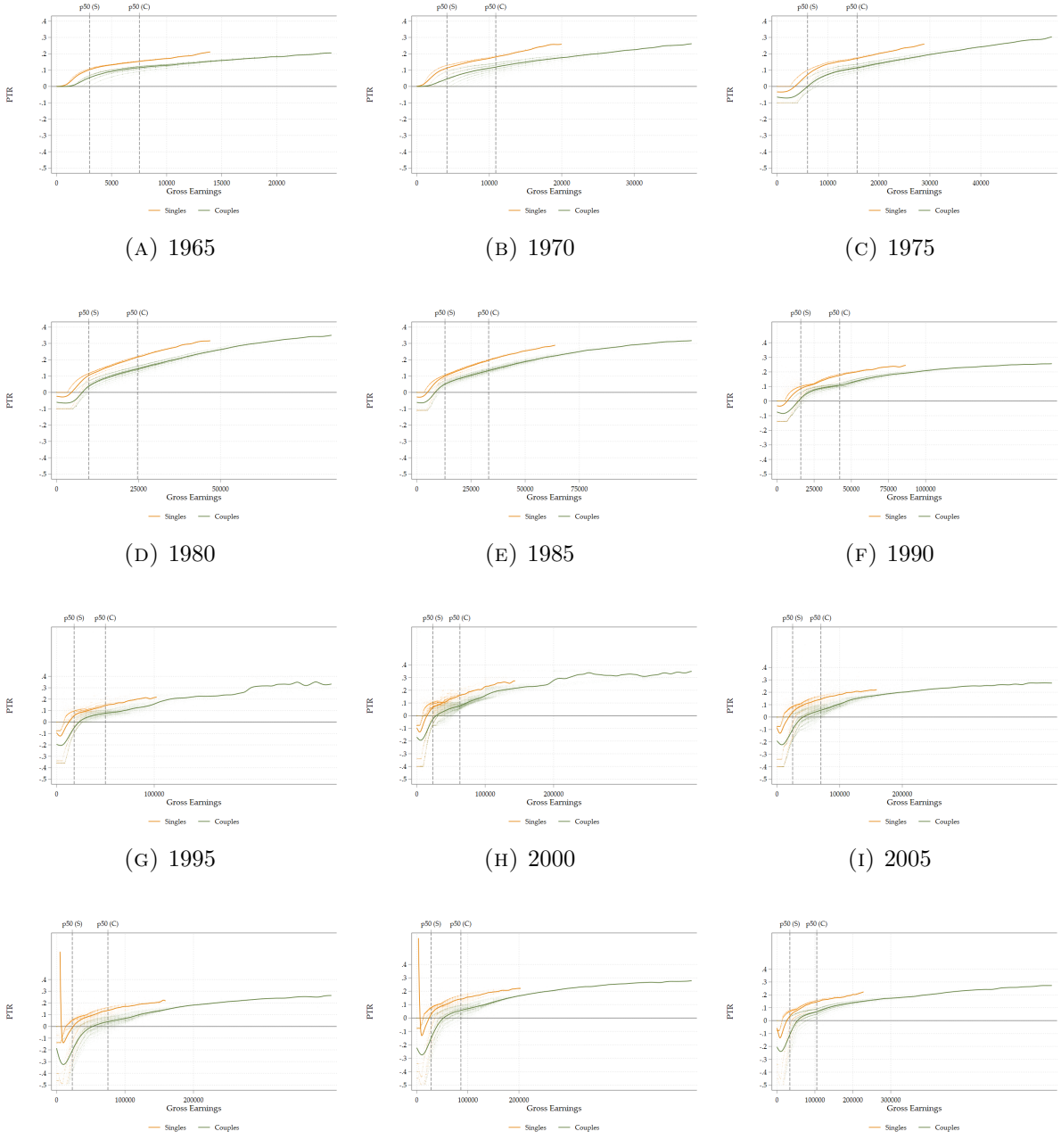
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure displays estimates of the probability density function of gross income for single earner (light green line) and dual earner couples (dark green line) in the respective year. Distributions are estimated using an adaptive kernel density estimator with a Gaussian kernel on an equally spaced grid between the first percentile and a value equal to the 99.9th percentile of the gross income distribution.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.17. – Participation Tax Rates



(A) 1965 (B) 1970 (C) 1975

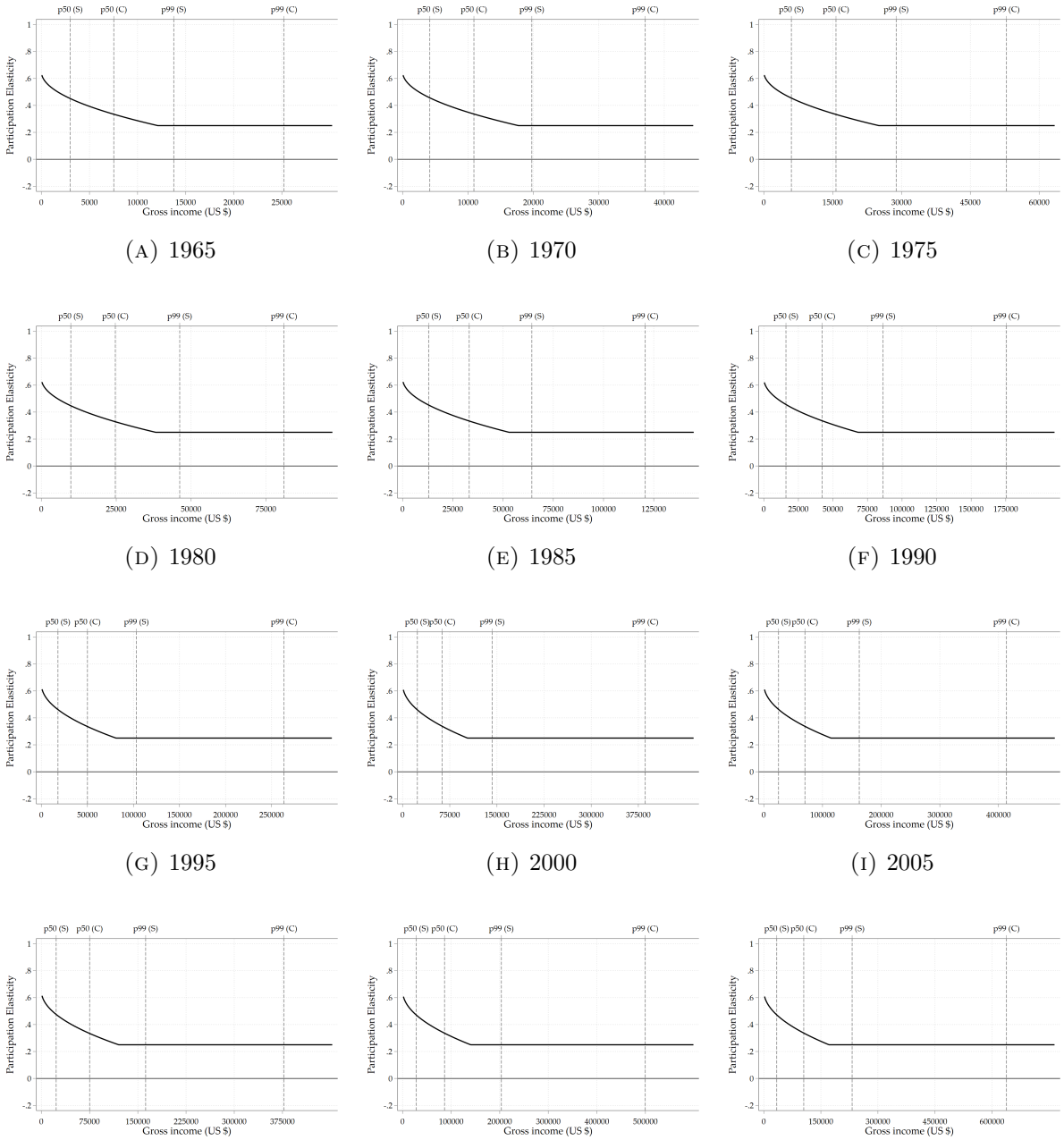
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure displays participation tax rates for every single (light orange dots) and every couple (light green dots) in the respective year. Solid orange (green) lines represent estimates of the average marginal tax rate schedule for singles (couples). Average participation tax rates at a given gross income level are estimated with a kernel-weighted local polynomial using the same bandwidth as for the estimation of the income distributions (see Figure C.11 and C.12).
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.18. – Participation Elasticities



(A) 1965 (B) 1970 (C) 1975

(D) 1980 (E) 1985 (F) 1990

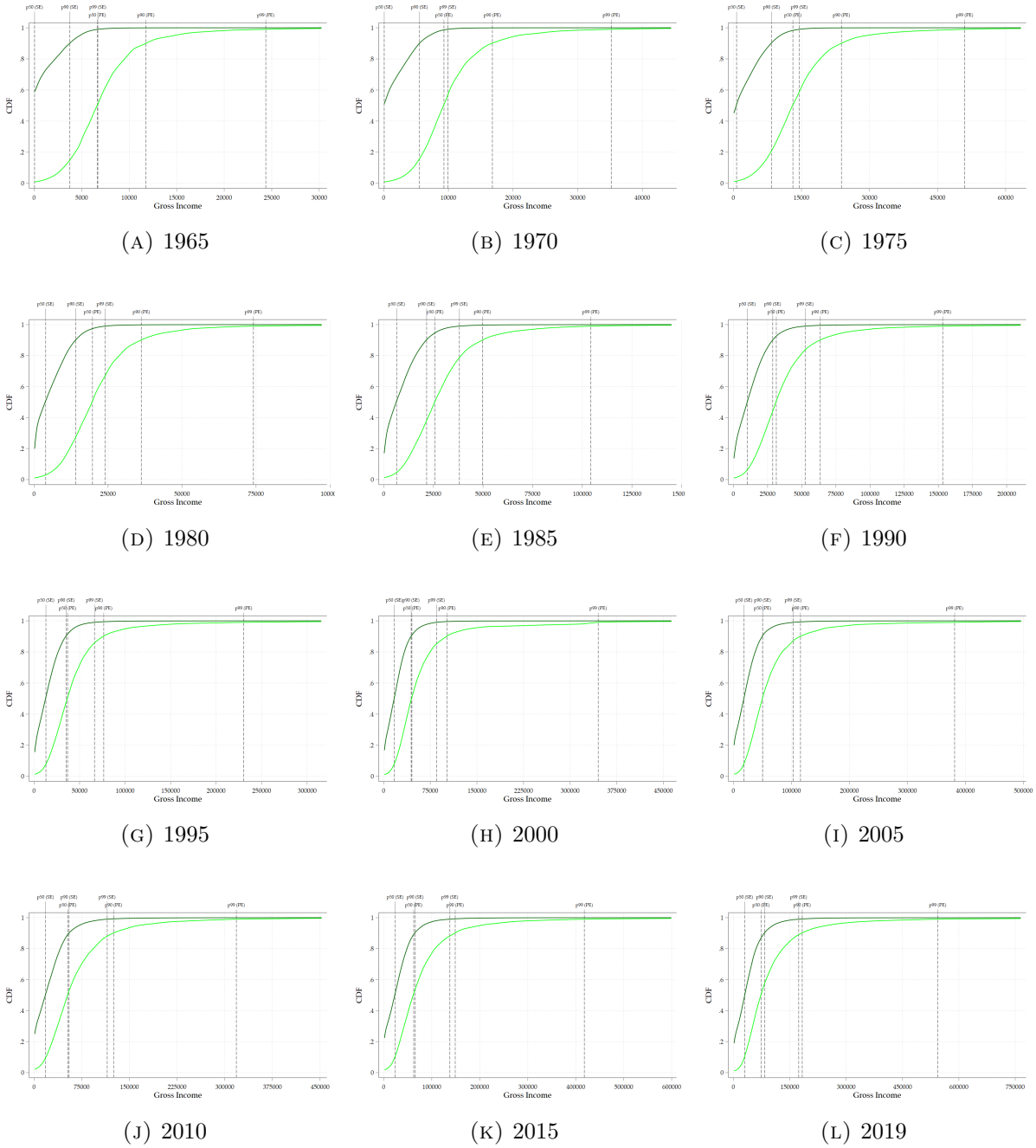
(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure displays for every year the evolution of the participation elasticity over income. The participation elasticity is assumed to decrease from 0.65 to 0.25 between zero and the 90th percentile of the gross income distribution based on the formula $\pi = 0.65 - 0.4\left(\frac{y}{y_{P90}}\right)^{\frac{1}{2}}$.

Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.19. – Cumulative Distribution Function, Primary and Secondary Earners



(A) 1965 (B) 1970 (C) 1975

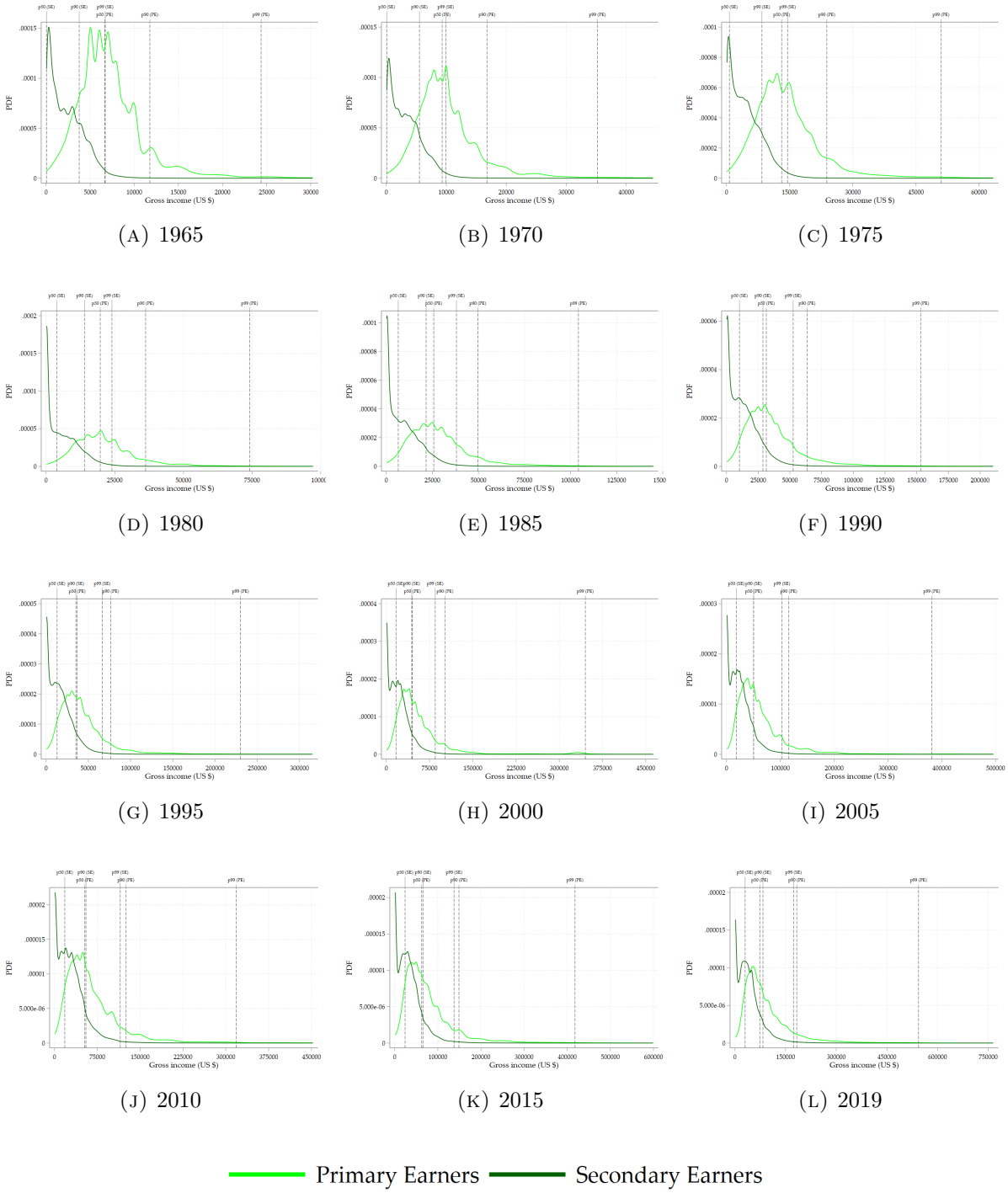
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure shows for selected years the cumulative density function of primary and secondary earnings. Distributions are estimated using an adaptive kernel density estimator with a Gaussian kernel on an equally spaced grid between the first percentile and a value equal to the 99.9th percentile of the gross income distribution.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.20. – Probability Density Function, Primary and Secondary Earners



(A) 1965 (B) 1970 (C) 1975

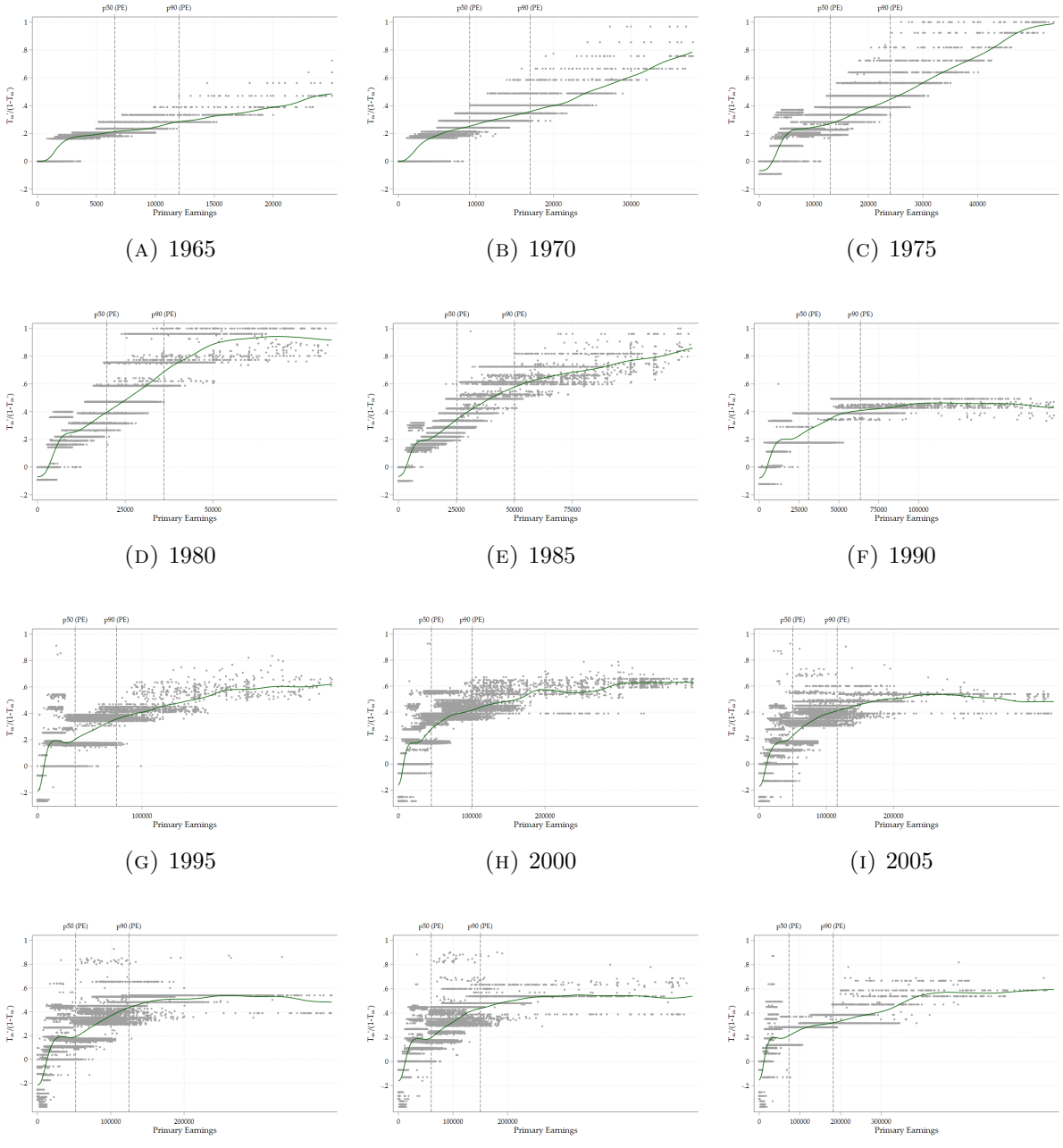
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

(J) 2010 (K) 2015 (L) 2019

Note: This figure shows for selected years the probability density function of primary and secondary earnings. Distributions are estimated using an adaptive kernel density estimator with a Gaussian kernel on an equally spaced grid between the first percentile and a value equal to the 99.9th percentile of the gross income distribution.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.21. – Average Marginal Tax Rates by Primary Earnings



(A) 1965 (B) 1970 (C) 1975

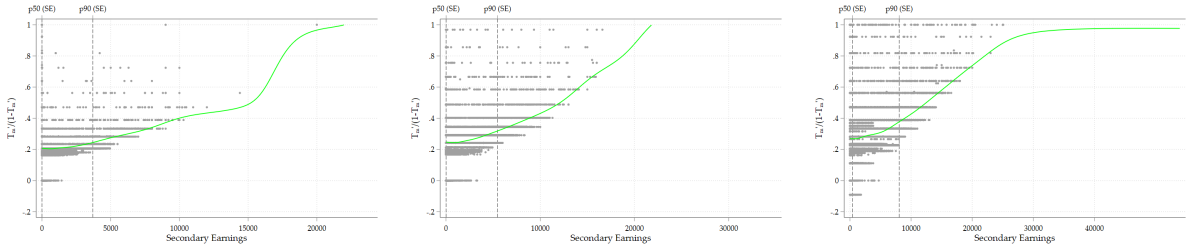
(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005

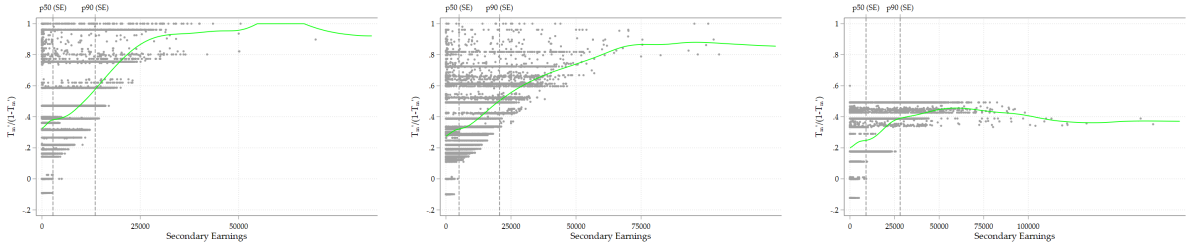
(J) 2010 (K) 2015 (L) 2019

Note: This figure shows for selected years the marginal tax rate ratio $\frac{T'}{1-T'}$ across primary earnings. The solid line represents an average calculated using a local polynomial regression.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.22. – Average Marginal Tax Rates by Secondary Earnings

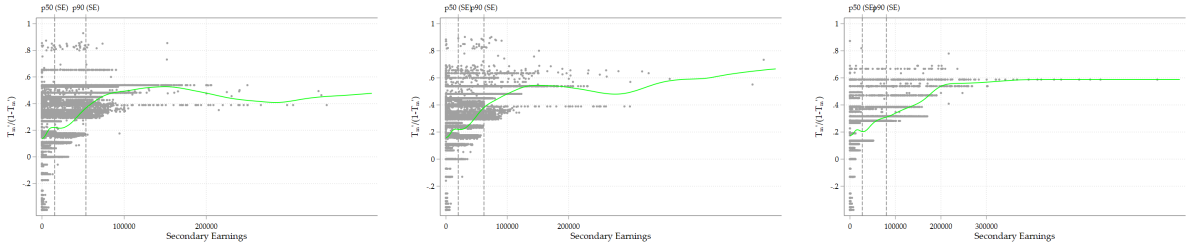


(A) 1965 (B) 1970 (C) 1975



(D) 1980 (E) 1985 (F) 1990

(G) 1995 (H) 2000 (I) 2005



(J) 2010 (K) 2015 (L) 2019

Note: This figure shows for selected years the marginal tax rate ratio $\frac{T'}{1-T'}$ across secondary earnings. The solid line represents an average calculated using a local polynomial regression.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

C.3 SUPPLEMENTARY MATERIAL: REFORMS *IN* THE SYSTEM

This part of the appendix provides additional explanations and supplementary results for the analysis of reforms in the system, i.e., the analysis of past reforms in Section 3.4 of the main text. We illustrate supplementary results on how past tax reforms affected the splitting function and marriage bonuses (Appendix C.3.1), and on the political feasibility (Appendix C.3.2) and welfare effects (Appendix C.3.3) of past reforms.

C.3.1 Marriage Bonuses and Penalties

In Section 3.4, we show, how the splitting function and marriage bonuses and penalties changed over time. In this part of the appendix, we provide details on how each of the eleven reforms between 1964 and 2017 directly affected penalties and bonuses through reforms in the system, i.e., reforms that differed in the tax treatment of couples and singles, but left the tax base as the sum of total household income untouched.

For this purpose, Figure C.23 first shows the mechanical effect of each tax reform for different household types. Some of the reforms provided very similar per-capita changes in tax liabilities for singles and couples, e.g., RA64. Other reforms stand out, since they affected singles and couples very differently. For instance, TRA69 provided much larger per-capita tax cuts to singles than for couples, while JGTRRA03 shows the reverse pattern. Under a constant splitting function, tax cuts for couples would thus have been much larger in the former, and much smaller under the latter reform.

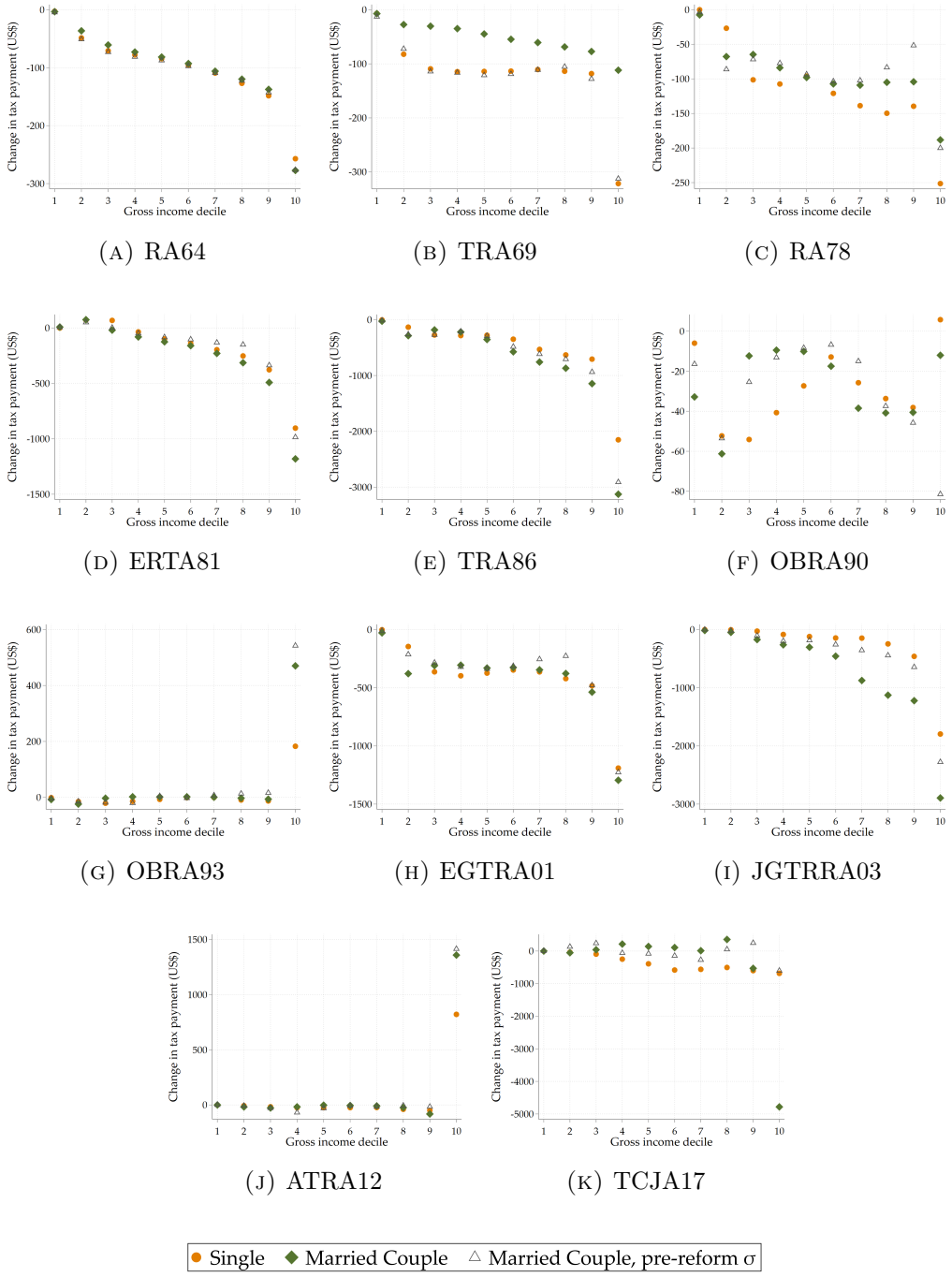
This differential tax changes for singles and couples therefore implicitly shifted the splitting function (see Figure C.24). In particular, TRA69 shifted the splitting function downwards, thereby increasing the potential for marriage penalties. In contrast, JGTRRA03 shifted the splitting function upwards and thereby increased the potential for marriage bonuses and made penalties less frequent. This is confirmed by a direct evaluation of the changes in marriage benefits and bonuses in Figure C.25. This figure also again illustrates that marriage penalties tend to occur when incomes of the two spouses are relatively equal while bonuses are particularly frequent for couples with very unequal incomes. Over time, the median couple has moved from the area associated with marriage bonuses to the area of marriage penalties.

Importantly, reforms did not always change marriage penalties across the board. For instance, the Trump tax cut in 2017 provided particularly large tax cuts for married

couples at the top of the income distribution while singles in the middle of the distribution received larger per capita tax cuts than couples (see Figure C.23k). This translated in an asymmetric shift of the splitting function (see Figure C.24k), and correspondingly particularly eliminations of marriage penalties for incomes at the very top (see Figure C.26e).

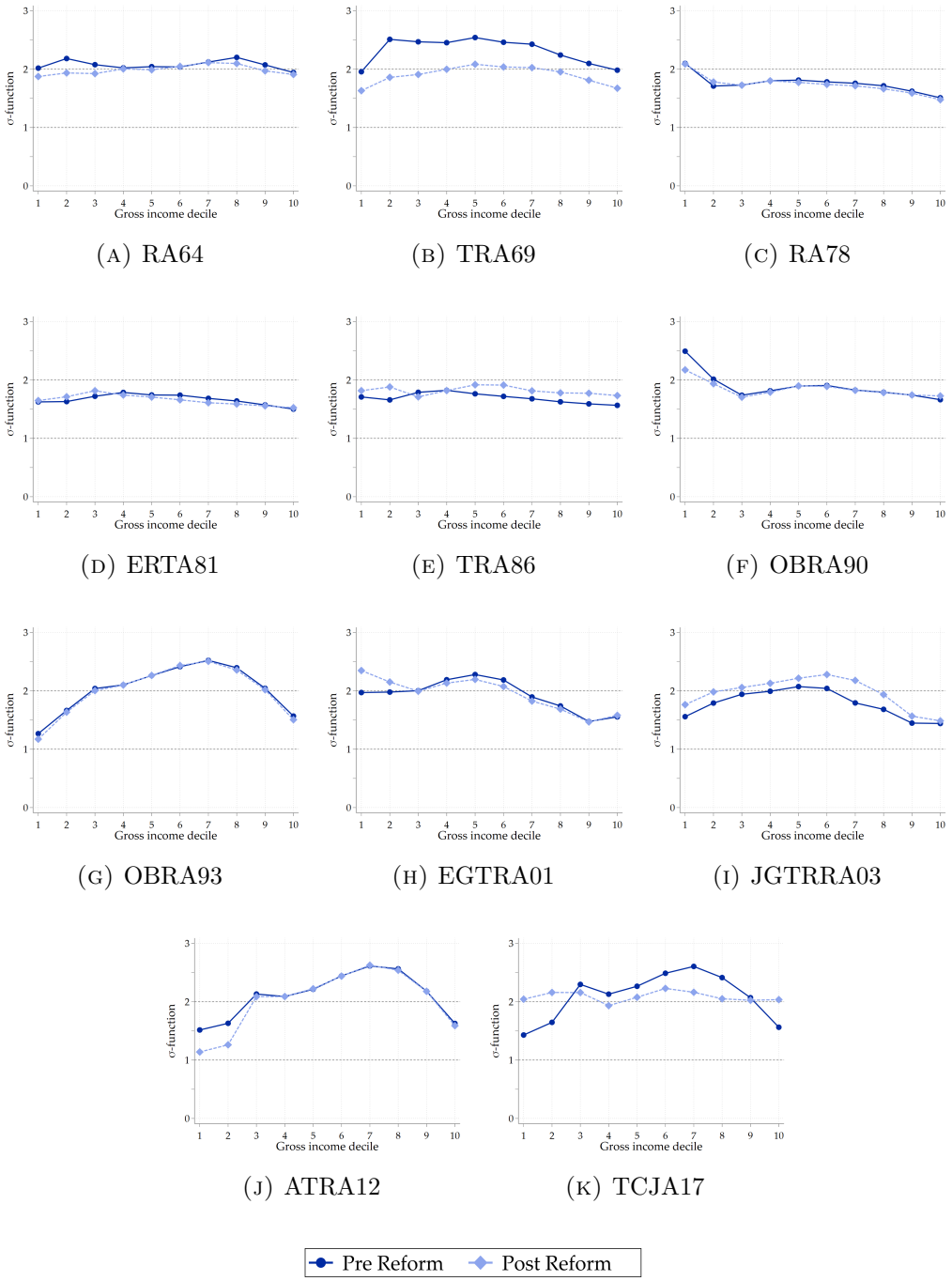
Through the 2017 reform, the splitting function became constant again and is now close to two across all income levels. It thereby returned to their original position at the beginning of the 1960s.

Figure C.23. – Mechanical Effect of Tax Reform



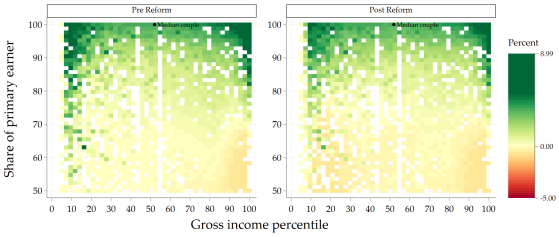
Note: This figure shows how tax reforms affected the per-capita tax burden of singles (orange circles) and couples (green diamonds), holding their income fixed at the pre-reform level, by deciles of the per capita gross income distribution. At the tax unit level, the change is equal to $T_{s1}(\hat{y}_{s0}) - T_{s0}(y_{s0})$ for singles and $T_{m1}(\hat{y}_{m0}) - T_{m0}(y_{m0})$ for couples. Post-reform tax payments $T_1(\hat{y}_0)$ are calculated based on the inflation-adjusted pre-reform income $\hat{y}_0$ using the CPI-U-RS deflator as uprating factor. In addition, the figure displays the hypothetical change in tax liability for couples under the assumption that observed tax changes of singles would have translated according to the empirical pre-reform splitting function σ to couples, i.e. $\sigma_0\, T_{s1}\left(\frac{\hat{y}_{m0}}{\sigma_0}\right) - T_{m0}(y_{m0})$ (grey triangles). For details on the methodology on the analysis of actual and hypothetical tax reforms, see Appendix C.2.4. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.24. – Change of σ

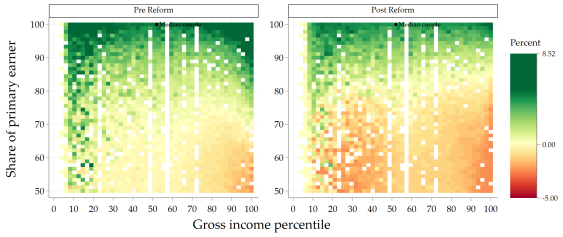


Note: This figure shows the effects on the splitting function σ, holding incomes fixed at the pre-reform level. Pre-reform (dark blue circles) and post-reform (light blue diamonds) splitting functions are calculated by estimating mean average tax rates of couples and singles in the respective year. Mean average tax rates are used to solve numerically for σ (see Appendix C.2.3)). Deciles refer to the gross income distribution of couples in the respective year. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

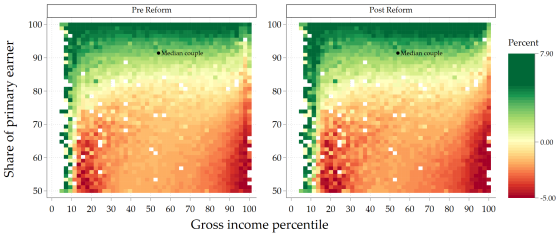
Figure C.25. – Change of Marriage Bonuses and Penalties



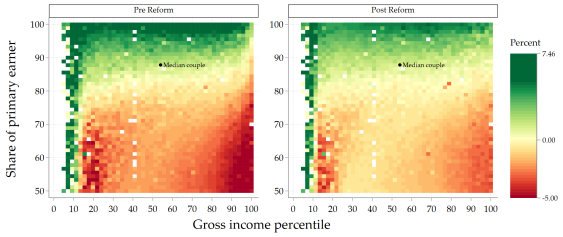
(A) RA64



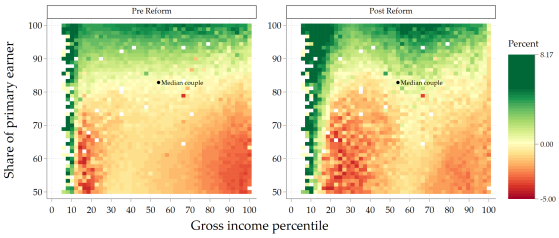
(B) TRA69



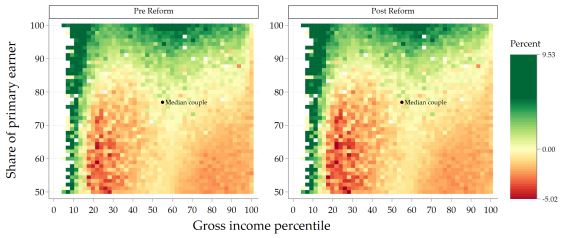
(C) RA78



(D) ERTA81



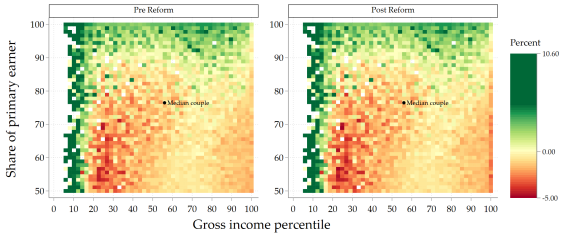
(E) TRA86



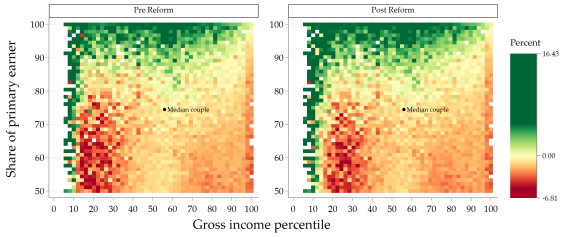
(F) OBRA90

Note: This figure shows marriage bonuses and penalties relative to gross income. Each square in a figure represents an average of marriage bonuses (green) or penalties (red) for a group of tax units at a particular income percentile (horizontal axis) and with a particular primary earner income share (vertical axis). Relative marriage bonuses/penalties relate the absolute monetary advantage from filing as a married couple to the total income of the couple, i.e. $\frac{T_m(y_1+y_2)-(T_s(y_1)+T_s(y_2))}{y_1+y_2}$ (see Appendix C.2.3 for details). The distribution of marriage penalties and bonuses is shown for the pre-reform year (left panel) and the post-reform year (right panel). Income percentiles at the horizontal axis refer to the per capita income distribution of the full sample, i.e. individuals in couples are assigned half of the joint income. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

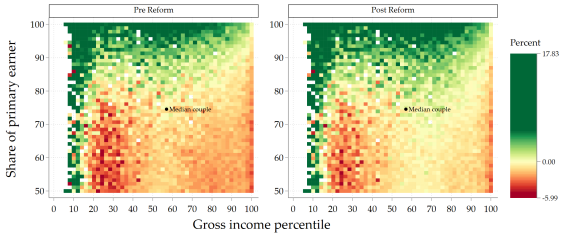
Figure C.26. – Change of Marriage Bonuses and Penalties (cont.)



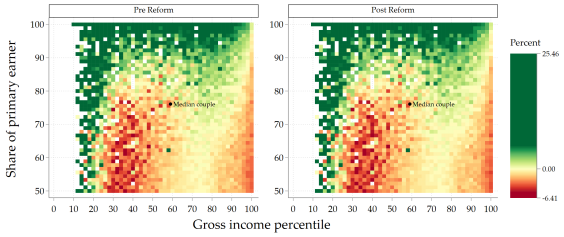
(A) OBRA93



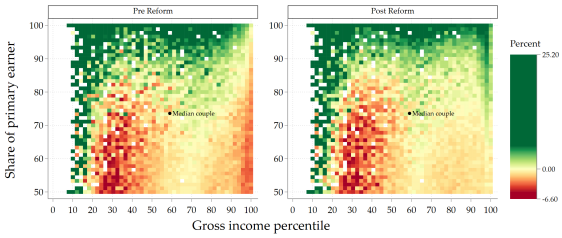
(B) EGTRA01



(C) JGTRRA03



(D) ATRA12



(E) TCJA17

Note: This figure shows marriage bonuses and penalties relative to gross income. Each square in a figure represents an average of marriage bonuses (green) or penalties (red) for a group of tax units at a particular income percentile (horizontal axis) and with a particular primary earner income share (vertical axis). Relative marriage bonuses/penalties relate the absolute monetary advantage from filing as a married couple to the total income of the couple, i.e. $\frac{T_m(y_1+y_2)-(T_s(y_1)+T_s(y_2))}{y_1+y_2}$ (see Appendix C.2.3 for details). The distribution of marriage penalties and bonuses is shown for the pre-reform year (left panel) and the post-reform year (right panel). Income percentiles at the horizontal axis refer to the per capita income distribution of the full sample, i.e. individuals in couples are assigned half of the joint income. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. *Source:* Own calculations based on NBER TAXSIM and CPS-ASEC.

C.3.2 Political Economy

In Section 3.4 of the main text, we summarize the results from an evaluation of past reforms from a political economy perspective. In the baseline scenario, we assume behavioral responses at the intensive and extensive margin. Intensive margin elasticities are displayed in Table C.5. We assume that the extensive margin elasticities are homogeneous across different types of households, but vary across the income distribution. In particular, we assume that participation elasticities decrease with income from 0.65 to 0.25 until the 90th percentile of the gross income distribution, and stays constant in the top decile (see Figure C.18). In the baseline analysis, we also assume that revenue is redistributed lump-sum at the tax unit level.

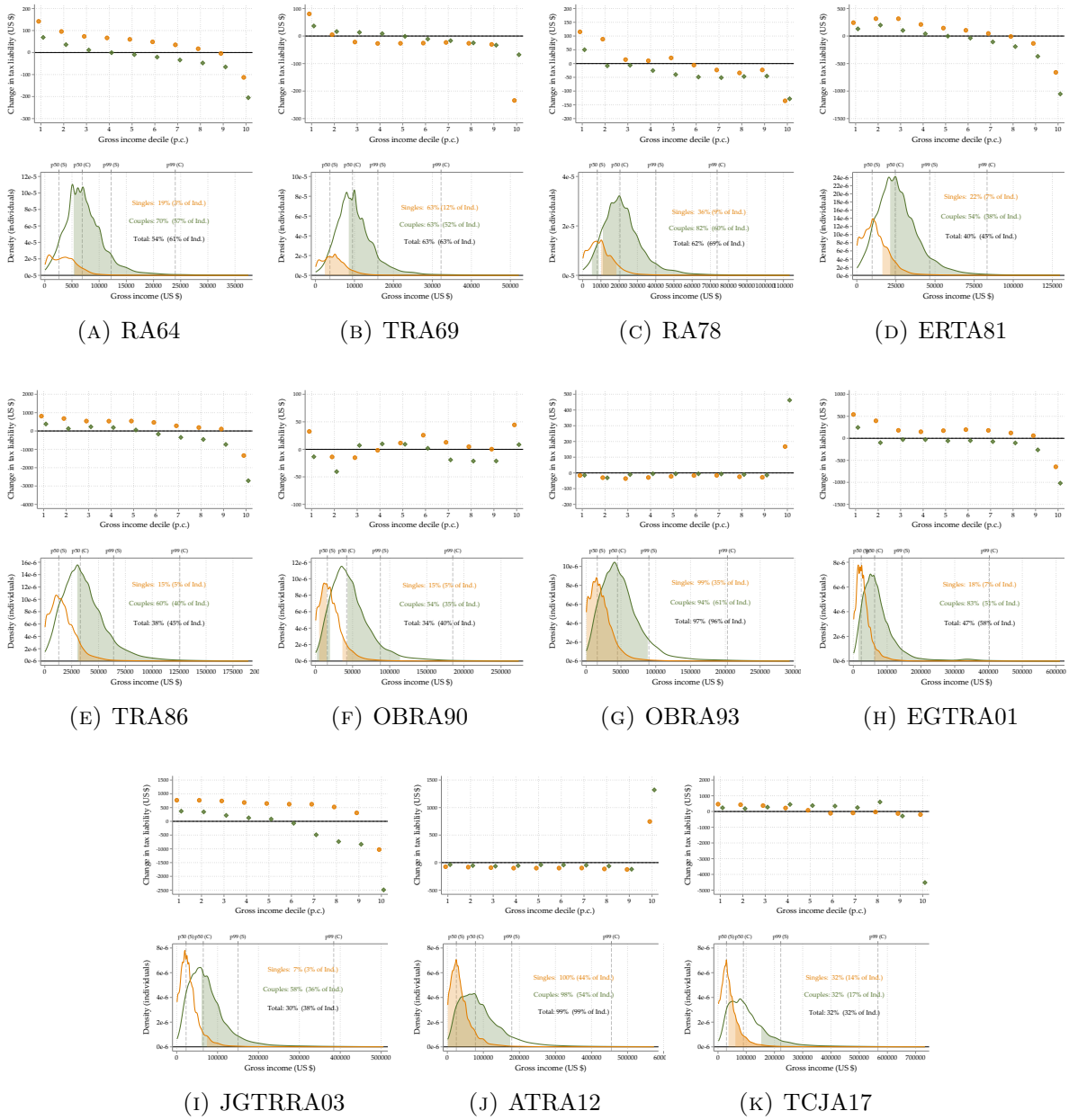
In this part of the appendix, we illustrate the political support for all reforms (see Figure C.27), and illustrate how political support in different groups changes under different assumptions about behavioral responses (see Table C.6). Finally, we also replicate the political economy analysis under the assumption that tax revenue is redistributed lump-sum at the individual level in per-capita terms (see Figure C.28) and compare the corresponding levels of political support to the baseline analysis (see Figure C.29).

Table C.5. – Assumptions about Labor Supply Elasticities

	Single	Couples	
		Prim. Earner	Sec. Earner
Low Elasticity Scenario	0.25	0.15	0.35
Baseline Elasticity Scenario	0.5	0.25	0.75
High Elasticity Scenario	1	0.5	1.5

Note: This table displays our assumptions about the labor supply elasticities for singles, as well as for primary and secondary earners in couples. Assumptions are guided by the range of estimates found in the literature, e.g. Gustafsson (1992), Blundell and Macurdy (1999), Blau and Kahn (2007), Eissa and Hoynes (2004), LaLumia (2008), Kaygusuz (2010), Saez et al. (2012), Bargain et al. (2014), and Neisser (2021).

Figure C.27. – Political Feasibility



(A) RA64 (B) TRA69 (C) RA78 (D) ERTA81

(E) TRA86 (F) OBRA90 (G) OBRA93 (H) EGTRA01

(I) JGTRRA03 (J) ATRA12 (K) TCJA17

Note: This figure replicates Figure 3.4 for all reforms. It shows the change in the tax liability (upper panel) and winners of the reform (lower panel) for singles (orange shaded area) and couples (green shaded area). The change in tax liability represents the average change in tax liability per capita (PC) for each of the 10 per capita gross income deciles. The location of winners and losers across the gross income distribution are identified by evaluating the tax liability change for each of the 25 gross income quantiles. The share of winners among tax units and the corresponding share of individuals is shown next to the distribution. We account for behavioral responses at both the intensive margin (baseline elasticity scenario from Table C.5) and the extensive margin. It is assumed that tax revenues are rebated lump sum at the tax unit level. Figure C.28 shows an alternative analysis based on lump-sum adjustments at the individual level. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Table C.6. – Political Economy of Past Reforms (Part 1)

1. Reform	2. Behavioral Responses		3. Couple Shares		4. Winners (TU)			5. Winners (Ind.)		
	Extensive	Intensive	TU	Ind.	Singles	Couples	All	Singles	Couples	All
RA64	Yes	low	68.9%	81.6%	16.7%	63.6%	49.0%	3.1%	51.9%	55.0%
	Yes	baseline	68.9%	81.6%	18.4%	69.8%	53.8%	3.4%	56.9%	60.3%
	Yes	high	68.9%	81.6%	26.0%	75.0%	59.8%	4.8%	61.2%	66.0%
	No	low	68.9%	81.6%	15.5%	59.1%	45.5%	2.9%	48.2%	51.0%
	No	baseline	68.9%	81.6%	16.6%	63.0%	48.6%	3.1%	51.4%	54.5%
	No	high	68.9%	81.6%	19.5%	71.9%	55.6%	3.6%	58.7%	62.3%
TRA69	Yes	low	68.7%	81.4%	60.8%	58.5%	59.2%	11.3%	47.7%	58.9%
	Yes	baseline	68.7%	81.4%	62.6%	63.0%	62.9%	11.6%	51.3%	62.9%
	Yes	high	68.7%	81.4%	64.7%	72.6%	70.1%	12.0%	59.1%	71.1%
	No	low	68.7%	81.4%	58.9%	54.1%	55.6%	10.9%	44.1%	55.0%
	No	baseline	68.7%	81.4%	60.8%	58.4%	59.2%	11.3%	47.6%	58.9%
	No	high	68.7%	81.4%	63.3%	69.0%	67.2%	11.8%	56.2%	67.9%
RA78	Yes	low	58.2%	73.6%	29.6%	76.5%	56.9%	7.8%	56.3%	64.1%
	Yes	baseline	58.2%	73.6%	34.9%	80.9%	61.7%	9.2%	59.6%	68.8%
	Yes	high	58.2%	73.6%	46.1%	87.3%	70.1%	12.2%	64.3%	76.5%
	No	low	58.2%	73.6%	21.8%	68.6%	49.0%	5.8%	50.5%	56.2%
	No	baseline	58.2%	73.6%	29.8%	76.9%	57.2%	7.9%	56.6%	64.5%
	No	high	58.2%	73.6%	40.3%	84.9%	66.3%	10.6%	62.5%	73.1%
ERTA81	Yes	low	55.0%	71.0%	19.5%	48.8%	35.6%	5.7%	34.6%	40.3%
	Yes	baseline	55.0%	71.0%	22.2%	53.6%	39.5%	6.4%	38.1%	44.5%
	Yes	high	55.0%	71.0%	37.2%	65.8%	52.9%	10.8%	46.7%	57.5%
	No	low	55.0%	71.0%	18.6%	46.5%	33.9%	5.4%	33.0%	38.4%
	No	baseline	55.0%	71.0%	21.0%	50.9%	37.4%	6.1%	36.1%	42.2%
	No	high	55.0%	71.0%	33.7%	61.6%	49.0%	9.8%	43.7%	53.5%
TRA86	Yes	low	50.8%	67.4%	10.5%	54.4%	32.8%	3.4%	36.7%	40.1%
	Yes	baseline	50.8%	67.4%	13.4%	58.0%	36.0%	4.4%	39.1%	43.4%
	Yes	high	50.8%	67.4%	20.2%	66.5%	43.7%	6.6%	44.8%	51.4%
	No	low	50.8%	67.4%	9.5%	52.5%	31.3%	3.1%	35.4%	38.5%
	No	baseline	50.8%	67.4%	12.1%	55.9%	34.3%	4.0%	37.7%	41.6%
	No	high	50.8%	67.4%	17.1%	63.5%	40.7%	5.6%	42.8%	48.4%
OBRA90	Yes	low	48.1%	64.9%	15.4%	53.5%	33.7%	5.4%	34.7%	40.1%
	Yes	baseline	48.1%	64.9%	15.4%	53.5%	33.7%	5.4%	34.7%	40.1%
	Yes	high	48.1%	64.9%	15.4%	53.5%	33.7%	5.4%	34.8%	40.2%
	No	low	48.1%	64.9%	15.4%	53.5%	33.7%	5.4%	34.7%	40.1%
	No	baseline	48.1%	64.9%	15.4%	53.5%	33.7%	5.4%	34.8%	40.1%
	No	high	48.1%	64.9%	15.4%	53.7%	33.8%	5.4%	34.9%	40.3%

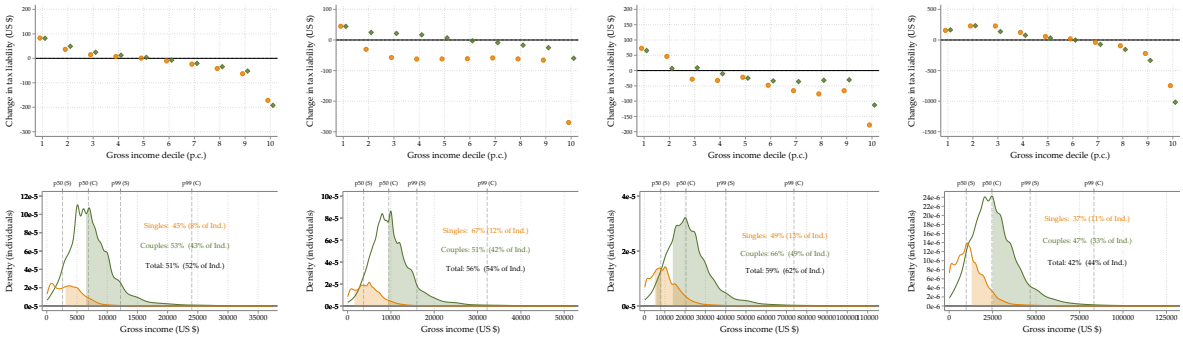
Notes: This table shows whether there was majority support for past reforms of the US federal income tax (column 1) under different assumptions regarding behavioral responses (column 2). Column 3 shows the share of married couples among all tax units and the share of individuals living in married couples. Column 4 shows the share of winners among tax units while column 5 shows the share of winners among individuals. Lump-sum adjustments are at the tax unit level. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. *Source:* Authors' calculations based on NBER TAXSIM and CPS-ASEC.

Table C.7. – Political Economy of Past Reforms (Part 2)

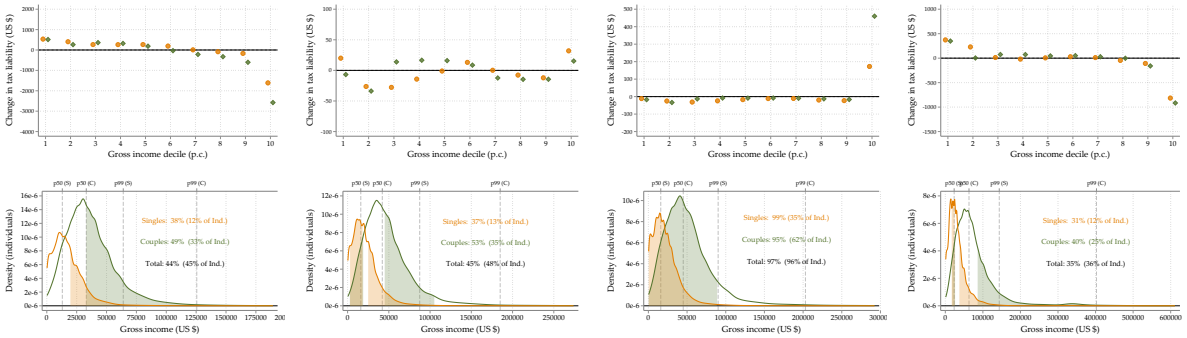
1. Reform	2. Behavioral Responses		3. Couple Shares		4. Winners (TU)			5. Winners (Ind.)		
	Extensive	Intensive	TU	Ind.	Singles	Couples	All	Singles	Couples	All
OBRA93	Yes	low	48.1%	65.0%	99.1%	95.4%	97.3%	34.7%	62.0%	96.7%
	Yes	baseline	48.1%	65.0%	99.0%	93.8%	96.5%	34.7%	61.0%	95.6%
	Yes	high	48.1%	65.0%	29.0%	34.4%	31.6%	10.1%	22.3%	32.5%
	No	low	48.1%	65.0%	99.2%	95.6%	97.4%	34.7%	62.1%	96.8%
	No	baseline	48.1%	65.0%	99.1%	95.1%	97.2%	34.7%	61.8%	96.5%
	No	high	48.1%	65.0%	29.7%	34.9%	32.2%	10.4%	22.7%	33.1%
EGTRRA01	Yes	low	45.0%	62.0%	16.0%	80.8%	45.1%	6.1%	50.1%	56.2%
	Yes	baseline	45.0%	62.0%	16.9%	81.8%	46.1%	6.4%	50.7%	57.1%
	Yes	high	45.0%	62.0%	19.0%	83.2%	47.9%	7.2%	51.6%	58.8%
	No	low	45.0%	62.0%	13.4%	74.3%	40.8%	5.1%	46.1%	51.2%
	No	baseline	45.0%	62.0%	15.9%	80.6%	45.0%	6.0%	50.0%	56.1%
	No	high	45.0%	62.0%	17.9%	82.3%	46.8%	6.8%	51.0%	57.8%
JGTRRA03	Yes	low	44.5%	61.6%	6.6%	56.3%	28.7%	2.5%	34.7%	37.2%
	Yes	baseline	44.5%	61.6%	6.9%	56.9%	29.2%	2.7%	35.0%	37.7%
	Yes	high	44.5%	61.6%	9.0%	61.1%	32.2%	3.4%	37.6%	41.1%
	No	low	44.5%	61.6%	6.4%	55.6%	28.3%	2.5%	34.2%	36.7%
	No	baseline	44.5%	61.6%	6.6%	56.3%	28.7%	2.5%	34.7%	37.2%
	No	high	44.5%	61.6%	8.0%	60.4%	31.3%	3.1%	37.2%	40.2%
ATRA12	Yes	low	38.4%	55.5%	99.6%	98.3%	99.1%	44.3%	54.5%	98.9%
	Yes	baseline	38.4%	55.5%	99.6%	97.9%	98.9%	44.3%	54.3%	98.7%
	Yes	high	38.4%	55.5%	99.3%	95.5%	97.8%	44.2%	53.0%	97.2%
	No	low	38.4%	55.5%	99.7%	98.4%	99.2%	44.3%	54.6%	98.9%
	No	baseline	38.4%	55.5%	99.6%	98.2%	99.1%	44.3%	54.5%	98.9%
	No	high	38.4%	55.5%	99.5%	96.7%	98.4%	44.3%	53.7%	98.0%
TCJA17	Yes	low	37.8%	54.9%	27.9%	29.9%	28.7%	12.6%	16.4%	29.0%
	Yes	baseline	37.8%	54.9%	31.1%	31.4%	31.2%	14.0%	17.2%	31.3%
	Yes	high	37.8%	54.9%	38.6%	36.4%	37.8%	17.4%	20.0%	37.4%
	No	low	37.8%	54.9%	27.1%	29.2%	27.9%	12.2%	16.0%	28.3%
	No	baseline	37.8%	54.9%	29.3%	30.7%	29.8%	13.2%	16.8%	30.1%
	No	high	37.8%	54.9%	37.1%	35.5%	36.5%	16.7%	19.5%	36.3%

Notes: This table shows the majority support for past reforms of the US federal income tax (column 1) under different assumptions regarding behavioral responses (column 2). Column 3 shows the share of married couples among all tax units and the share of individuals living in married couples. Column 4 shows the share of winners among tax units while column 5 shows the share of winners among individuals. Lump-sum adjustments are at the tax unit level. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Authors' calculations based on NBER TAXSIM and CPS-ASEC.

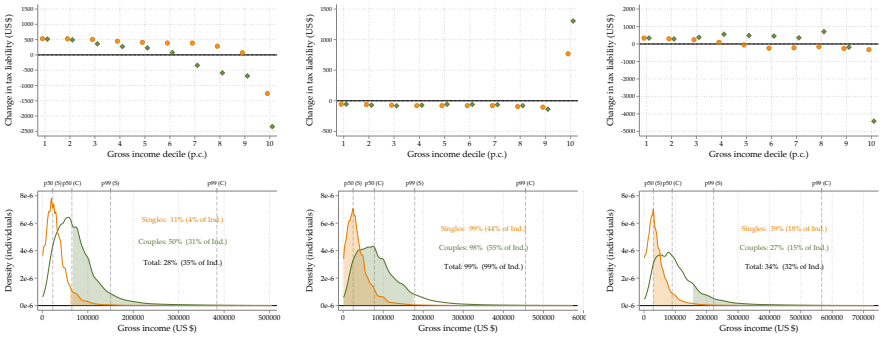
Figure C.28. – Political Feasibility, Lump-Sum Adjustment Per-Capita



(A) RA64 (B) TRA69 (C) RA78 (D) ERTA81



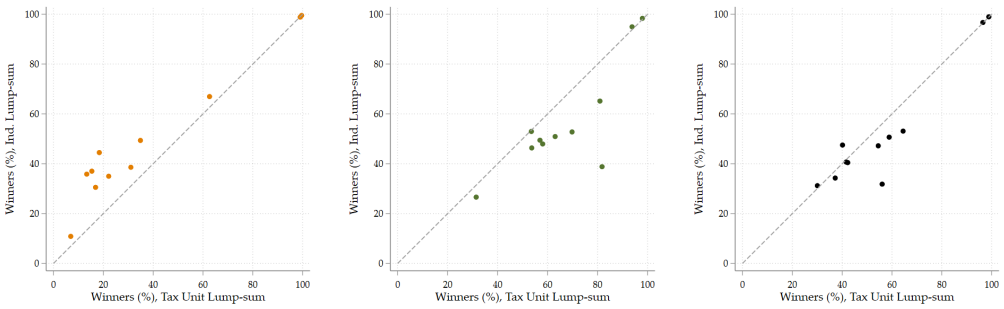
(E) TRA86 (F) OBRA90 (G) OBRA93 (H) EGTRA01



(I) JGTRRA03 (J) ATRA12 (K) TCJA17

Note: This figure replicates Figure C.27 using lump-sum adjustment at the individual level instead of the tax unit level. It shows the change in the tax liability (upper panel) and winners of the reform for singles (orange shaded area) and couples (green shaded area). The change in tax liability represents the average change in tax liability per capita (PC) for each of the 10 per capita gross income deciles. The location of winners and losers across the gross income distribution are identified by evaluating the tax liability change for each of the 25 gross income quantiles. The share of winners among tax units and the corresponding share of individuals is shown next to the distribution. We account for behavioral responses at the intensive margin (baseline elasticity scenario from Table C.5) and extensive margin responses. Further, reforms are made revenue neutral by distributing any gains or losses lump-sum. Lump-sum adjustments are implemented at the per capita level. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.29. – Choice of Lump-Sum Adjustment, Political Economy



(A) Singles (B) Couples (C) Aggregate

Note: This figure displays how majority support among singles, couples, and in the aggregate population under lump-sum adjustment at the tax unit level compares to majority support under a per-capita lump-sum adjustment. Every dot represents a specific reform. The figure displays majority support under extensive and intensive margin responses using the baseline elasticity scenario from Table C.5). Detailed graphical analyses on the majority support under tax-unit (individual) lump-sum adjustment are shown in Figure C.27 (C.28). All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

C.3.3 Welfare Implications

In Section 3.4 of the main text, we summarize the results from an evaluation of past reforms from a welfare perspective. Here, we show, which exact specification of social welfare weights are used for the evaluation (see Table C.8), and provide detailed results for all reforms (see Table C.9).

Table C.8. – Welfare Weights for Reforms *in* the System

Welfare weights	Singles	Couples
Equal	$\forall y_s,\ g_s(y_s) = 1$	$\forall y_m,\ g_m(y_m) = 2$
Singles Only	$\forall y_s,\ g_s(y_s) = 1$	$\forall y_m,\ g_m(y_m) = 0$
Single Women Only	$\forall y_s,\ g_s(y_s) = \begin{cases} 1, & \text{for } female \\ 0, & \text{for } male \end{cases}$	$\forall y_m,\ g_m(y_m) = 0$
Couples Only	$\forall y_s,\ g_s(y_s) = 0$	$\forall y_m,\ g_m(y_m) = 2$
Decreasing	$\forall y_s,\ g_s(y_s) = y_s^{-a}$	$\forall y_m,\ g_m(y_m) = 2\left(\frac{y_m}{2}\right)^{-a}$
Rawlsian	$\forall y_s,\ g_s(y_s) = \begin{cases} 1, & \text{for } y_s \leq \text{P} \\ 0, & \text{for } y_s \geq \text{P} \end{cases}$	$\forall y_m,\ g_m(y_m) = \begin{cases} 2, & \text{for } \frac{y_m}{2} \leq \text{P} \\ 0, & \text{for } \frac{y_m}{2} \geq \text{P} \end{cases}$
Rawlsian (Single Only)	$\forall y_s,\ g_s(y_s) = \begin{cases} 1, & \text{for } y_s \leq \text{P} \\ 0, & \text{for } y_s \geq \text{P} \end{cases}$	$\forall y_m,\ g_m(y_m) = 0$
Rawlsian (Single Women Only)	$\forall y_s,\ g_s(y_s) = \begin{cases} 1, & \text{for } y_s \leq \text{P and } female \\ 0, & \text{for } y_s \geq \text{P or } male \end{cases}$	$\forall y_m,\ g_m(y_m) = 0$
Rawlsian (Couples Only)	$\forall y_s,\ g_s(y_s) = 0$	$\forall y_m,\ g_m(y_m) = \begin{cases} 2, & \text{for } \frac{y_m}{2} \leq \text{P} \\ 0, & \text{for } \frac{y_m}{2} \geq \text{P} \end{cases}$
Affirmative Action Feminist	$\forall y_s,\ g_s(y_s) = \begin{cases} 1, & \text{for } female \\ 0, & \text{for } male \end{cases}$	$\forall y_m,\ g_m(y_m) = \frac{y_{woman}}{y_{man}+y_{woman}}$
Rawlsian Affirmative Action Feminist	$\forall y_s,\ g_s(y_s) = \begin{cases} 1, & \text{for female and } y_s \leq \text{P} \\ 0, & \text{for male or } y_s > \text{P} \end{cases}$	$\forall y_m,\ g_m(y_m) = \begin{cases} \frac{y_{woman}}{y_{man}+y_{woman}}, & \text{for } \frac{y_m}{2} \leq \text{P} \\ 0, & \text{for } \frac{y_m}{2} > \text{P} \end{cases}$

Note: This table shows different specifications of welfare weights to evaluate reforms *in* the system. The sum of weights over the whole population is normalized to 1. P refers to specific percentiles of the per capita income distribution and the parameter a is strictly positive.

The "Equal" welfare function assigns equal weights to all individuals. The welfare function "Couples" ("Singles") assigns equal weights to all couples (singles), and none to singles (couples). It is meant to be descriptive, not normatively appealing. Welfare goes up if and only if the social surplus among the respective group (total output minus total effort costs) goes up. These welfare functions are all maximized by a *laissez-faire* outcome without distortionary taxation in the group. Thus, a positive evaluation by such a welfare function indicates that tax distortions have gone down, and a negative evaluation indicates that the tax system has become more distortionary in the considered group. In contrast, the Rawlsian welfare function concentrates welfare weights at the bottom of the income distribution while the Affirmative Action Feminist welfare function concentrates weights on those households with a high female income share. We also include a welfare function

that we refer to as "Rawlsian Affirmative Feminist". This measure concentrates weights at the bottom. For couples at the bottom, moreover, weights increase in the female income share.

Table C.9 contains an evaluation of all tax reforms since the 1960s according to these welfare weights. Some of the reforms, e.g. RA64, TRA86, and ERTA81, reduced distortions in the system: the "Equal" welfare function approves them. It also shows that these reforms are unambiguously rejected by Rawlsian welfare functions and the ones with weights that are decreasing functions of income. Thus, with inequality aversion, the loss of tax revenue trumps the effect that some taxpayers benefit from a tax cut. Possibly, such reform are approved, however, by an Affirmative Feminist social welfare measure, e.g. TRA69 and ERTA81. Reducing distortions in the system may be desirable from the perspective of secondary earners who face high marginal tax rates under joint taxation. However, since the gains from the reduction in the distortion among dual earner couples needs to outweigh the negative effects on single women (about which the Affirmative Action Feminist also cares a lot), strong behavioral responses are needed for an approval of the Affirmative Action Feminist. Some reforms like OBRA93 and ATRA12 show the reverse pattern. By raising marginal tax rates, these reforms increased distortions of the system, and are therefore rejected by the welfare function putting equal weight on all individuals. However, these reforms are unanimously supported by Rawlsian welfare measures, since the distortionary effects are for them outweighed by the increase in tax revenue that benefits poor households.

In Appendix C.1.3 and Appendix C.2.4, we show that there are two different measures of individual welfare that can be used for an evaluation of large tax reforms. Figure C.30 shows that the choice of welfare measure is relevant for assessing the magnitude of the welfare effect, but is not consequential for an assessment of the direction of the welfare effect of a specific reform under a particular set of welfare weights.

Table C.9. – Welfare Implications of Past Reforms (Part 1)

Reform	Welfare Weights	Int. Margin			Ext. + Int. Margin		
		Low	Baseline	High	Low	Baseline	High
RA64	Equal	>	>	>	>	>	>
	Singles Only	<	<	<	<	<	<
	Single Women Only	<	<	<	<	<	<
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	<	<	<	<	<	<
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<
TRA69	Equal	>	>	>	>	>	>
	Singles Only	>	>	>	>	>	>
	Single Women Only	<	<	<	<	<	<
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	<	<	<	<	<	<
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	<	<	>	<	<	>
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<
RA78	Equal	>	>	>	>	>	>
	Singles Only	<	<	<	<	<	>
	Single Women Only	<	<	<	<	<	<
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	<	<	<	<	<	<
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<
ERTA81	Equal	>	>	>	>	>	>
	Singles Only	<	<	<	<	<	>
	Single Women Only	<	<	<	<	<	<
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	<	<	<	<	<	<
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	<	<	>	<	>	>
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<

Notes: This table shows the welfare implications under different welfare weights for past reforms of the US federal income tax under different assumptions regarding behavioral responses (extensive + intensive margin, intensive margin only), and different scenarios for the intensive margin elasticity (see Table 3.1). Lump sum adjustments have been implemented on a per-tax-unit basis. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Authors' calculations based on NBER TAXSIM and CPS-ASEC.

Table C.10. – Welfare Implications of Past Reforms (Part 2)

Reform	Welfare Weights	Int. Margin			Ext. + Int. Margin		
		Low	Baseline	High	Low	Baseline	High
TRA86	Equal	>	>	>	>	>	>
	Singles Only	<	<	<	<	<	<
	Single Women Only	<	<	<	<	<	<
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	<	<	<	<	<	<
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	<	<	>	<	<	>
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<
OBRA90	Equal	>	>	>	>	>	>
	Singles Only	<	<	<	<	<	<
	Single Women Only	>	>	>	>	>	>
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	>	>	>	>	>	>
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	>	>	>	>	>	>
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<
OBRA93	Equal	<	<	<	<	<	<
	Singles Only	>	>	<	>	>	<
	Single Women Only	>	>	<	>	>	<
	Couples Only	<	<	<	<	<	<
	Decreasing, a=.8	>	>	<	>	>	<
	Decreasing, a=2	>	>	<	>	>	<
	Rawlsian, p10	>	>	<	>	>	<
	Rawlsian, p5	>	>	<	>	>	<
	Rawlsian (Single Only), p10	>	>	<	>	>	<
	Rawlsian (Single Only), p5	>	>	<	>	>	<
	Rawlsian (Single Woman Only), p10	>	>	<	>	>	<
	Rawlsian (Single Woman Only), p5	>	>	<	>	>	<
	Rawlsian (Couples Only), p10	>	>	>	>	>	>
	Rawlsian (Couples Only), p5	>	>	<	>	>	<
	Affirmative Action Feminist	>	>	<	>	>	<
	Rawlsian Affirmative Action Feminist, p10	>	>	<	>	>	<
	Rawlsian Affirmative Action Feminist, p5	>	>	<	>	>	<
EGTRRA01	Equal	>	>	>	>	>	>
	Singles Only	<	<	<	<	<	<
	Single Women Only	<	<	<	<	<	<
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	<	<	<	<	<	<
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<

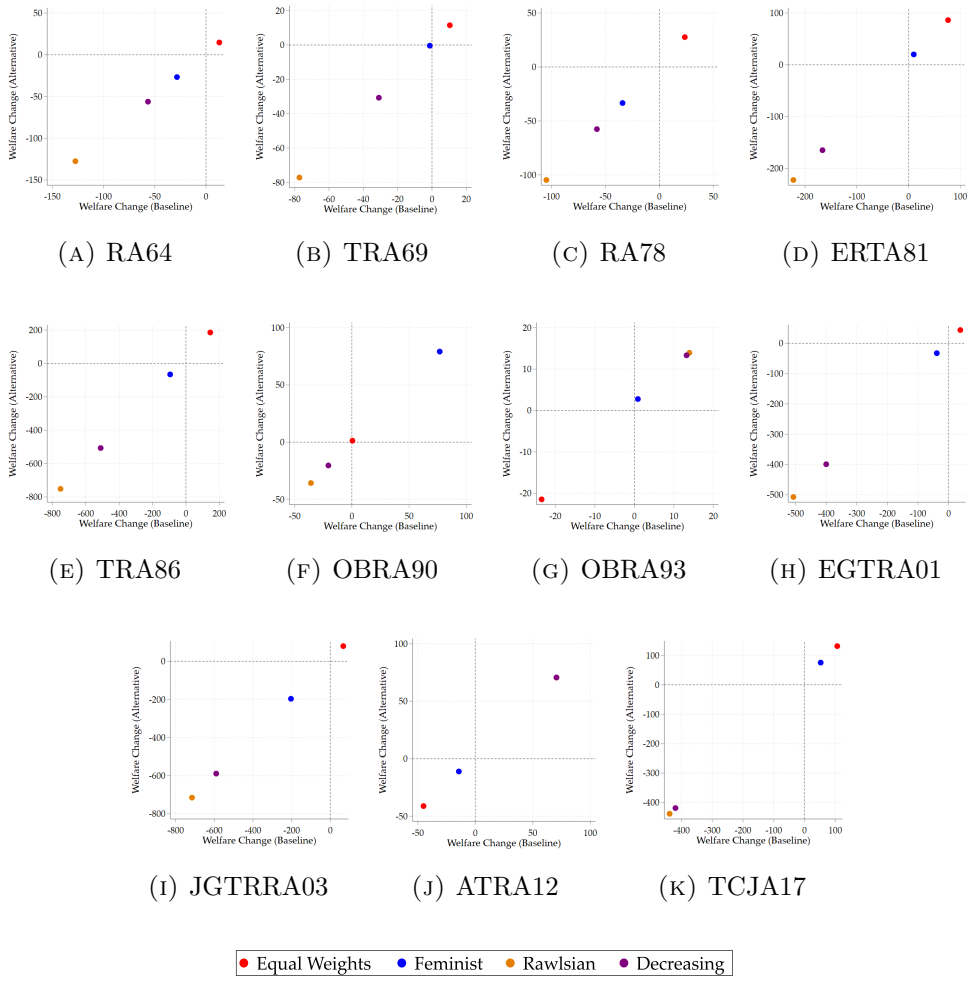
Notes: This table shows the welfare implications under different welfare weights for past reforms of the US federal income tax under different assumptions regarding behavioral responses (extensive + intensive margin, intensive margin only), and different scenarios for the intensive margin elasticity (see Table 3.1). Lump sum adjustments have been implemented on a per-tax-unit basis. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Authors' calculations based on NBER TAXSIM and CPS-ASEC.

Table C.11. – Welfare Implications of Past Reforms (Part 3)

Reform	Welfare Weights	Int. Margin			Ext. + Int. Margin		
		Low	Baseline	High	Low	Baseline	High
JGTRRA03	Equal	>	>	>	>	>	>
	Singles Only	<	<	<	<	<	<
	Single Women Only	<	<	<	<	<	<
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	<	<	<	<	<	<
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<
ATRA12	Equal	<	<	<	<	<	<
	Singles Only	>	>	>	>	>	<
	Single Women Only	>	>	>	>	>	>
	Couples Only	<	<	<	<	<	<
	Decreasing, a=.8	>	>	>	>	>	>
	Decreasing, a=2	>	>	>	>	>	>
	Rawlsian, p10	>	>	>	>	>	>
	Rawlsian, p5	>	>	>	>	>	>
	Rawlsian (Single Only), p10	>	>	>	>	>	>
	Rawlsian (Single Only), p5	>	>	>	>	>	>
	Rawlsian (Single Woman Only), p10	>	>	>	>	>	>
	Rawlsian (Single Woman Only), p5	>	>	>	>	>	>
	Rawlsian (Couples Only), p10	>	>	>	>	>	>
	Rawlsian (Couples Only), p5	>	>	>	>	>	>
	Affirmative Action Feminist	>	>	<	>	<	<
	Rawlsian Affirmative Action Feminist, p10	>	>	>	>	>	>
	Rawlsian Affirmative Action Feminist, p5	>	>	>	>	>	>
TCJA17	Equal	>	>	>	>	>	>
	Singles Only	<	<	<	<	<	<
	Single Women Only	<	<	<	<	<	<
	Couples Only	>	>	>	>	>	>
	Decreasing, a=.8	<	<	<	<	<	<
	Decreasing, a=2	<	<	<	<	<	<
	Rawlsian, p10	<	<	<	<	<	<
	Rawlsian, p5	<	<	<	<	<	<
	Rawlsian (Single Only), p10	<	<	<	<	<	<
	Rawlsian (Single Only), p5	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p10	<	<	<	<	<	<
	Rawlsian (Single Woman Only), p5	<	<	<	<	<	<
	Rawlsian (Couples Only), p10	<	<	<	<	<	<
	Rawlsian (Couples Only), p5	<	<	<	<	<	<
	Affirmative Action Feminist	<	>	>	>	>	>
	Rawlsian Affirmative Action Feminist, p10	<	<	<	<	<	<
	Rawlsian Affirmative Action Feminist, p5	<	<	<	<	<	<

Notes: This table shows the welfare implications under different welfare weights for past reforms of the US federal income tax under different assumptions regarding behavioral responses (extensive + intensive margin, intensive margin only), and different scenarios for the intensive margin elasticity (see Table 3.1). Lump sum adjustments have been implemented on a per-tax-unit basis. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Authors' calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.30. – Welfare, Baseline and Alternative Measure



(A) RA64 (B) TRA69 (C) RA78 (D) ERTA81

(E) TRA86 (F) OBRA90 (G) OBRA93 (H) EGTRA01

(I) JGTRRA03 (J) ATRA12 (K) TCJA17

Note: This figure shows for different welfare measures, how the baseline measure based on equation (C.28) compares to the alternative measure based on equation (C.29). Rawlsian weights are based on p5 while decreasing weights are based on $\alpha = 0.8$. The table shows welfare effects for the case with behavioral responses at the intensive margin (baseline elasticity scenario from Table 3.1) and extensive margin responses. Further, reforms are made revenue neutral by distributing any gains or losses lump-sum. Lump-sum adjustments are implemented at the tax unit level. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

C.4 SUPPLEMENTARY MATERIAL: REFORMS *OF* THE SYSTEM

In Section 3.5 of the main text, we analyze hypothetical reforms towards individual taxation from a political economy perspective and under social welfare considerations. This appendix provides supplementary information on these reforms (Appendix C.4.1) and on how we discipline hypothetical reforms with reference to past reforms (Appendix C.4.2).

C.4.1 Stylized Hypothetical Reforms

In Section 3.5, we assess the political support of a reform towards individual taxation. Figure C.31 replicates the analysis for five-year intervals between 1965 and 2019. It serves as the ingredient to Figure 3.6 that visualizes the evolution of political support over time. We also discuss the welfare consequences of reforms towards individual taxation. The exact specifications of welfare weights is shown in Table C.12. Figure C.33 shows the corresponding welfare evaluations for selected years between 1965 and 2019.

Table C.12. – Welfare Weights for Reforms *of* the System

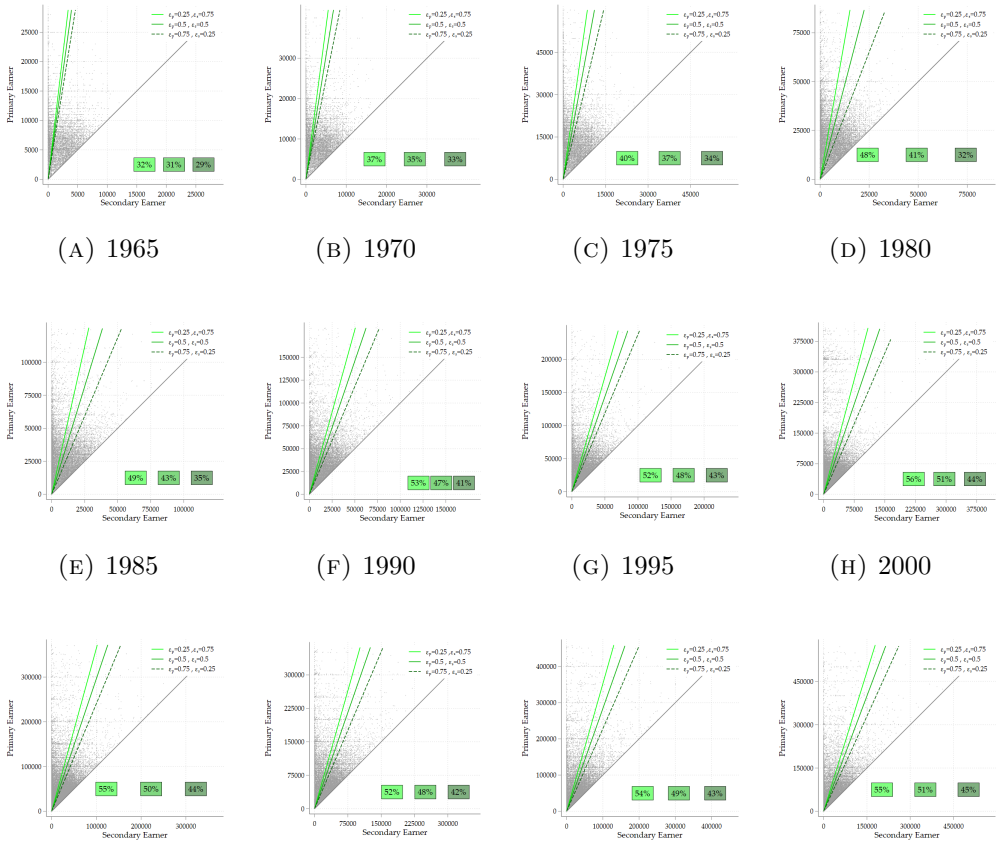
Welfare Weights	Details
Equal (Feminist)	$\forall y_m,\ g_m(y_m) = 1$
Decreasing	$\forall y_m,\ g_m(y_m) = (y_1 + y_2)^{-a}$
Rawlsian	$\forall y_m,\ g_m(y_m) = \begin{cases} 1, & \text{for } y_m \leq \text{P} \\ 0, & \text{for } y_m \geq \text{P} \end{cases}$
Affirmative Action Secondary Earner	$\forall y_m,\ g_m(y_m) = \frac{y_2}{y_m}$
Affirmative Action Feminist	$\forall y_m,\ g_m(y_m) = \frac{y_{woman}}{y_{man}+y_{woman}}$
Rawlsian Affirmative Action Feminist	$\forall y_m,\ g_m(y_m) = \begin{cases} \frac{y_{woman}}{y_{man}+y_{woman}}, & \text{for } y_m \leq \text{P} \\ 0, & \text{for } y_m \geq \text{P} \end{cases}$

Note: This table shows different specifications of welfare weights to evaluate reforms *of* the system. The sum of weights over the whole population of married couples is normalized to 1. P refers to specific percentiles of the couple income distribution and the parameter a is strictly positive. Note that our sample consists also of a small share of same-sex married couples (in 2019 around 0.8 percent of all married couples). While homosexual couples are included for the welfare analysis using Affirmative Action Secondary Earner welfare weights, they are not considered in the analysis using Affirmative Action Feminist welfare weights.

In the main text, we also describe an alternative hypothetical reform towards individual taxation, where marginal tax rates are decreased for all secondary earners, but the increase of marginal tax rates for primary earners is restricted to the upper half of the

couple income distribution. Figure C.34 visualizes the political feasibility and welfare consequences of this reform. In comparison to the baseline, this reform achieves broad political support and unites the welfare concerns for the poor and the working women.

Figure C.31. – Reform towards Individual Taxation, Political Economy



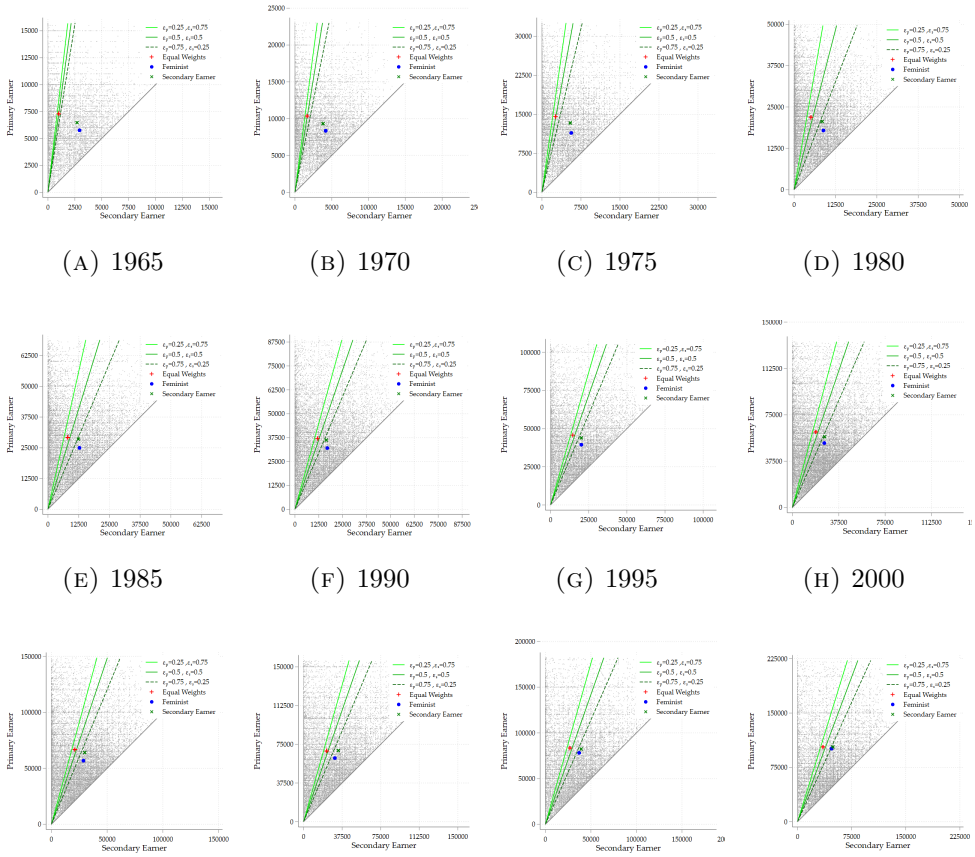
(A) 1965 (B) 1970 (C) 1975 (D) 1980

(E) 1985 (F) 1990 (G) 1995 (H) 2000

(I) 2005 (J) 2010 (K) 2015 (L) 2019

Note: This figure replicates Figure 3.5 for more years. It shows how the political support for a revenue neutral reform towards individual taxation varies with behavioral responses to taxation. Each grey dot represents a couple in the data with specific income of the primary (secondary) earner displayed on the vertical (horizontal) axis. All results are displayed including extensive margin responses. The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). The figure also displays the respective share of couples than benefits from a reform towards individual taxation. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.32. – Reform towards Individual Taxation, Welfare (Middle)



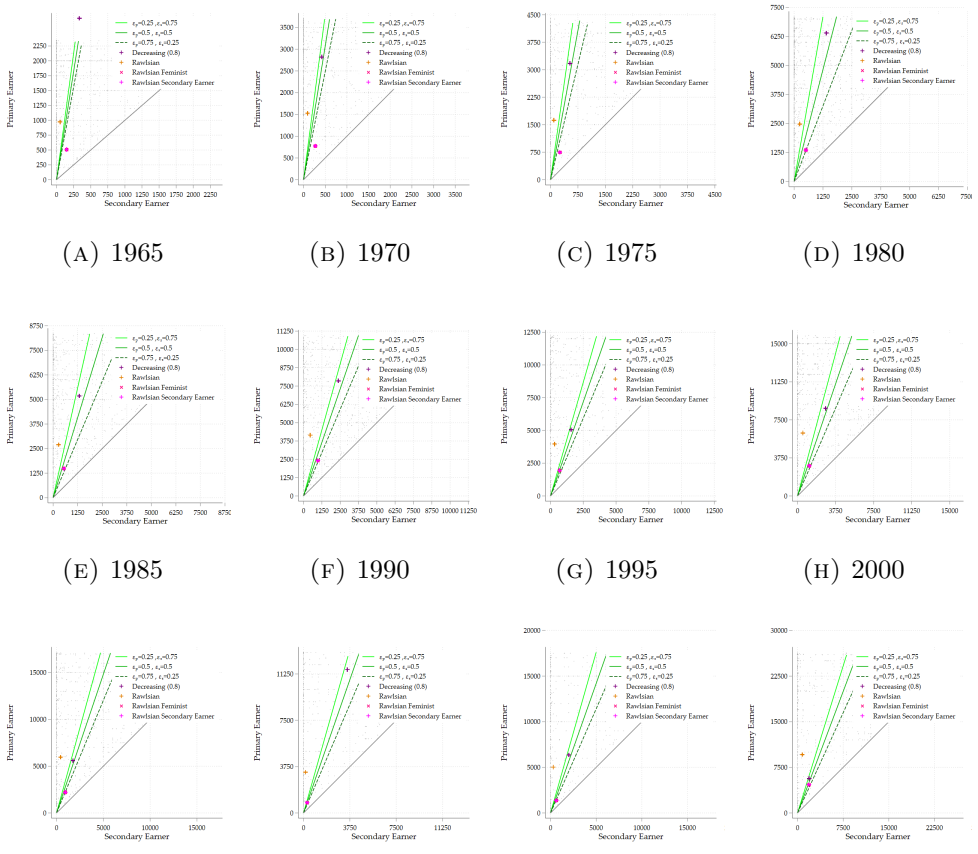
(A) 1965 (B) 1970 (C) 1975 (D) 1980

(E) 1985 (F) 1990 (G) 1995 (H) 2000

(I) 2005 (J) 2010 (K) 2015 (L) 2019

Note: This figure replicates the left panel of Figure 3.8 for more years. It shows how a reform towards individual taxation is evaluated from a welfare perspective under different exogenous welfare weights. It displays welfare implications for welfare weights centered in the middle of the income distribution. Each gray dot represents a couple in the data with specific income of the primary (resp. secondary) earner displayed on the vertical (resp. horizontal) axis. Welfare evaluations with different welfare weights are plotted as a colored dot the location of which is defined via the average welfare-weighted primary earnings (vertical axis) and the average welfare-weighted secondary earnings (horizontal axis). All results are displayed including extensive margin responses. The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). For detailed information on welfare weight specification, see Table C.12. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.33. – Reform towards Individual Taxation, Welfare (Bottom)



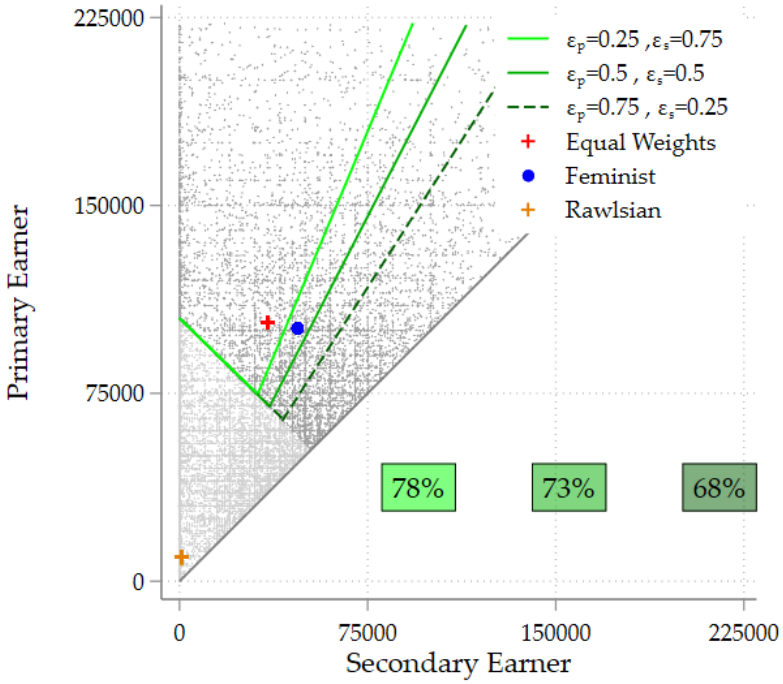
(A) 1965 (B) 1970 (C) 1975 (D) 1980

(E) 1985 (F) 1990 (G) 1995 (H) 2000

(I) 2005 (J) 2010 (K) 2015 (L) 2019

Note: This figure replicates the right panel of Figure 3.8 for more years. It shows how a reform towards individual taxation is evaluated from a welfare perspective under different exogenous welfare weights. It displays welfare implications for welfare weights centered at the bottom of the income distribution. Each gray dot represents a couple in the data with specific income of the primary (resp. secondary) earner displayed on the vertical (resp. horizontal) axis. Welfare evaluations with different welfare weights are plotted as a colored dot the location of which is defined via the average welfare-weighted primary earnings (vertical axis) and the average welfare-weighted secondary earnings (horizontal axis). All results are displayed including extensive margin responses. The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). For detailed information on welfare weight specification, see Table C.12. The specific percentile used for Rawlsian weights is P5 and $a = 0.8$ for decreasing welfare weights. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.34. – Reconciling Rawlsian and Feminist Welfare (2019)



Note: This figure shows for the current tax system, how a partial reform towards individual taxation is evaluated from a welfare perspective under different exogenous welfare weights. A partial reform lowers marginal tax rates for all secondary earners, but raises marginal tax rates only for primary earners above the median of the couple income distribution. Each grey dot represents a couple in the data with specific income of the primary (resp. secondary) earner displayed on the vertical (resp. horizontal) axis. All results are displayed including extensive margin responses. The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). For detailed information on welfare weight specification, see Table C.12. The specific percentile used for Rawlsian weights is P5. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

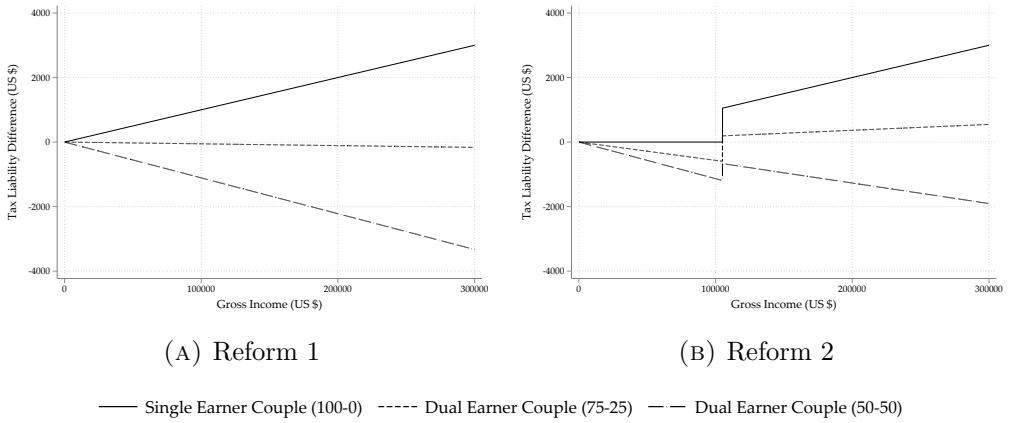
C.4.2 Disciplining Hypothetical Reforms

In Section 3.5 we describe the consequences of hypothetical reforms towards individual taxation, both from a stylized perspective and replicating observed reforms in Sweden and the United Kingdom. Here, we explain the details of the stylized reform and how we adapt our framework to replicate observed reforms towards individual taxation from other countries in the United States.

Under baseline assumptions about behavioral responses shown in Table 3.1, the hypothetical reform towards individual taxation would require a change in primary and secondary tax rates under the ratio $\frac{\tau_2}{\tau_1} = 3.22$ (see Figure 3.5). This implies that a revenue neutral reform that increases the marginal tax rate for primary earners by one percentage point can finance a 3.22 percentage point decrease in secondary earners marginal tax rates. The corresponding hypothetical change in tax liabilities is shown in Figure C.35a.

Figure C.35b depicts the case where we limit the increase in primary earners marginal tax rate to couples in the upper half of the income distribution (see Figure C.34). In this case, the corresponding ratio decreases to $\frac{\tau_2}{\tau_1} = 2.27$, because the reduction in τ_2 that can be financed through increases in τ_1 for rich couples only is now lower.

Figure C.35. – Hypothetical Reforms towards Individual Taxation, US



(A) Reform 1 (B) Reform 2

Note: The figure shows the change in tax liabilities for the hypothetical reforms towards individual taxation across different levels of gross income and for different types of married couples. The reforms increase marginal tax rate on primary earnings by 1 percentage point. The corresponding reduction in secondary earners marginal tax rate is based on the ratio of revenue functions displayed in Equation (C.4). Revenue functions are calculated under both extensive and intensive margin behavioral responses and baseline assumptions about intensive margin elasticities shown in Table 3.1. Figure C.35a shows results for the general reform towards individual taxation while Figure C.35b shows results for the reform, in which marginal tax rate on primary earnings are only increased for couples above the median.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

The two reforms that we discuss are just two exemplary reforms out of infinitely many hypothetical reforms that represent *reforms towards individual taxation*. We therefore now turn to countries that switched to individual taxation in the past.[14] We (i) analyze the shape of the respective reform, and (ii) test whether a reform towards individual taxation of this shape would achieve majority support in the US as of today. We focus on two countries that illustrate the contrast between different types of reforms towards individual taxation: Sweden and the United Kingdom.

Table C.13. – Tax Reforms in OECD Countries

Country	**Year**	**Reform**
Denmark	1970	The unit of taxation changed from the family to individuals.
Sweden	1971	For married couples, earned income was taxed separately under a new unified tax schedule, regardless of marital status.
Austria	1973	Family taxation was replaced by individual taxation.
The Netherlands	1973	Joint taxation was replaced by individual taxation.
Finland	1976	Joint taxation was replaced by individual taxation.
Italy	1977	Joint taxation of husbands and wives was abolished, allowing spouses to be assessed separately.
Belgium	1988	The system changed from aggregated taxation for married couples to individual taxation, allowing the lower-earning spouse to transfer up to 30% of the couple's combined earned income.
Canada	1988	A spousal tax exemption was replaced by a non-refundable tax credit, reducing the system's reliance on joint taxation.
Spain	1988	The joint filing requirement was eliminated, giving married couples the option of filing jointly or separately.
United Kingdom	1990	Joint taxation was replaced by individual taxation.

Source: Authors compilation from OECD Tax Reform Database.

Sweden

Sweden went from joint to individual taxation in 1971. Compared to the US, at the time of the reform, the prevalence of dual-earner couples was already very high. Among married couples with positive gross income, almost 50 percent were dual earner couples with a relatively equal within-couple income distribution while only around 10 percent

[14]For an overview of the tax treatment of couples around the world, see Table C.2. Tax reforms in OECD countries are shown in Table C.13.

were single earner couples (see Table C.14). This is significantly lower than the share of single earner couples in the US today which might explain the political feasibility of the reform.[15]

Table C.14. – Sweden, Distribution of Married Couples, 1969

	100%	75%-100%	50%-75%
Share	9.58%	43.12%	47.3%
Median Income	28,651 SEK	35,639 SEK	43,534 SEK
Mean Income	34,789 SEK	40,793 SEK	46,535 SEK

Note: This table displays the distribution of married couples in Sweden in 1969. We focus on all married couples where both spouses are between 25 and 55 years old. Income shares are calculated based on total gross earnings consisting of labor earnings and self-employment income. Empty income cells reflecting mismatching between the population data and income data are disregarded for the calculation. Among all adult individuals of age 25 to 55, 78% were married persons in 1969.
Source: Special analyses by Daniel Waldenström based on the Swedish population database for the income tax register and population register.

The reform introduced joint taxation, but at the same time modified existing marginal tax rates, tax brackets, and deduction possibilities. In particular, to limit the detrimental consequences for single earner couples, the reform introduced a spousal tax reduction that was phased out with secondary earnings.

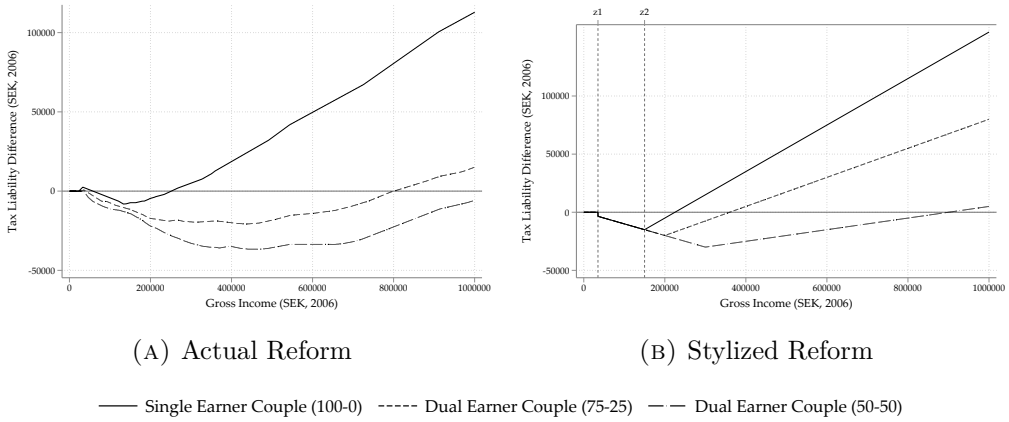
Figure C.36 shows the aggregate impact of the Swedish reform incorporating various elements discussed in Selin (2014).[16] The Swedish reform resembles more the hypothetical reform in C.35b, since the negative consequences of increased marginal tax rates for primary earners only became relevant at the upper part of the income distribution.

To represent the complexities of the Swedish reform such that it can be applied to the United States, we provide a simplified, but accurate representation of the reform in Figure C.36b. It captures three key ingredients of the reform. First, for very poor couples with joint income below the first dashed line, there are no changes in marginal tax rates. Second, after the first dashed line, couples' tax liability decreases, both for single earner

[15]See Gustafsson (1992) for details and political context of the reform at the time. Gunnarsson and Eriksson (2017) discuss policy implications from this reform and conclude that "a majority of taxpayers needs to gain from a reform" as their most important policy lesson learned.

[16]We account for changes in the tax schedule, changes in local taxes, changes in the standard deduction, the deduction for work, and the introduction of the spousal and special tax reduction. We abstract, however, from changes in the treatment of owner occupied housing, interest rate expenses and the sickness insurance fee. We also abstract from changes in public transfers, most notably housing allowances and child allowances. Following Selin (2014), we compare joint filing under the tax system of 1969 to separate filing in 1975.

Figure C.36. – Sweden's Tax Reform of 1971



(A) Actual Reform (B) Stylized Reform

Note: The figure shows the change in tax liabilities for the 1971 reform to individual taxation in Sweden across different levels of gross income and for different types of married couples. We follow the description of the reform by Selin (2014) and compare joint filing before the reform in 1969 to separate filing after the reform in 1975. Figure C.36a shows the impact of the reform calculated based on the detailed specifications in the tax law while Figure C.36b replicates the reform in a stylized manner according to equations (C.38) and (C.39) and for $\tau_1 = 0.2, \tau_2 = 0.1, z_1 = 35000, z_2 = 150000$.
Source: Own calculations based on Selin (2014).

couples and for dual earner couples. Finally, the tax liability increases again, first for single earner couples and then for couples with a more equal income distribution. In particular, the reform can be described as follows

$$T^1 - T^0 = \begin{cases} 0 & \text{if } y_1 + y_2 \leq z_1, \\ -\tau_2 y_1 - \tau_2 y_2 & \text{if } y_1 + y_2 > z_1, y_1 \leq z_2, \\ \tau_1 y_1 - \tau_2 y_2 & \text{if } y_1 + y_2 > z_1, y_1 > z_2, \end{cases} \tag{C.38}$$

with $\tau_1 > 0$ and $\tau_2 > 0$ describing changes of marginal tax rates, z_1 and z_2 describing the income thresholds above which the changes apply.

Since the reform described by equation (C.38) would imply a discrete jump in the tax liability at the second threshold, the desired reform drawn in Figure C.36b is not revenue neutral. Instead, it needs to raise enough revenue to provide a conditional lump sum transfer l to all couples with primary earnings larger than z_2 to ensure a smooth function. Formally, the conditional lump sum transfer l is defined by

$$l = \begin{cases} -\tau_1 z_2 - \tau_2 z_2 & \text{if } y_1 + y_2 > z_1, y_1 > z_2, \\ 0 \quad \text{otherwise.} \end{cases} \tag{C.39}$$

Revenue Implications. The revenue implications of the reform can be expressed in terms of revenue functions:

$$\int_{\mathbb{K}} \mathcal{R}^2\left(y_2\right) dy_2 \tau_2 + \int_{\mathbb{K}}^{z_2} \mathcal{R}^1\left(y_1\right) dy_1 \tau_2$$
$$+ \int_{\mathbb{K},z_2} \mathcal{R}^1\left(y_1\right) dy_1 \tau_1 - z_2(\tau_1 + \tau_2) \int_{\mathbb{K},z_2} f(y_1) d(y_1). \tag{C.40}$$

All integrals are calculated only for couples in $\mathbb{K}$, where $\mathbb{K}$ is the set of couples with $y_1 + y_2 > z_1$. The first integral covers the revenue loss from decreasing secondary earners marginal tax rates by τ_2 for all couples. The second integral covers the revenue loss from decreasing primary earners marginal tax rate by τ_2 for couples with primary earnings up to z_2. The third integral covers the revenue gain from increasing primary earner marginal tax rates by τ_1 for couples with primary earnings larger than z_2. The final part of the equation captures the revenues necessary to finance the lump sum tax in Equation (C.39) for all couples with primary earnings larger than z_2.

Bringing the Swedish Reform to the US. The stylized Swedish tax reform has four parameters, τ_1, τ_2, z_1, z_2. To bring the Swedish reform to the US, we specify the thresholds z_1, z_2 exogenously by transforming the Swedish threshold values to their US dollar value as of 2019. The first Swedish threshold z_1 lies at around 35,000 SEK (2006 prices) which is equivalent to approximately 5000 USD in 2019 prices. The second Swedish threshold z_2 lies at around 150,000 SEK (2006 prices) which is equivalent to approximately 21,000 USD in 2019 prices. We then assume that the secondary earners marginal tax rate is decreased by 1 percentage point and that the primary earners marginal tax rate is increased by 1.25 percentage point. Any additional revenue loss or gain is redistributed lump-sum to all couples.[17]

Figure 3.9 in the main text shows the results. While the tax reduction at the bottom of the income distribution also applies to single earner couples, the following increase in tax liabilities occurs faster for single earner couples. As a consequence, equal earning dual earner couples still benefit from the reform at relatively high income levels. Figure 3.9b shows that about 63% of couples would benefit from such a reform indicating majority support. The winners are concentrated at the bottom of the income distribution and for couples with equal earnings between the two spouses.

[17]Different values of the primary earners marginal tax rate change will influence the curvature of the tax liability difference and can yield different losses/gains in tax revenue.

United Kingdom

The United Kingdom went from joint to individual taxation in 1990.[18] While the reform left the tax schedule in place and did neither change tax brackets nor marginal tax rates on taxable income, the reform operated through changes in the tax exemptions for which married couples were eligible. Similar to Sweden, the reform cushioned the detrimental consequences of the reform for single earner couples through an additional allowance (tax exemption) for married couples. Figure C.37a displays the reform impact for different types of couples following the reform descriptions in Stephens and Ward-Batts (2004) and comparing the pre reform tax year of 1989/1990 under joint filing to the post reform year 1990/1991 under separate filing.[19] The reform is different compared to both our stylized reforms towards individual taxation and the Swedish reform, since it left τ_1 untouched and only operated effectively through a reduction in τ_2.

Figure C.37b shows the stylized version of the reform that we can apply in our framework to the US. It captures three ingredients of the reform. First, for very poor couples with joint income below the first dashed line, there are no changes in the tax liability. Second, there is a small lump-sum reduction in taxes for all couples between the first and the second dashed line. Finally, after the second dashed line, couples with secondary earnings get a tax reduction. This tax cut occurs faster for couples with a more equal income distribution and at a slower pace for couples with a lower secondary earner share. Importantly, once secondary earnings have reached a specific threshold, there are no further changes in the tax liability. The stylized reform can be described as follows (τ_2 defined positively).

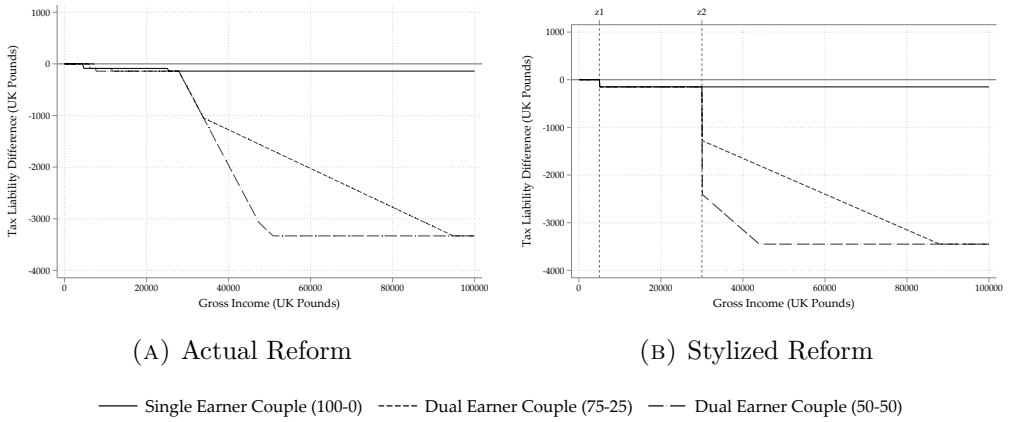
$$T^1 - T^0 = \begin{cases} 0 & \text{if } y_1 + y_2 \leq z_2, \\ -\tau_2 y_2 & \text{if } y_1 + y_2 > z_2, y_2 \leq z_3, \\ 0 & \text{if } y_2 > z_3. \end{cases} \tag{C.41}$$

Since the reform described by equation (C.41) would imply discrete jumps in the tax liability, the reform requires some additional lump sum transfer.

[18] See Stephens and Ward-Batts (2004) for details of the reform and an analysis of it on the intra-household allocation of assets. To the best of our knowledge, there is no direct evaluation of this reform on labor supply. Blundell et al. (1998) use individual taxation along with other income tax reforms in the UK to estimate labor supply of married women.

[19] Comparing these two years has the advantage to isolate the effect of the reform switch from joint to individual taxation from other changes in the tax and transfer system in the 1980s and 1990s. For an overview of changes in the tax and transfer system, see van de Ven and Hérault (2022).

Figure C.37. – United Kingdom's Tax Reform of 1990



(A) Actual Reform (B) Stylized Reform

Note: The figure shows the change in tax liabilities for the 1990 reform to individual taxation in UK across different levels of gross income and for different types of married couples. We follow the description of the reform by Stephens and Ward-Batts (2004) and compare joint filing before the reform in the tax year 1989/1990 to separate filing after the reform in the tax year 1990/1991. Figure C.37a shows the impact of the reform calculated based on the detailed specifications in the tax law while Figure C.37b replicates the reform in a stylized manner according to equations (C.41) and (C.42) and for $\tau_2 = 0.15, l_1 = 150, z_1 = 5000, z_2 = 30000, z_3 = 22000$.
Source: Own calculations based on Stephens and Ward-Batts (2004).

$$l = \begin{cases} 0 & \text{if } y_1 + y_2 \leq z_1, \\ -l_1 & \text{if } y_1 + y_2 > z_1, y_2 \leq z_3, \\ -l_1 - \tau_2 z_3 & \text{if } y_1 + y_2 > z_1, y_2 > z_3. \end{cases} \tag{C.42}$$

Revenue Implications. The revenue implications of the reform can be expressed in terms of revenue functions:

$$\int_{\mathbb{K}}^{z_3} \mathcal{R}^2\left(y_2\right) dy_2 \tau_2 + \int_{\mathbb{L}} -l_1 dy_1 + \tau_2 z_3 \int_{\mathbb{K}, z_3} f(y_2) d(y_2). \tag{C.43}$$

The first integral captures the revenue loss from decreasing secondary earners marginal tax rates by τ_2. This applies to all couples in $\mathbb{K}$, where $\mathbb{K}$ is the set of couples with $y_1 + y_2 > z_2$ and only up to secondary earnings z_3. The second integral covers the revenue loss from the lump sum tax cut for all couples in the set $\mathbb{L}$, where $\mathbb{L}$ is the set of couples with $y_1 + y_2 > z_1$. The third integral covers the revenue loss from the transfer to all couples in $\mathbb{K}$ $(y_1 + y_2 > z_2)$ with secondary earnings larger than z_3.

Bringing the UK Reform to the US. The stylized UK tax reform has four parameters, z_1, z_2, z_3, l_1, and τ_2. To bring the Swedish reform to the US, we specify the thresholds exogenously by transforming the UK threshold values in 1989 to their US dollar value as of 2019 ($z_1 = 15,000$ USD, $z_2 = 92,000$ USD, $z_3 = 68,000$ USD, $l_1 = 500$ USD). We then assume that the secondary earners marginal tax rate is decreased by 5 percentage points. Any additional revenue/loss that the reform generates is redistributed lump sum to all couples.

Figure 3.10 in the main text shows the results. The reform works the opposite way compared to the Swedish reform. In particular, support comes from rich dual earner couples who benefit from the tax cut while poor and single earner couples are made worse off through the loss in tax revenue. Accounting for this loss, the UK type of reform towards individual taxation does not achieve majority support in the US as of today (only 43% in favor).

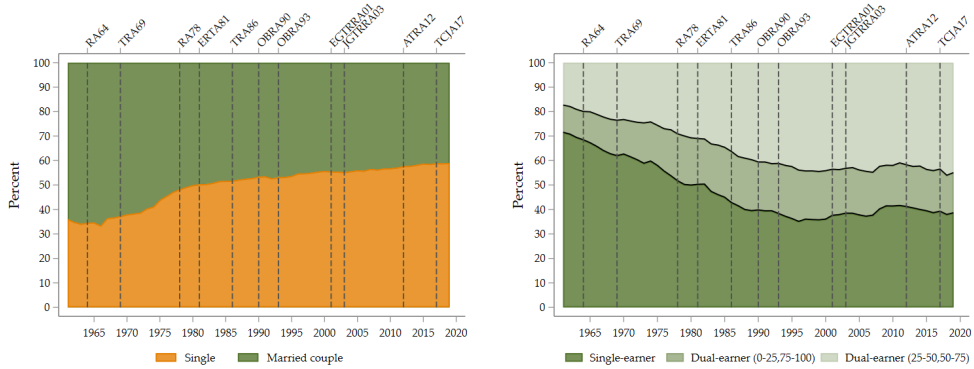
C.5 SUPPLEMENTARY MATERIAL: ALTERNATIVE SAMPLE RESTRICTION

The main analysis focuses on the working age population, i.e. we restrict the sample to tax units with non-negative gross income in which both spouses are between 25 and 55 years old. This sample restriction follows from our model that does not include retirement and education decisions. In addition, since labor force attachment is much lower among old and young groups, our assumptions on behavioral responses to taxation do not apply straightforwardly to these groups.

In this section, as a robustness check, we replicate the figures and tables presented in the main text for an alternative sample restriction in which we consider all adults in tax units with non-negative gross income.[20] The main takeaway from this analysis is that the qualitative properties of our results remain valid. In general, the full population contains more tax units with zero gross income, more singles, and more single-earner couples. The main quantitative differences are based on the latter fact. Given that single-earner couples tend to loose from a reform towards individual taxation, this reform has less support than in our main analysis (47 percent instead of 55 percent for our baseline scenario).

[20]Under this sample restriction, all singles and both spouses in a couple are at least 18 years old.

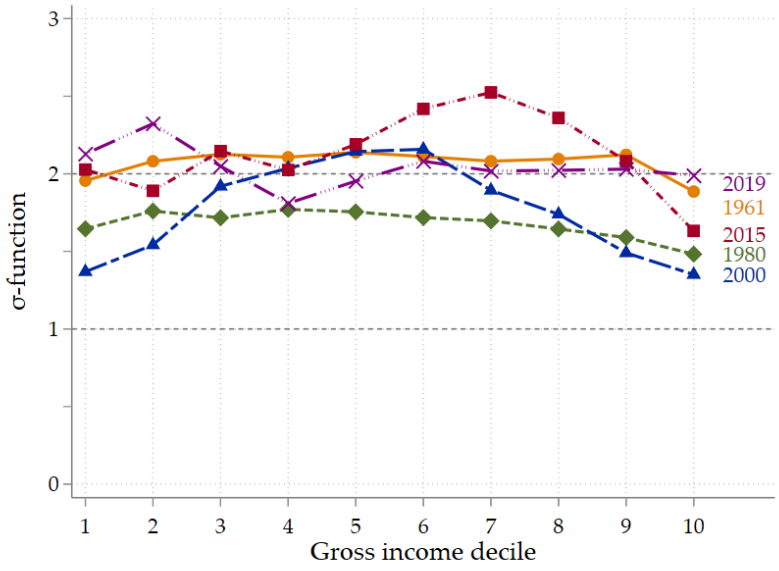
Figure C.38. – Demographic Change, Alternative Sample



(A) Tax Unit Types (B) Couple Types

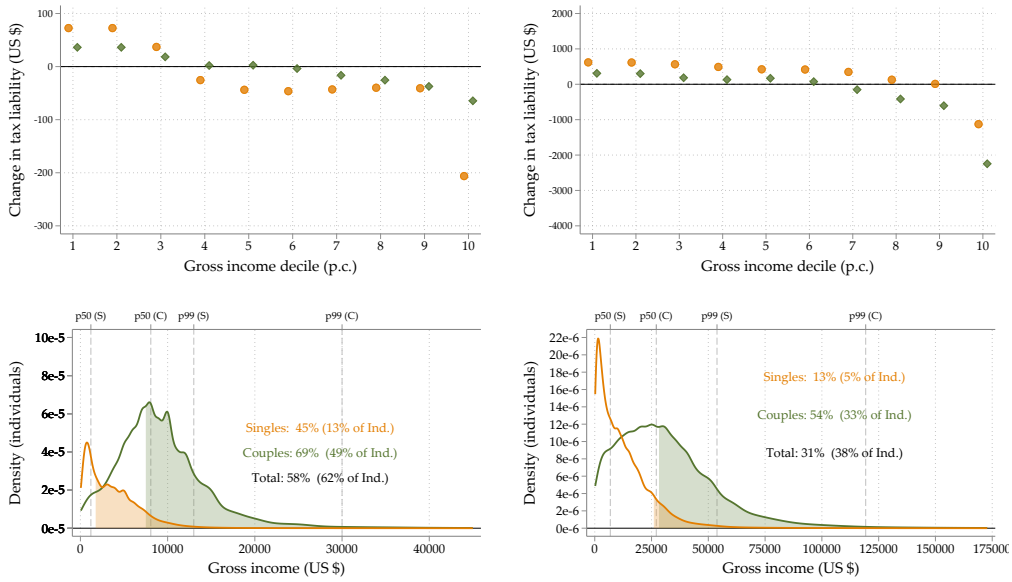
Note: This figure replicates Figure 3.1 for the full adult population instead of the working age population. It shows the distribution of tax unit types over time. Figure C.38a displays the share of single tax units (orange area) and the share of couple tax units (green area). Figure C.38b displays the share of single-earner and dual-earner couples. A single-earner couple refers to a married couple, in which one spouse is not employed (dark green area). The figure further displays the share of dual-earner couples in which both spouses are employed and (i) one spouse earns between 0 and 25 percent (mid green area) and (ii) between 25 and 50 percent of total earnings (light green area). Earnings shares are computed on the basis of wage, business and farm income. Reforms of the federal income tax code as described in Table C.4 are displayed as vertical lines. All estimates are based on tax units with strictly positive gross income in which both spouses are at least 18 years old.
Source: Own calculations based on CPS-ASEC.

Figure C.39. – The Splitting Function σ over Time

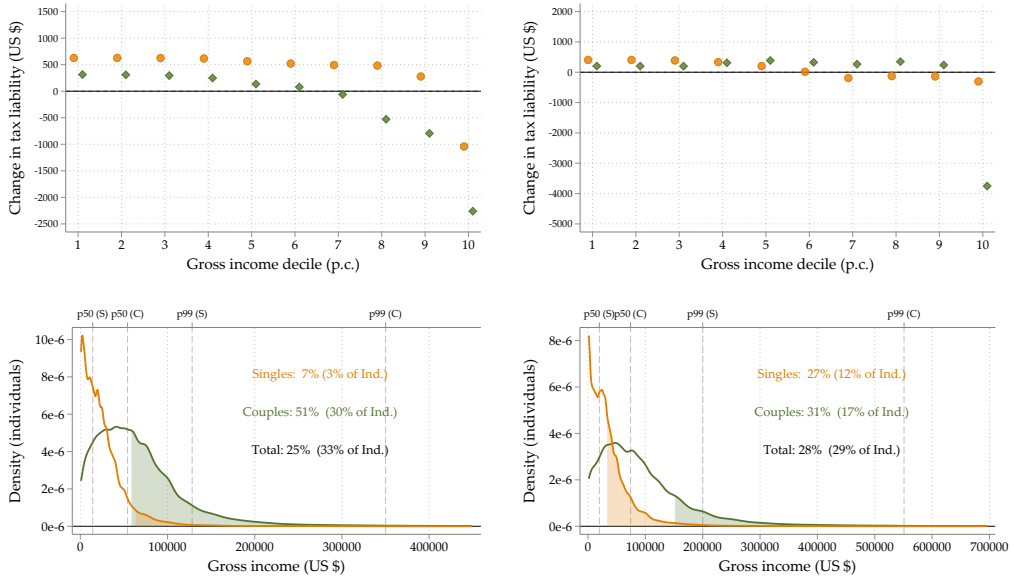


Note: This figure replicates Figure 3.2 for the full adult population instead of the working age population. It shows estimates of the splitting function σ for selected years. The σ-function is calculated by estimating mean average tax rates of couples and singles. Mean average tax rates are used to solve numerically for σ (see Appendix C.2.3). Deciles refer to the gross income distribution of couples in the respective year. All estimates are based on tax units with strictly positive gross income in which both spouses are at least 18 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.40. – Political Feasibility, Alternative Sample



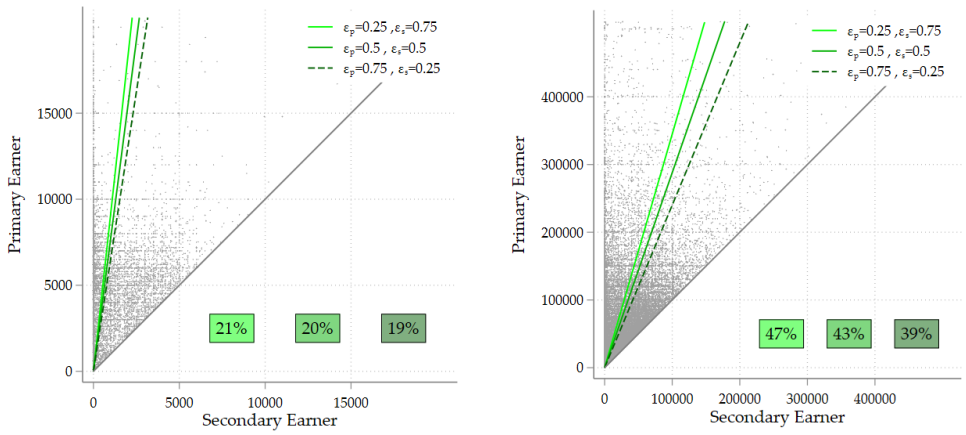
(A) TRA69 (B) TRA86



(C) JGTRRA03 (D) TCJA17

Note: This figure replicates Figure 3.4 for the full adult population instead of the working age population. It shows the change in the tax liability (upper panel) and winners of the reform (lower panel) for singles and couples and for an alternative sample restriction, i.e. the full adult population. The change in tax liability represents the average change in tax liability per capita (PC) for each of the 10 per capita gross income deciles. The location of winners and losers across the gross income distribution are identified by evaluating the tax liability change for each of the 25 gross income quantiles. We account for behavioral responses at the intensive margin (baseline elasticity scenario from Table 3.1) and extensive margin responses. Further, reforms are made revenue neutral by distributing any gains or losses lump-sum. Lump-sum adjustments are implemented at the tax unit level. All estimates are based on tax units with non-negative gross income in which both spouses are at least 18 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

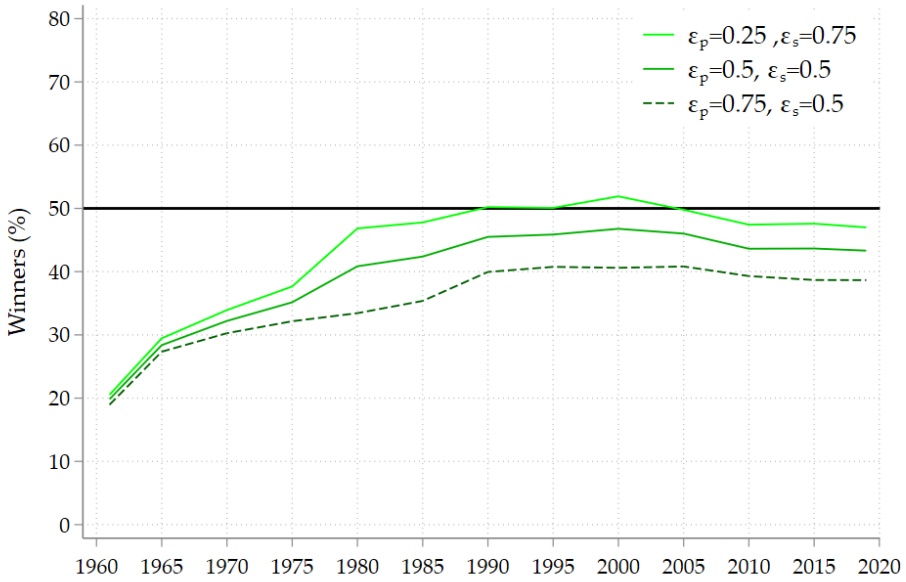
Figure C.41. – Reform towards Individual Taxation: Political Economy, Alternative Sample



(A) 1961 (B) 2019

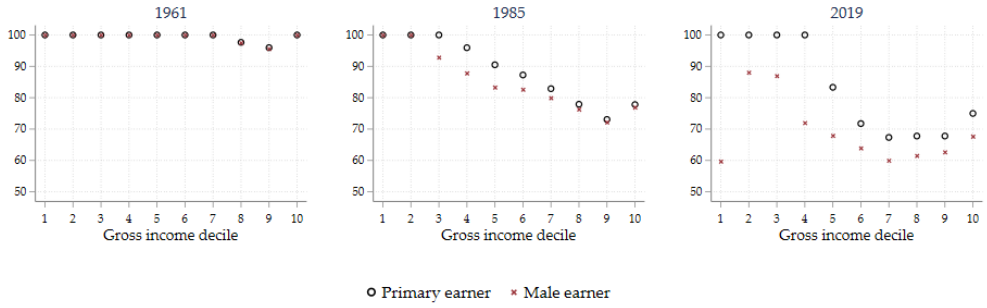
Note: This figure replicates Figure 3.5 for the full adult population instead of the working age population. It shows for 1961 and 2019, how the political support for a revenue neutral reform towards individual taxation varies with behavioral responses to taxation. Each grey dot represents a couple in the data with specific income of the primary (secondary) earner displayed on the vertical (horizontal) axis. All results are displayed including extensive margin responses. The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). The figure also displays the respective share of couples than benefits from a reform towards individual taxation. All estimates are based on tax units with non-negative gross income in which both spouses are at least 18 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.42. – Reform towards Individual Taxation: Share of Winners over Time, Alternative Sample



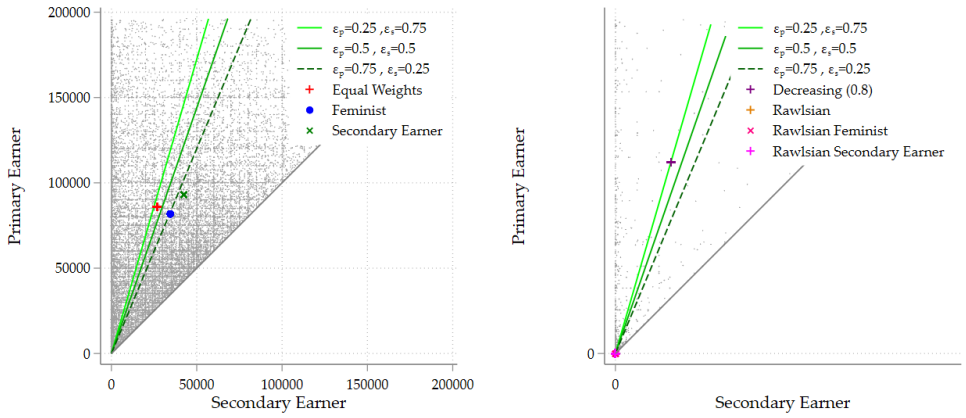
Note: This figure replicates Figure 3.6 for the full adult population instead of the working age population. It shows how the political support for a revenue neutral reform towards individual taxation evolved over time. All results are displayed including extensive margin responses. The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). All estimates are based on tax units with non-negative gross income in which both spouses are at least 18 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.43. – Median Share of Primary and Male Earner, Alternative Sample



Note: This figure replicates Figure 3.7 for the full adult population instead of the working age population. It shows the median income share of the primary earner in the couple by income decile. Earnings shares are computed on the basis of non-negative wage, business and farm income. All estimates are based on tax units with non-negative gross income in which both spouses are at least 18 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

Figure C.44. – Reform towards Individual Taxation: Welfare (2019), Alternative Sample



(A) Welfare i (B) Welfare ii

Note: This figure replicates Figure 3.8 for the full adult population instead of the working age population. It shows for the current tax system, how a reform towards individual taxation is evaluated from a welfare perspective under different exogenous welfare weights. Figure C.44a (C.44b) displays welfare implications for welfare weights centered in the middle (bottom) of the income distribution. Each gray dot represents a couple in the data with specific income of the primary (resp. secondary) earner displayed on the vertical (resp. horizontal) axis. Welfare evaluations with different welfare weights are plotted as a colored dot the location of which is defined via the average welfare-weighted primary earnings (vertical axis) and the average welfare-weighted secondary earnings (horizontal axis). All results are displayed including extensive margin responses. The light green solid line illustrates the result under the baseline elasticity scenario of Table 3.1 in which the primary (resp. secondary) earner has an elasticity of 0.25 (resp. 0.75). For illustrative purposes, the dark green solid line refers to the case where primary and secondary earners' elasticities coincide (0.5) while the dashed green line shows the results under the assumption that the primary earner's elasticity (0.75) is higher than for the secondary earner (0.25). For detailed information on welfare weight specification, see Table C.12. The specific percentile used for Rawlsian weights is P5 and $a = 0.8$ for decreasing welfare weights. All estimates are based on tax units with non-negative gross income in which both spouses are at least 18 years old.
Source: Own calculations based on NBER TAXSIM and CPS-ASEC.

C.6 SUPPLEMENTARY MATERIAL: NARRATIVES ABOUT TAX REFORMS

The particularities of the tax treatment of couples and singles in tax reforms cannot only be observed through an analysis of the implemented tax changes, but also by how narratives among tax reforms in different times were shaped by, e.g., the concern of marriage penalties or treatment of secondary earners. We use word clouds to get a sense of the underlying discussion around three US tax reforms (TRA69, ERTA81 and TCJA17) that led to significant changes in penalties and bonuses. The goal of this analysis is to reconstruct the narrative surrounding these reforms and to uncover to what extent the public discussion at the time is reflected in the effects of the final tax bill on the tax treatment of couples.

We source data from various sources displayed in Table C.15. Among the most prevalent source types are congressional records, newspaper articles, and policy documents prepared by think tanks. The raw data is preprocessed in the following way: first, we reduce the text data to include only letters and hyphens and transform it to lowercase. We then split strings to receive a list of words and remove all stopwords.[21] The remaining words are then reduced to their roots via lemmatization.[22] Finally, we correct for obvious spelling mistakes and compile one data set for each time around a reform.[23]

We construct three different types of word clouds for each of the three reforms. *Type 1* includes raw unigrams (i.e. single words) and bigrams (i.e. collocations of two words). Word clouds of *Type 2* only includes raw bigrams. *Type 3* also focuses on bigrams, but involves further adjustments of the data. In particular, inverted bigrams (e.g. "rate tax" and "tax rate"), synonyms (e.g. "single person", "single individual") are grouped together. In a last step, bigrams referring to institutions or persons as well as doubled terms ("tax tax") are removed.

Across all three reforms, the treatment singles and couples was a prominent topic in public debate about tax reforms. Around TRA69, the arguments mainly circled around the unequal tax treatment of single persons and married couples (see Figure C.45). The displayed "income splitting" procedure in place prior to the reform was perceived to unfairly discriminate against single people (see for example "discrimination single", "burden single", "inequity single"). In later years, the focus shifted towards the concern about unfair

[21] The stopwords document is taken from Lisa Chalaguine's GitHub representation.

[22] In this case, we use the function WordNetLemmatizer.

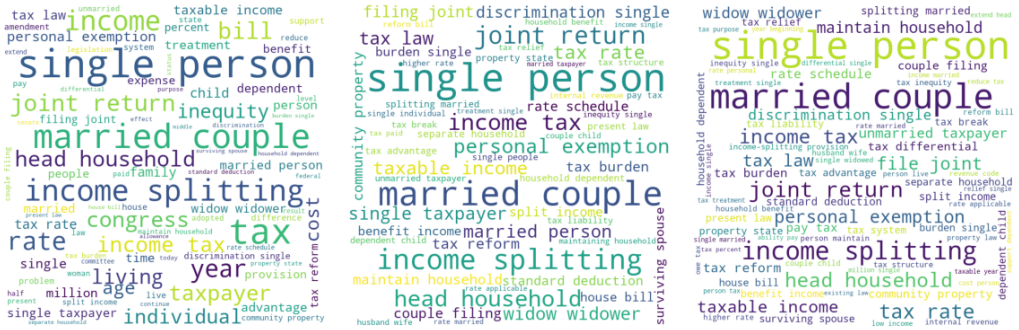
[23] Tables C.16, C.17, and C.18 display the data sources compiled for each reform.

treatment of couples as terms like "marriage penalty" (or "marriage tax penalty") become prominent in the debate (see Figures C.46 and C.47). A lingering concern to "discourage marriage" is also discernible. The term "marriage neutrality" evolves as an objective. In addition, the incentive structure for the second earner, often a married woman, was also an important part of the discussion.

Table C.15. – Text Data Source Types

Type of Source	TRA69	ERTA81	TCJA17
Congressional record	15	17	6
Official	4	8	2
Newspapers	4	87	10
President statement	3	9	2
Journal	3	0	3
Think tank	0	1	40
Blog	0	0	1
Total	29	122	64

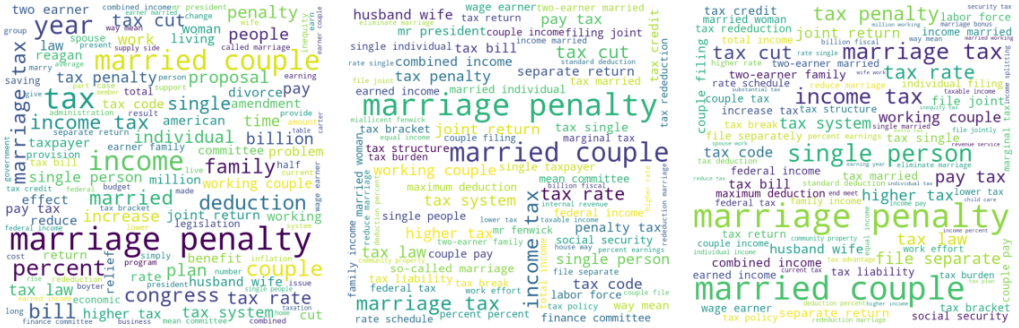
Figure C.45. – Narratives for TRA69



(A) Type 1 (B) Type 2 (C) Type 3

Note: This figure shows narratives around the TRA69. Word clouds have been generated by pre-processing and compiling the textual sources displayed in Table C.16 to one large document and analyzing unigrams and bigrams within this data. Type 1 shows all unigrams and bigrams while Type 2 and Type 3 focus on bigrams only. The word cloud of Type 3 involves a further processing of the data by eliminating synonyms and inverse bigrams.
Source: Own calculations based on sources in Table C.16.

Figure C.46. – Narratives for ERTA81



(A) Type 1 (B) Type 2 (C) Type 3

Note: This figure shows narratives around ERTA81. Word clouds have been generated by pre-processing and compiling the textual sources displayed in Table C.16 to one large document and analyzing unigrams and bigrams within this data. Type 1 shows all unigrams and bigrams while Type 2 and Type 3 focus on bigrams only. The word cloud of Type 3 involves a further processing of the data by eliminating synonyms and inverse bigrams.
Source: Own calculations based on sources in Table C.17.

Figure C.47. – Narratives for TCJA17

(A) Type 1 (B) Type 2 (C) Type 3

Note: This figure shows narratives around TCJA17. Word clouds have been generated by pre-processing and compiling the textual sources displayed in Table C.18 to one large document and analyzing unigrams and bigrams within this data. Type 1 shows all unigrams and bigrams while Type 2 and Type 3 focus on bigrams only. The word cloud of Type 3 involves a further processing of the data by eliminating synonyms and inverse bigrams.
Source: Own calculations based on sources in Table C.16.

Table C.16. – Underlying Text Data for TRA69 (1967–1975)

Type of source	Source name	URL
congressional record	May 8, 1967, 90th Congress, 1st Session, Vol.113, Part 9	Link
congressional record	June 6, 1967, 90th Congress, 1st Session, Vol.113, Part 11	Link
official	Annual Report of the Secretary of the Treasury on the State of the Finances (Fiscal Year Ended June 30, 1968).	Link
newspaper	San Bernardino Sun	Link
President statement	Richard Nixon, Letter to Senate Leaders Mike Mansfield and Hugh Scott on the Tax Reform Bill.	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
official	Tax Reform Act of 1969, Report of the Committee on Finance	Link
congressional record	August 7, 1969, 91st Congress, 1st Session, Vol. 115, Part 17	Link
congressional record	December 10, 1969, 91st Congress, 1st Session, Vol.115, Part 28	Link
congressional record	October 30, 1969, 91st Congress, 1st Session, Vol.115, Part 24	Link
congressional record	August 6, 1969, 91st Congress, 1st Session, Vol. 115, Part 17	Link
congressional record	February 19, 1969, 91st Congress, 1st Session, Vol.115, Part 3	Link
congressional record	September 11, 1969, 91st Congress, 1st Session, Vol.115, Part 19	Link
congressional record	October 2, 1969, 91st Congress, 1st Session, Vol.115, Part 21	Link
congressional record	February 5, 91st Congress, 1st Session, Vol.115, Part 3	Link
congressional record	February 17, 1969, 91st Congress, 1st Session, Vol.115, Part 3	Link
congressional record	December 3, 1969, 91st Congress, 1st Session, Vol.115, Part 27	Link
congressional record	June 5, 1969, 91st Congress, 1st Session, Vol.115, Part 11	Link
congressional record	August 13, 1969, 91st Congress, 1st Session, Vol.115, Part 18	Link
congressional record	April 15, 1969, 91st Congress, 1st Session, Vol.115, Part 7	Link
official	Annual Report of the Secretary of the Treasury on the State of the Finances (Fiscal Year Ended June 30, 1970).	Link
official	General Explanation of the Tax Reform Act of 1969, Joint Committee on Internal Revenue Taxation	Link
newspaper	San Bernardino Sun	Link
journal	Richards, 1970	Link
journal	Richards, 1971	Link
journal	Betz, 1974	Link

Table C.17. – Underlying Text Data for ERTA81 (1978–1985)

Type of Source	Source Name	URL
newspaper	U.S. News and World Report	Link
newspaper	U.S. News and World Report	Link
newspaper	The Associated Press	Link
official	Annual Report of the Secretary of the Treasury on the State of the Finances (Fiscal Year Ended September 30, 1980).	Link
official	Exhibit 51.—Statement of Deputy Assistant Secretary Sunley, August 5, 1980, before the Subcommittee on Taxation and Debt Management of the Senate Finance Committee, on the tax treatment of married and single taxpayers	
President statement	Jimmy Carter tax proposals	Link
newspaper	Columbia Missourian	Link
newspaper	Catholic News Service	Link
newspaper	Desert Sun	Link
newspaper	Catholic News Service	Link
newspaper	The Broomfield Enterprise	Link
newspaper	San Bernardino Sun	Link
newspaper	Bronxville Review Press and Reporter	Link
newspaper	San Bernardino Sun	Link
newspaper	The Stanford Daily	Link
newspaper	San Bernardino Sun	Link

newspaper	Desert Sun	Link
newspaper	Desert Sun	Link
newspaper	Desert Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	Lancaster Farming	Link
newspaper	Santa Cruz Sentinel	Link
newspaper	The Steamboat Pilot	Link
newspaper	Santa Cruz Sentinel	Link
newspaper	Santa Cruz Sentinel	Link
official	Administration'es (Carter) tax proposal	Link
newspaper	U.S. News and World Report	Link
newspaper	The Christian Science Monitor	Link
newspaper	The New York Times	Link
newspaper	The Associated Press	Link
newspaper	U.S. News and World Report	Link
newspaper	The Associated Press	Link
newspaper	The New York Times	Link
newspaper	The Christian Science Monitor	Link
newspaper	The Associated Press	Link
official	Economic Report of the President, 1981	Link
official	General Explanation of the Economic Recovery Tax Act of 1981, Joint Committee on Taxation	Link
President statement	Jimmy Carter, Budget Message Message to the Congress Transmitting the Fiscal Year 1982 Budget.	Link
President statement	Jimmy Carter, Annual Message to the Congress: The Economic Report of the President	Link
President statement	Ronald Reagan, Address Before a Joint Session of the Congress on the Program for Economic Recovery	Link
President statement	Ronald Reagan, Remarks on Federal Tax Reductions Following Meetings With Members of Congress	Link
President statement	Ronald Reagan, The President's News Conference	Link
President statement	Ronald Reagan, Remarks About Federal Tax Reduction Legislation at a Meeting With State Legislators and Local Government Officials	Link
President statement	Ronald Reagan, Address to the Nation on Federal Tax Reduction Legislation	Link
newspaper	Lancaster Farming	Link
newspaper	The Daily Collegian	Link
newspaper	The Daily Collegian	Link
newspaper	Desert Sun	Link
newspaper	Catholic News Service	Link
newspaper	Catholic News Service	Link
newspaper	Desert Sun	Link
newspaper	Catholic News Service	Link
newspaper	Desert Sun	Link
newspaper	Douglas County News-Press	Link
newspaper	Calexico Chronicle	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	Healdsburg Tribune, Enterprise and Scimitar	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	Recorder	Link
newspaper	Desert Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link

newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	San Bernardino Sun	Link
newspaper	Desert Sun	Link
newspaper	Santa Cruz Sentinel	Link
newspaper	Desert Sun	Link
newspaper	Santa Cruz Sentinel	Link
newspaper	Desert Sun	Link
newspaper	Santa Cruz Sentinel	Link
newspaper	Santa Cruz Sentinel	Link
newspaper	Santa Cruz Sentinel	Link
newspaper	Santa Cruz Sentinel	Link
official	Economic Recovery Tax Act of 1981, Report of the Committee on Finance	Link
newspaper	Cherokeean	Link
congressional record	July 18, 1981, 97th Congress, 1st Session, Vol.127, Part 12	Link
congressional record	July 29, 1981, 97th Congress, 1st Session, Vol. 127, Part 14	Link
congressional record	March 11, 1981, 97th Congress, 1st Session, Vol.127, Part 3	Link
congressional record	February 5, 97th Congress, 1st Session, Vol.127, Part 2	Link
congressional record	April 10, 1981, 97th Congress, 1st Session, Vol.127, Part 6	Link
congressional record	July 27, 1981, 97th Congress, 1st Session, Vol.127, Part 13	Link
congressional record	March 24, 1981, 97th Congress, 1st Session, Vol.127, Part 4	Link
congressional record	February 26, 1981, 97th Congress, 1st Session, Vol.127, Part 3	Link
congressional record	July 15, 1981, 97th Congress, 1st Session, Vol.127, Part 12	Link
congressional record	January 6, 1981, 97th Congress, 1st Session, Vol. 127, Part 1	Link
congressional record	January 20, 1981, 97th Congress, 1st Session, Vol. 127, Part 1	Link
congressional record	August 3, 1981, 97th Congress, 1st Session, Vol.127, Part 15	Link
congressional record	June 10, 1981, 97th Congress, 1st Session, Vol.127, Part 9	Link
congressional record	January 5, 1981, 97th Congress, 1st Session, Vol. 127, Part 1	Link
congressional record	April 30, 1981, 97th Congress, 1st Session, Vol.127, Part 6	Link
congressional record	July 22, 1981, 97th Congress, 1st Session, Vol.127, Part 13	Link
congressional record	November 12, 1981, 97th Congress, 1st Session, Vol.127, Part 21	Link
newspaper	The Christian Science Monitor	Link
newspaper	The Associated Press	Link
newspaper	The Associated Press	Link
newspaper	The New York Times	Link
newspaper	The New York Times	Link
newspaper	The Associated Press	Link
official	Economic Report of the President, 1982	Link
President statement	Ronald Reagan, The President's News Conference	Link
newspaper	The Daily Collegian	Link
newspaper	The Daily Collegian	Link
newspaper	Desert Sun	Link
newspaper	Desert Sun	Link
newspaper	The Louisville Times	Link
newspaper	Rappahannock Record	Link
newspaper	San Bernardino Sun	Link
newspaper	Indianapolis Recorder	Link
newspaper	Desert Sun	Link
newspaper	Rappahannock Record	Link
newspaper	Smithfield Times	Link
newspaper	San Bernardino Sun	Link
official	The Treasury Department Report to the President, Volume 2: General Explanation of the Treasury Department Proposals	Link
thinktank	American Enterprise Institute, Feenberg	Link
newspaper	Santa Cruz Sentinel	Link

Table C.18. – Underlying Text Data for TCJA17 (2016–2019)

Type of Source	Source Name	URL
thinktank	Center on Budget and Policy Priorities	Link
thinktank	Cato Institute	Link
thinktank	Tax Foundation	Link
thinktank	Independent Institute	Link
thinktank	Manhattan Institute	Link
thinktank	Urban Institute	Link
thinktank	American Enterprise Institute	Link
thinktank	American Enterprise Institute	Link
thinktank	R Street Institute	Link
thinktank	Urban Institute	Link
thinktank	Demos	Link
thinktank	American Enterprise Institute	Link
thinktank	American Enterprise Institute	Link
thinktank	Brookings Institution	Link
thinktank	Brookings Institution	Link
thinktank	Cato Institute	Link
thinktank	Center on Budget and Policy Priorities	Link
thinktank	The Century Foundation	Link
thinktank	Manhattan Institute	Link
thinktank	R Street Institute	Link
thinktank	Tax Foundation	Link
thinktank	Urban Institute	Link
thinktank	American Enterprise Institute	Link
newspaper	USNEWS.com	Link
thinktank	Institute for Family Studies	Link
thinktank	American Enterprise Institute	Link
thinktank	American Enterprise Institute	Link
thinktank	American Enterprise Institute	Link
thinktank	The Heartland Institute	Link
thinktank	Hudson Institute	Link
thinktank	Hudson Institute	Link
thinktank	Hudson Institute	Link
thinktank	Urban Institute	Link
newspaper	Deseret Morning News	Link
newspaper	States News Services	Link
journal	National Law Review	Link
journal	The ANNALS of the American Academy of Political and Social Science	Link
official	Joint Economic Committee	Link
thinktank	Brookings Institution	Link
thinktank	The Heritage Foundation	Link
thinktank	Manhattan Institute	Link
thinktank	Niskanen Center	Link
thinktank	American Enterprise Institute	Link
congressional record	October 17, 2017, 115th Congress, 1st Session, Vol.163, No.167	Link
congressional record	November 29, 2017, 115th Congress, 1st Session, Vol.163, No.194	Link
congressional record	October 11, 2017, 115th Congress, 1st Session, Vol.163, No.163	Link
congressional record	November 15, 2017, 115th Congress, 1st Session, Vol.163, No.187	Link
congressional record	December 1, 115th Congress, 1st Session, Vol.163, No.196	Link
congressional record	November 16, 2017, 115th Congress, 1st Session, Vol.163, No.188	Link
newspaper	Tribune Review	Link
newspaper	Tampa Bay Times	Link
newspaper	States News Services	Link
newspaper	Western Free Press	Link
newspaper	Western Free Press	Link

President statement	Donald J. Trump, The President's Weekly Address	Link
President statement	Donald J. Trump, Remarks at the Family Research Council's Values Voter Summit	Link
official	Economic Report of the President, 2018	Link
official	General Explanation of Public Law 115-97, Joint Committee on Taxation	Link
thinktank	Tax Foundation	Link
newspaper	The Oregon Catalyst	Link
thinktank	Georgia Center for Opportunity	Link
newspaper	The American Prospect	Link
blog	Newstex Blogs	Link
thinktank	Baker Institute	Link

Appendix to Chapter 4

D.1 SUPPLEMENTARY MATERIAL

D.1.1 Demographics

Table D.1. – Selected Demographics, SOEP, 2017

Variable	Share (%)
Sex (N=28565)	
Male	49.13
Female	50.87
Age (N=28565)	
18-29 years	14.13
30-44 years	23.22
45-59 years	28.74
60 years and older	33.90
Highest Education (N=28565)	
Lower Education	34.11
Intermediate Education	53.41
Higher Education	12.48
Marital status (N=28565)	
Single	49.63
Married	50.37

Note: This table presents summary statistics for demographics for adult indiviudals in Germany based on information from the German Socioeconomic Panel for the year 2017. We use a version of GSOEP (v35) that has been prepared for the utilization of the ifo microsimulation model.
Source: Own calculation based on GSOEP (v35).

Table D.2. – Selected Demographics, forsa Survey, 2021

Variable	Share (%)
Sex (N=8080)	
Male	48.93
Female	51.07
Age (N=8080)	
18-29 years	10.59
30-44 years	28.30
45-59 years	26.94
60 years and older	34.17
Highest Educational Degree (N=8080)	
Primary / Lower Secondary Education	16.08
Intermediate Secondary Education	41.08
Upper Secondary Education	18.96
Tertiary Education	21.66
Other / No response	2.23
Marital status (N=8080)	
Married, living together	52.72
Married, living separated	2.57
Widowed	6.13
Divorced	9.26
Single	28.32
Don't know	1.00

Note: This table presents summary statistics for demographics for adult indiviudals in Germany based on information from the *forsa* survey module in 2021. We use survey weights provided by *forsa*.
Source: Own calculation based on *forsa* survey module.

D.1.2 Material Self-Interest

Extensive Margin Responses. Equation (4.1) in Section 4.3 describes the sufficient statistics used to estimate revenue functions under intensive margin responses as derived in Bierbrauer et al. (2025). The formula $\mathcal{R}_1^{\text{int}}(y_1)$ can be extended to include also extensive margin responses.

$$\mathcal{R}_1^{\text{int}+ext}(y_1) = \mathcal{R}_1^{\text{int}}(y_1) + \mathcal{X}_{dec}(y_1) + \mathcal{X}_{sec}(y_1) \tag{D.1}$$

$$\mathcal{X}_{dec}(y_1) = -\int_{y_1}^{\bar{y}} \mathbf{E}\left[\frac{T_{m0}(y_m^0)}{y_m^0 - T_{m0}(y_m^0)} \times \pi_{dec}\left(y_m^0\right) \mid y_1^0 = y_1'\right] m_{dec}^{y_1}(y_1')\, dy_1', \tag{D.2}$$

$$\mathcal{X}_{sec}(y_1) = -\int_{y_1}^{\bar{y}} \mathbf{E}\left[\frac{T_{m0}(y_m^0)}{y_m^0 - T_{m0}(y_m^0)} \times \pi_{sec}\left(y_m^0\right) \mid y_1^0 = y_1'\right] m_{sec}^{y1}(y_1')\, dy_1'. \tag{D.3}$$

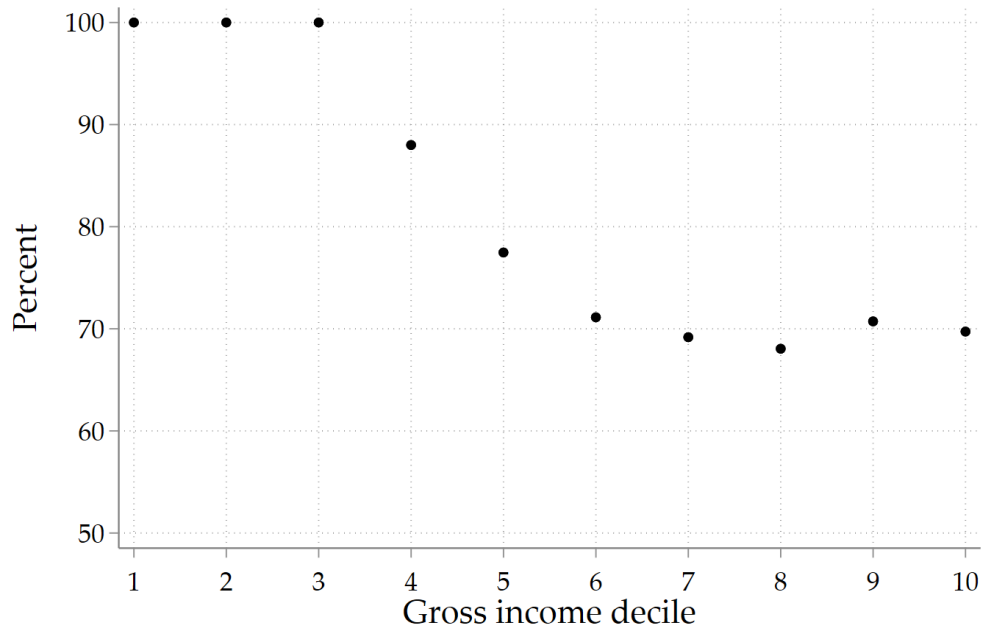
The extensive margin response consists of two parts. Equation (D.2) describes the revenue effects based on those dual earner couples that become single earner couples as a response to the tax change. Equation (D.3) describes the revenue effects based on those single earner couples that stop working completely.

Table D.3. – Welfare Weights

Welfare Weights	Specification
Equal (Feminist)	$\forall y_m,\ g_m(y_m) = 1$
Rawlsian	$\forall y_m,\ g_m(y_m) = \begin{cases} 1, & \text{for } y_m \leq \text{P} \\ 0, & \text{for } y_m \geq \text{P} \end{cases}$
Feminist	$\forall y_m,\ g_m(y_m) = \frac{y_{woman}}{y_{man}+y_{woman}}$

Note: This table shows different specifications of welfare weights to evaluate reforms towards individual taxation. The sum of weights over the whole population of married couples is normalized to 1. P refers to specific percentiles of the couple income distribution.

Figure D.1. – Median Income Share of Primary Earner by Deciles



Note: This figure displays the median income share of the primary earner across per capita gross income deciles. The estimates are based on all married couples with strictly positive gross income and in which both spouses are between 25 and 55 years old. Gross income consists of wage and self-employment income.
Source: Own calculation based on GSOEP v35.

D.1.3 Stated Preferences

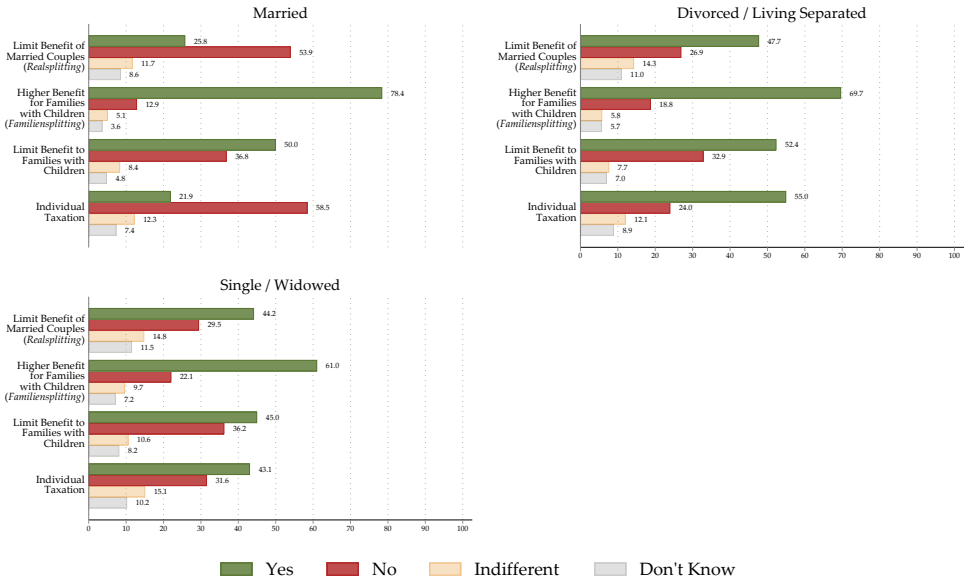
Table D.4. – Statements from Party Proposals

Party	Statements (German)	Statements (English)	Source
CDU/ CSU	Wir halten am Ehegattensplitting fest und wollen unabhängig davon zusätzlich Ansätze entwickeln, um Kinder positiv zu berücksichtigen. Wir haben die finanzielle Situation von Familien spürbar verbessert, indem wir den Kinderfreibetrag und das Kindergeld zum 1. Januar 2021 deutlich erhöht haben. Perspektivisch streben wir den vollen Grundfreibetrag für Kinder an und finden damit den Einstieg in ein Kindersplitting.	We will stick with income splitting for married couples and, independently of that, we want to develop additional approaches to positively consider children. We have significantly improved the financial situation of families by increasing the child tax exemption and child benefits on January 1, 2021. In the long term, we aim to achieve the full basic tax exemption for children, thus paving the way for child income splitting.	p. 72 / Translation
SPD	Das geltende Steuerrecht befördert die klassische Arbeitsteilung zwischen Männern und Frauen. Das werden wir ändern und das Steuerrecht stärker auf Partnerschaftlichkeit ausrichten sowie die Steuerlast bei unterschiedlich hohen Einkommen zwischen den Eheleuten gerecht verteilen. Das Ehegattensplitting bildet die gesellschaftliche Realität nicht mehr ab und schließt viele Haushalte mit Kindern von dem gewährten Steuervorteil aus. Nutznießer sind stattdessen vor allem Alleinverdiener-Ehepaare mit hohen Einkommen unabhängig von der Kinderzahl. Wir werden das Ehegattensplitting für neu geschlossene Ehen ändern. Die allermeisten Haushalte mit Kindern werden durch die Kindergrundsicherung finanziell bessergestellt werden. Normalverdienende auch ohne Kinder werden keine Einbußen erleiden. Für bestehende Ehen werden wir zudem ein Wahlrecht einführen.	Current tax law encourages the traditional division of labor between men and women. We will change this situation and gear tax legislation more strongly towards a partnership model. Tax splitting between married couples no longer reflects our social reality and excludes many households with children from tax benefits, whilst the primary beneficiaries are single-earner couples with high incomes, regardless of the number of children they have. We will change the marital splitting system for newly-married couples to ensure that the vast majority of households with children benefit financially from a basic child allowance, and that earners of a normal income will suffer no losses even if they do not have children. We will also introduce a right for existing married couples to choose between the various models.	p. 23 / p. 17
FDP	Am Splittingverfahren für Ehe- und eingetragene Lebenspartnerschaften wollen wir festhalten. Ebenso kann es sinnvoll sein, künftig stärker mit - von der Steuerschuld abzuziehenden - Steuergutschriften zu arbeiten. Dadurch wirken Freibeträge besser für die niedrigen und mittleren Einkommen.	We want to maintain the splitting procedure for married couples and registered life partnerships. It may also be useful to work more with tax credits that can be deducted from the tax debt. This makes exemptions more effective for low and middle incomes.	p. 33 / Translation
DIE LINKE	Das Ehegattensplitting wird mit sozialverträglichen Übergangsregelungen durch eine geschlechtergerechte Individualbesteuerung ersetzt. Dabei muss das nicht ausgeschöpfte steuerliche Existenzminimum zwischen Eheleuten bzw. Lebenspartner*innen übertragbar sein.	The joint taxation of married couples will be replaced with a gender-equitable individual taxation system, accompanied by socially compatible transitional regulations. In this context, the unused tax-free minimum between spouses or life partners must be transferable.	p. 88 / Translation

Table D.5. – Statements from Party Proposals (cont.)

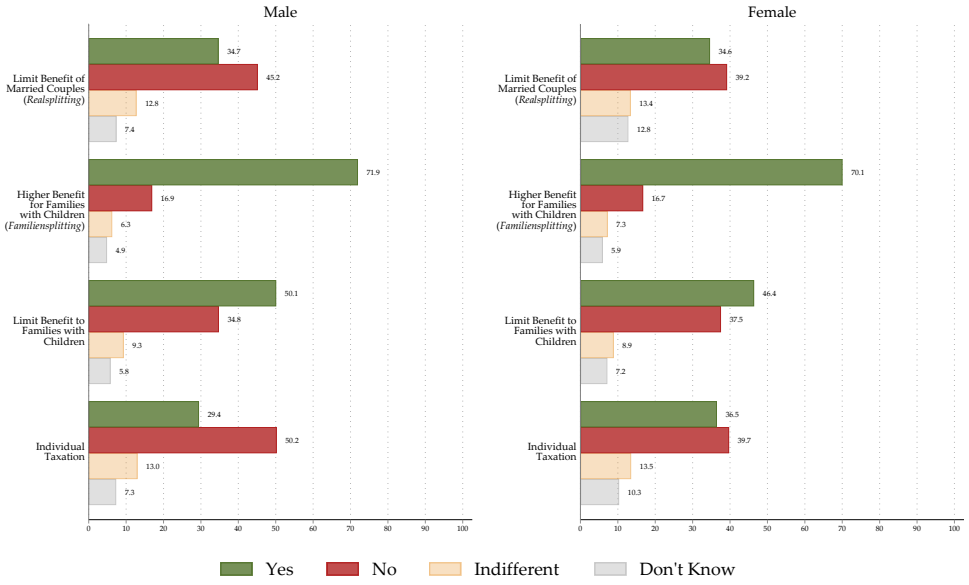
Party	Statements (German)	Statements (English)	Source
Bündnis 90/Die Grünen	Das deutsche Steuerrecht steckt noch im letzten Jahrhundert fest. Während sich viele Paare Familien- und Erwerbsarbeit gleichberechtigter aufteilen, als es noch vor Jahren der Fall war, gilt bei der Steuer nach wie vor das Modell eines männlichen Ernährers und einer Frau, die höchstens dazuverdient und sich hauptsächlich um Haushalt und Kinder kümmert. Dieses Modell ist ungerecht, weil es Ehen privilegiert, Ein-Eltern-Familien (Alleinerziehende) und nicht verheiratete Paare außen vor lässt, die Erwerbstätigkeit von Frauen hemmt und Frauen gleichzeitig nicht wirklich absichert. In Krisen bekommen vor allem Frauen die Nachteile zu spüren, zum Beispiel durch weniger Kurzarbeiter- oder Arbeitslosengeld. Im Zusammenspiel mit Minijobs und der kostenlosen Mitversicherung wirken sich diese Maßnahmen negativ auf die Erwerbstätigkeit von Frauen aus. Deshalb wollen wir für neu geschlossene Ehen eine individuelle Besteuerung mit übertragbarem Grundfreibetrag einführen. Bei der Lohnsteuer soll die/der heute über Gebühr belastete Zweitverdienerin entlastet werden, indem das Faktorverfahren zur Regel und die Steuerklasse 5 für Zuverdienerinnen abgeschafft wird. So sorgen wir dafür, dass gleichberechtigte Lebensentwürfe nicht länger benachteiligt werden. Paare, die bereits verheiratet sind, können sich entscheiden, ob sie sich einzeln veranlagen oder weiterhin das Ehegattensplitting nutzen wollen. Zugleich stärken wir mit der Kindergrundsicherung Familien. Alleinerziehende, die heute am stärksten von Armut betroffen sind, entlasten wir mit einer Steuergutschrift.	The German tax laws are still stuck in the last century. While many couples divide up family work and gainful employment more equally than was the case many years ago, taxation is still based on the model of a male breadwinner and a wife who at most earns a small amount extra and is primarily takes care of the household and children. This model is unfair because it favours married couples, leaves out single-parent families and unmarried couples, limits the participation of women in the labor market and at the same time does not truly give women full security. In crises, it is women who are most heavily disadvantaged, for example, through lower short-time working allowances or unemployment benefits. Together with minijobs and free co-insurance, these measures have a negative effect on the participation of women in the labor market. We therefore want to introduce individual taxation with a transferable tax-free allowance for newly married couples. In the case of income tax, the second earner, who currently bears an undue tax burden, will have this burden relieved by making the so-called 'work-factor method' the rule and scrapping tax class 5 for second earners. In this way, we will ensure that equal life plans are no longer disadvantaged. Couples who are already married can decide whether they want to be assessed individually or continue to split their taxation. At the same time, we will strengthen families with the guaranteed child allowance. Single parents, who are currently the hardest hit by poverty, will be supported with tax credits.	p. 111 / p. 107
AfD	Vor allem Familien der mittleren Einkommensschicht werden in Deutschland steuerlich stärker belastet als in fast allen anderen europäischen Ländern. Die Geburtenrate ist bei diesen Familien erschreckend gering. Die AfD will hier mit dem steuerlichen Familiensplitting einen Paradigmenwechsel einleiten, verbunden mit einer Anhebung des Kinderfreibetrages.	Especially families in the middle-income bracket are taxed more heavily in Germany than in almost all other European countries. The birthrate in these families is alarmingly low. The AfD aims to initiate a paradigm shift here with family income splitting, along with an increase in the child tax exemption.	p. 106 / Translation

Figure D.2. – Approval Rates, by Household Type



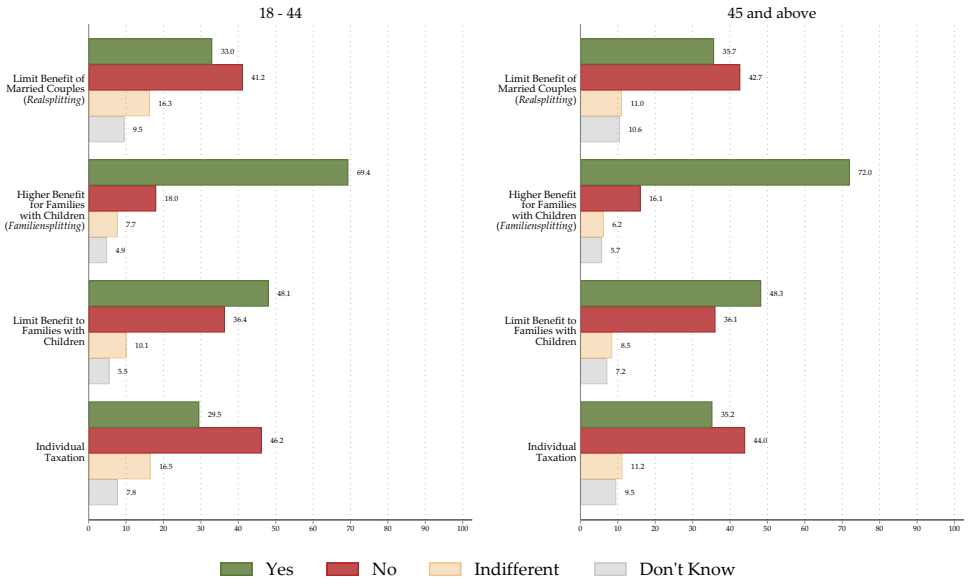
Note: This figure displays for different household types the approval rates concerning reform options for the system of joint taxation in Germany.
Source: Own calculation based

Figure D.3. – Approval Rates, by Sex



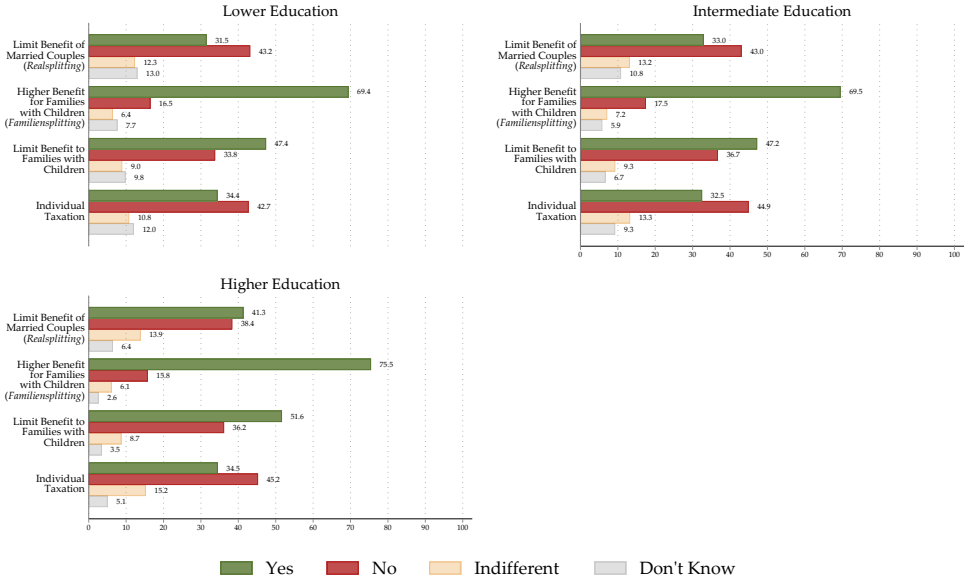
Note: This figure displays for men and women the approval rates concerning reform options for the system of joint taxation in Germany.
Source: Own calculations based on *forsa* survey module.

Figure D.4. – Approval Rates, by Age



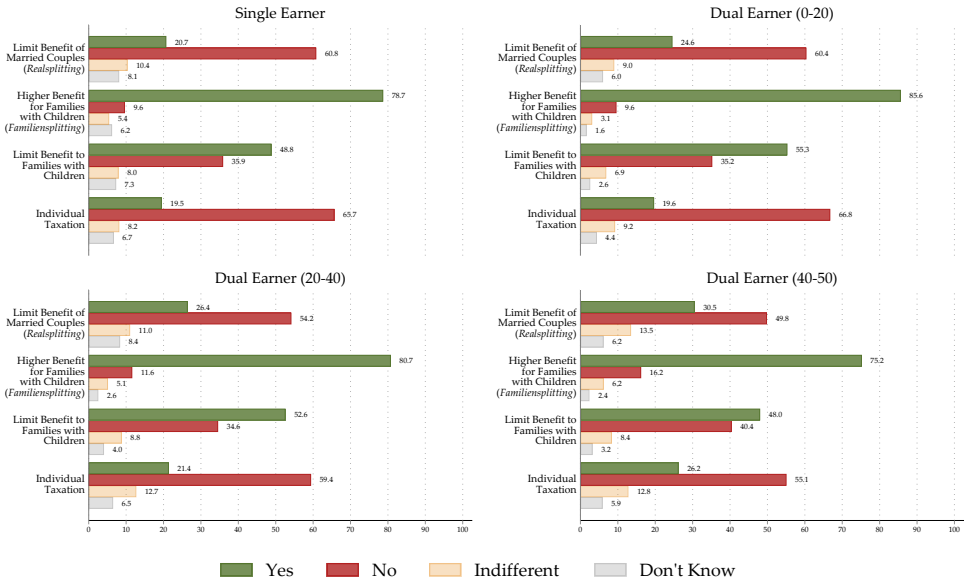
Note: This figure displays for young and old respondents the approval rates concerning reform options for the system of joint taxation in Germany.
Source: Own calculations based on *forsa* survey module.

Figure D.5. – Approval Rates, by Education



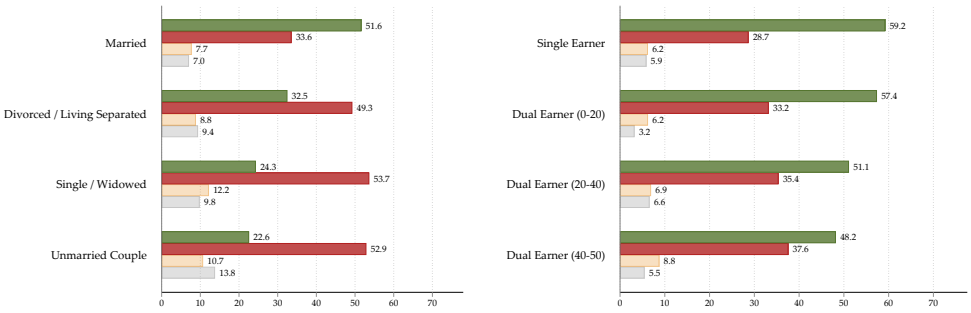
Note: This figure displays for different education groups the approval rates concerning reform options for the system of joint taxation in Germany.
Source: Own calculations based on *forsa* survey module.

Figure D.6. – Approval Rates, by Type of Married Couple



Note: This figure displays for all married couples the approval rates concerning reform options for the system of joint taxation in Germany for different types of married couples.
Source: Own calculations based on *forsa* survey module.

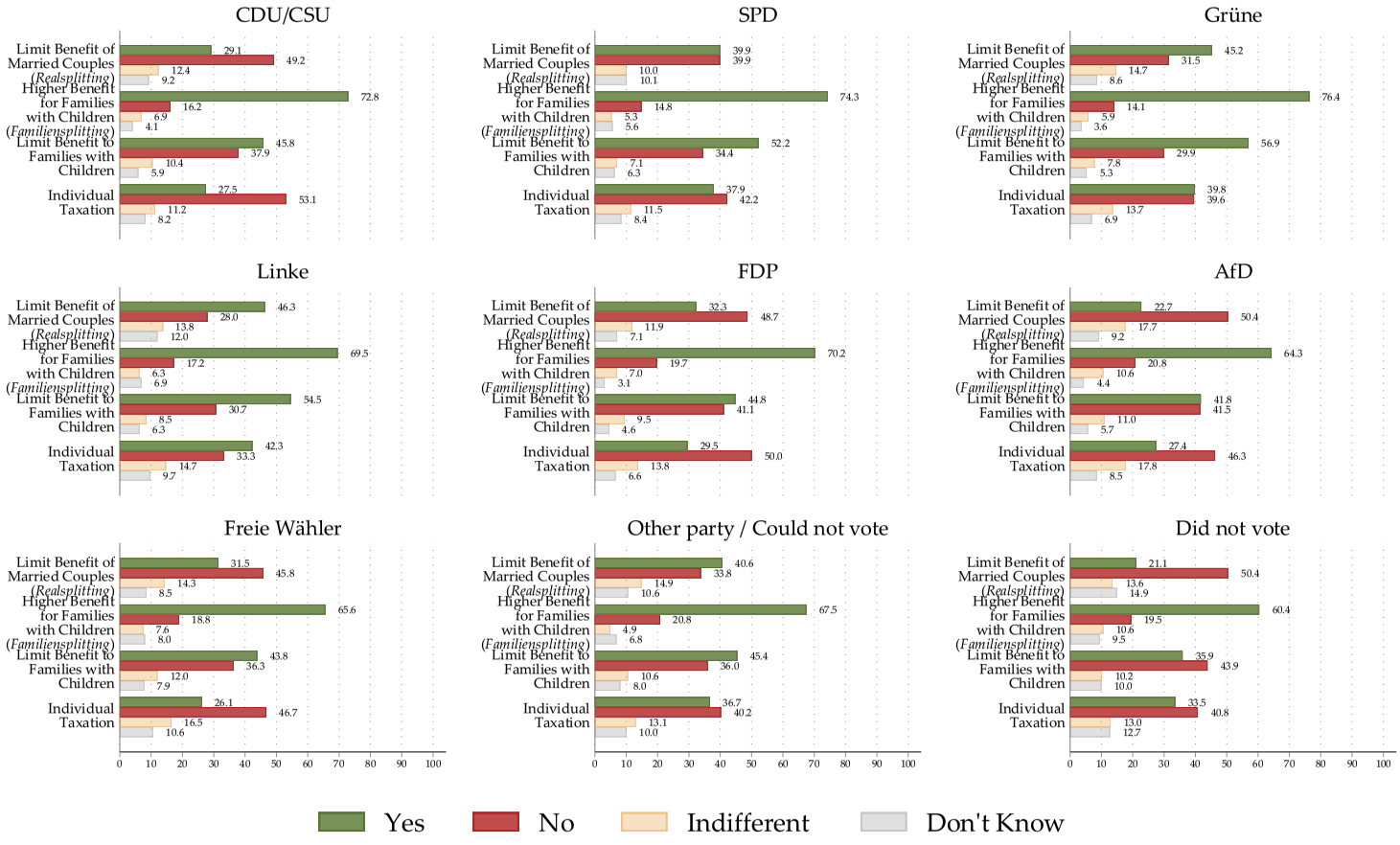
Figure D.7. – Limit Reform to Newly Married Couples, Approval Rates



(A) By marital status (B) By couple type

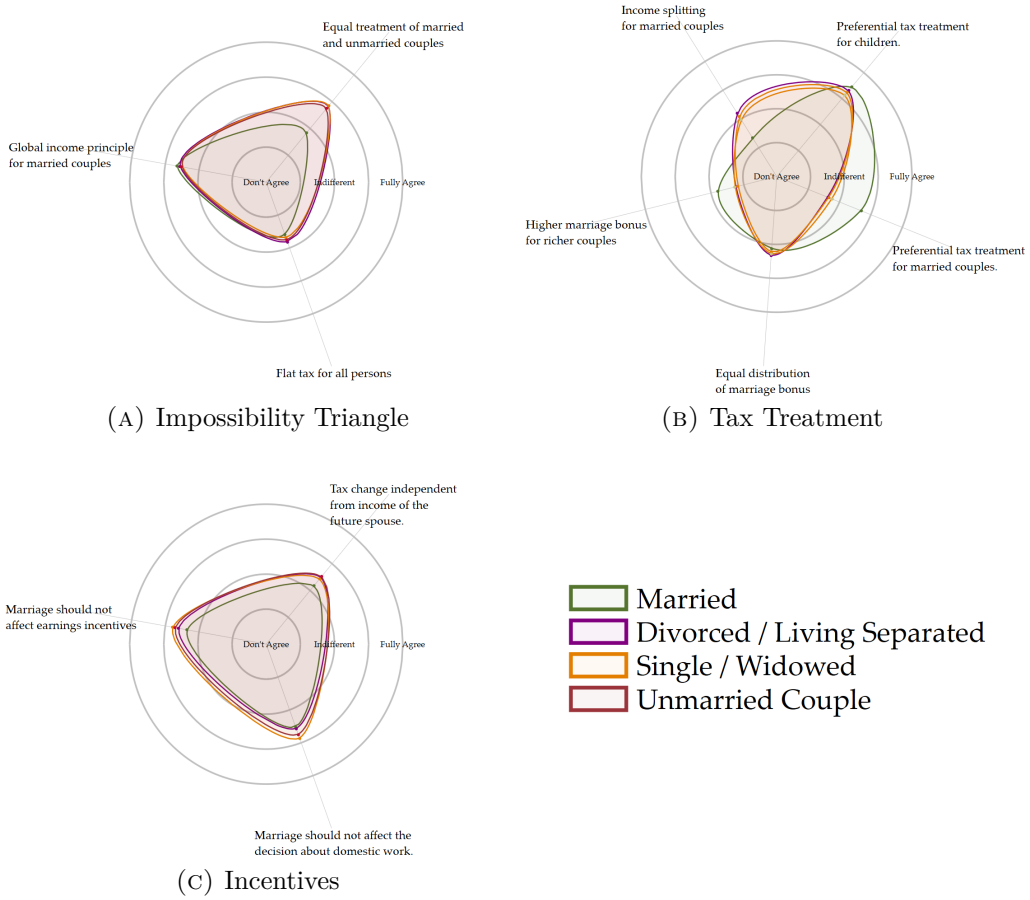
Note: This figure displays the approval rates for limiting a reform of joint taxation to newly married couples. Figure D.7a displays the results for the full population by marital status, while Figure D.7b displays results for married couples by the couple type answering the question. We exclude respondents that did not reveal its marital status and income. We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

Figure D.8. – Approval Rates, by Political Party Preference



Note: This figure displays for different political preference groups the approval rates concerning reform options for the system of joint taxation in Germany. Political party preference is measured by the stated party vote for the German federal election in 2021.
Source: Own calculations based on *forsa* survey module.

Figure D.9. – General Principles, by Household Type



(A) Impossibility Triangle

(B) Tax Treatment

(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation by household type. The first three principles cover the impossibility triangle principles (see Figure D.9a). Figure D.9b displays results for detailed principles of the tax treatment of couples while the last three principles capture principles specified in terms of different incentive margins (see Figure D.9c). Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question.
Source: Own calculations based on *forsa* survey module.

Figure D.10. – General Principles, by Earner Type

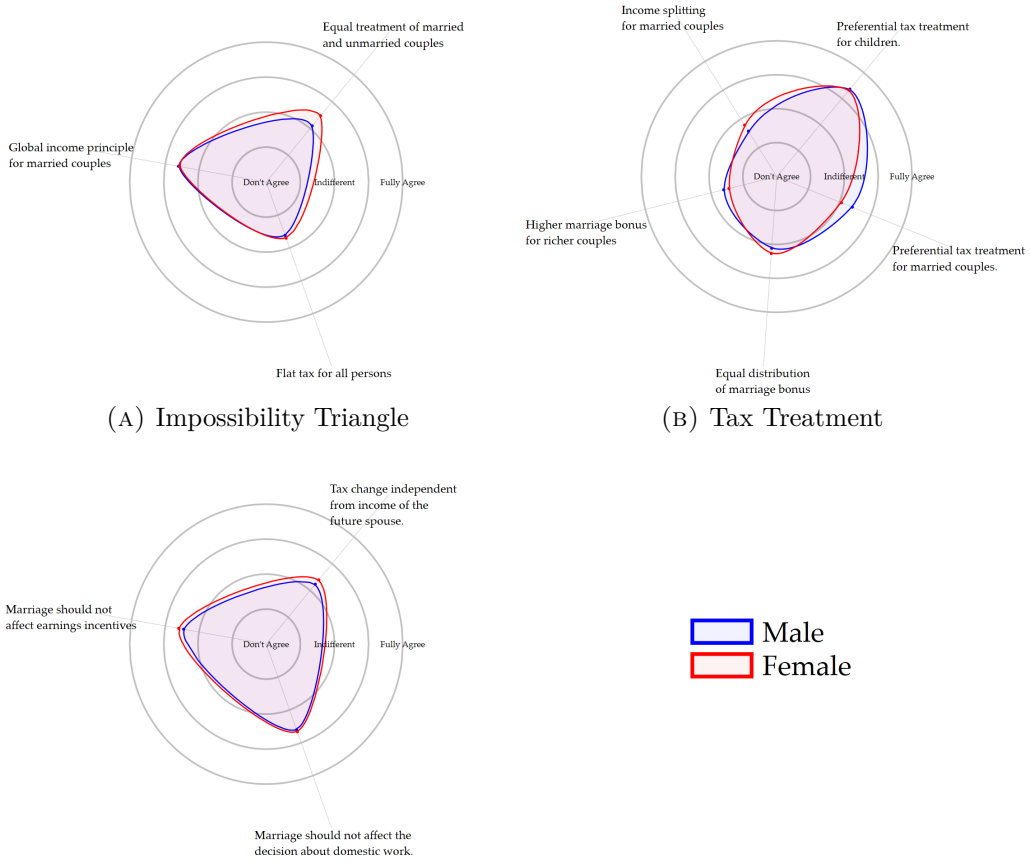
(A) Impossibility Triangle

(B) Tax Treatment

(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation for all married couples by earner type. The first three principles cover the impossibility triangle principles (see Figure D.10a). Figure D.10b displays results for detailed principles of the tax treatment of couples while the last three principles capture principles specified in terms of different incentive margins (see Figure D.10c). Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question.
Source: Own calculations based on *forsa* survey module.

Figure D.11. – General Principles, by Sex



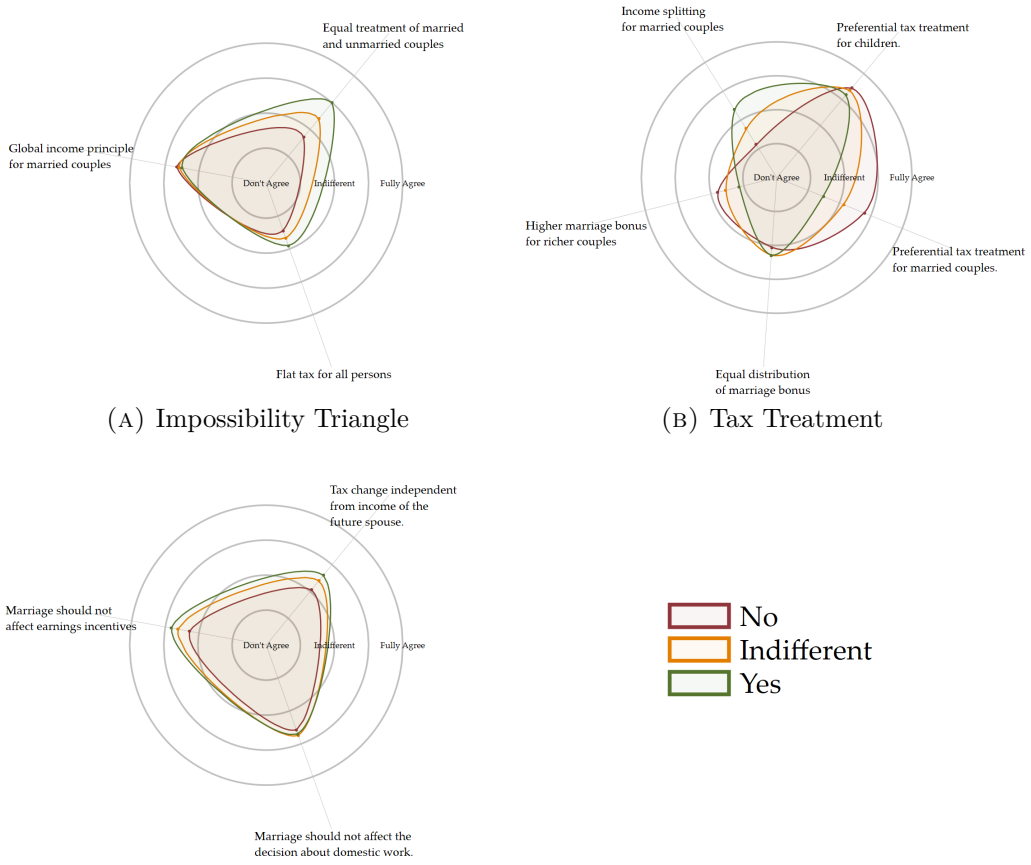
(A) Impossibility Triangle

(B) Tax Treatment

(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation by sex. The first three principles cover the impossibility triangle principles (see Figure D.11a). Figure D.11b displays results for detailed principles of the tax treatment of couples while the last three principles capture principles specified in terms of different incentive margins (see Figure D.11c). Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question.
Source: Own calculations based on *forsa* survey module.

Figure D.12. – General Principles, by Approval of Individual Taxation



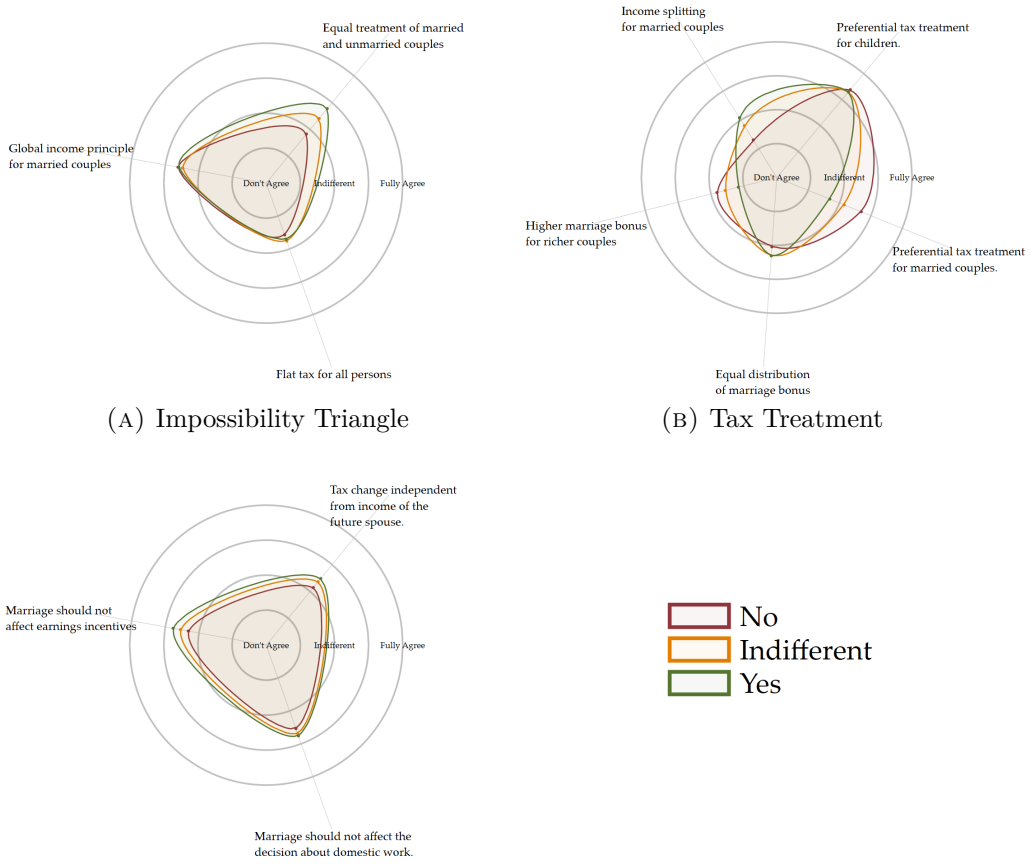
(A) Impossibility Triangle

(B) Tax Treatment

(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation by approval of individual taxation. Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question.
Source: Own calculations based on *forsa* survey module.

Figure D.13. – General Principles, by Approval of Realsplitting



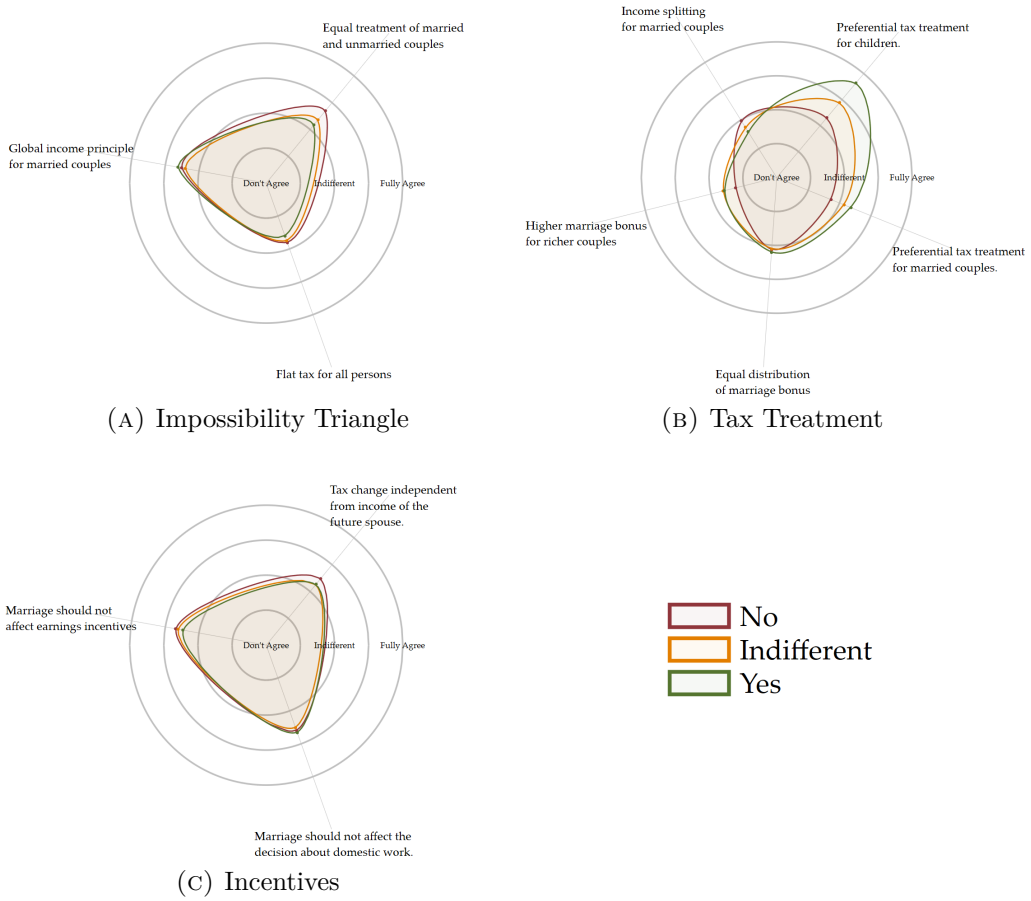
(A) Impossibility Triangle

(B) Tax Treatment

(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation by approval of Realsplitting. Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question.
Source: Own calculations based on *forsa* survey module.

Figure D.14. – General Principles, by Approval of Family Splitting



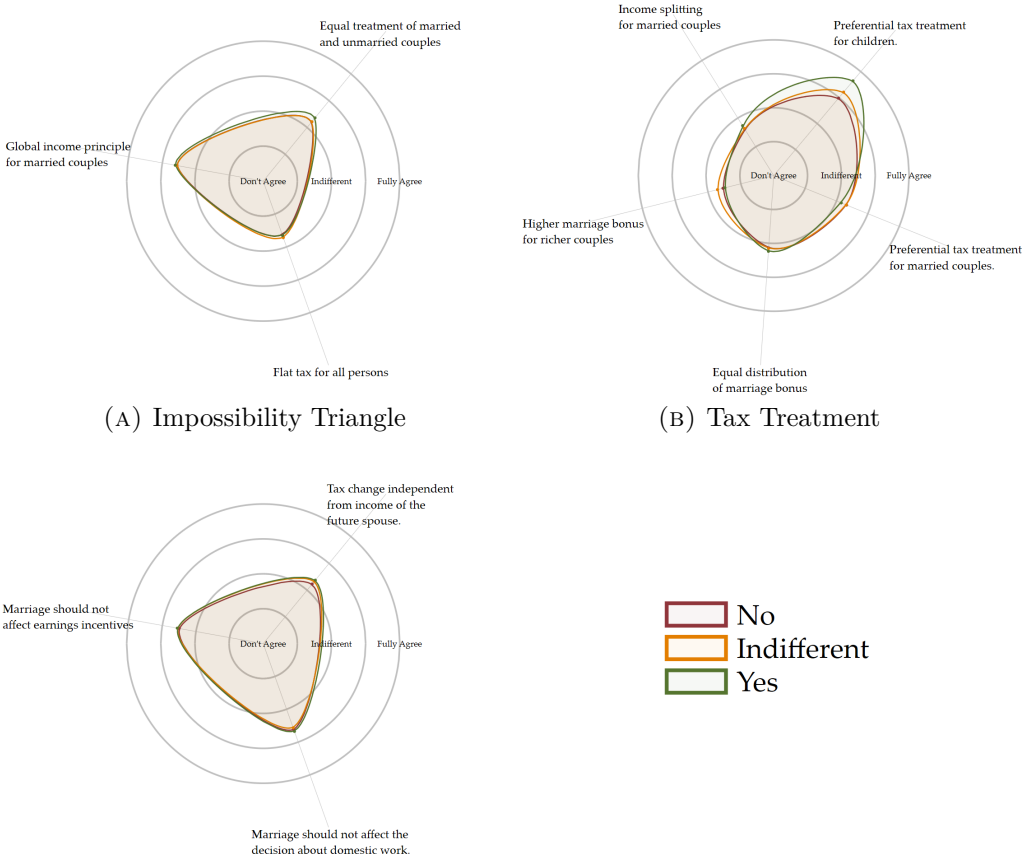
(A) Impossibility Triangle

(B) Tax Treatment

(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation by approval of family splitting. Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question.
Source: Own calculations based on *forsa* survey module.

Figure D.15. – General Principles, by Approval to Limit Marriage Benefit to Families with Children



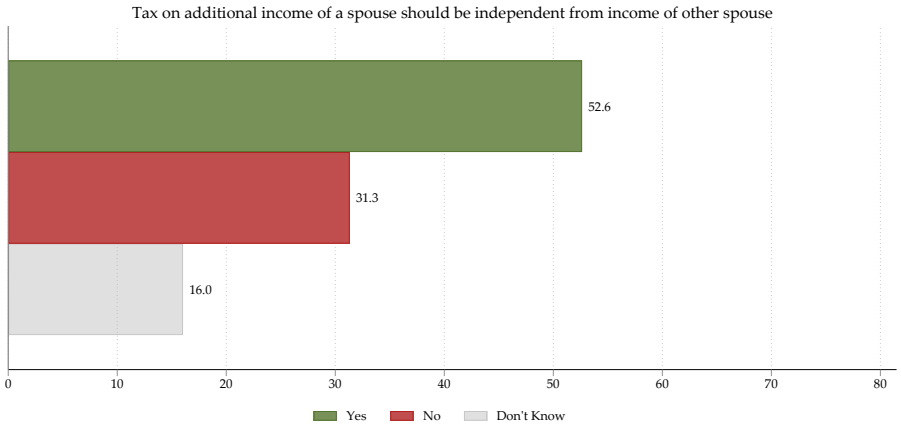
(A) Impossibility Triangle

(B) Tax Treatment

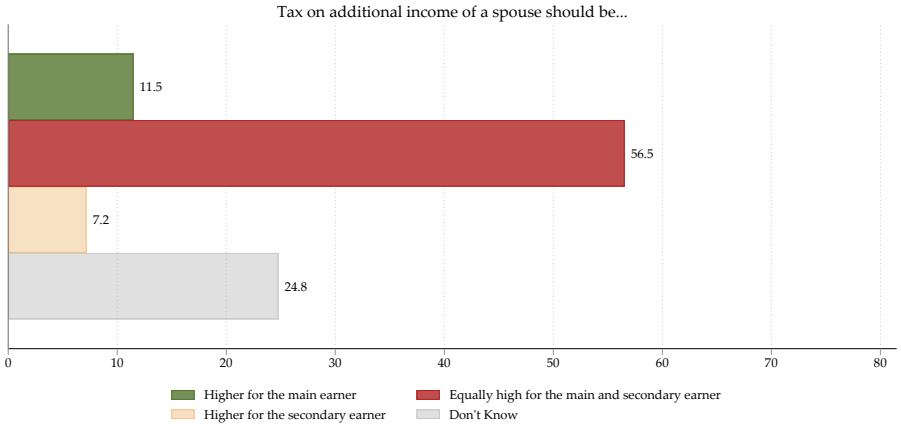
(C) Incentives

Note: This figure displays the average approval responses for different principles of couple taxation by approval to limit marriage benefit to families with children. Approval to each principle was elicited using a five point scale ranging from "Full Disagreement" (1) to "Full Agreement" (5). Details on the response scales and exact statements can be found in Appendix D.3. We exclude respondents that decided not to answer the question.
Source: Own calculations based on *forsa* survey module.

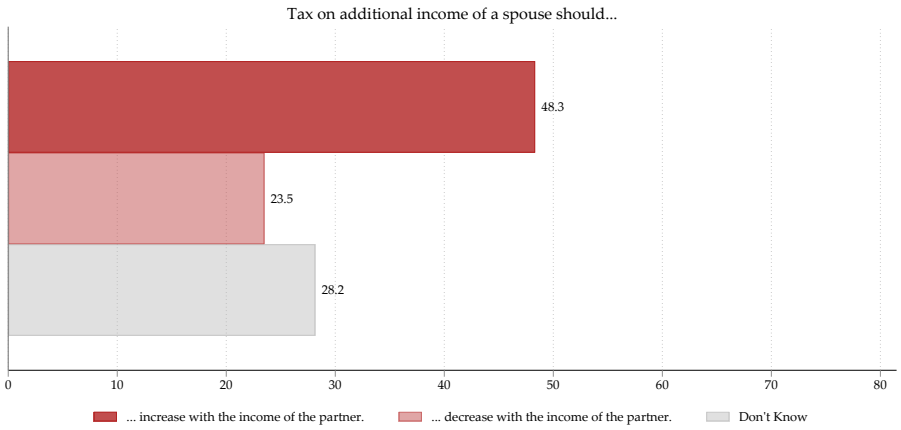
Figure D.16. – Tax on Additional Income



(A) Independency



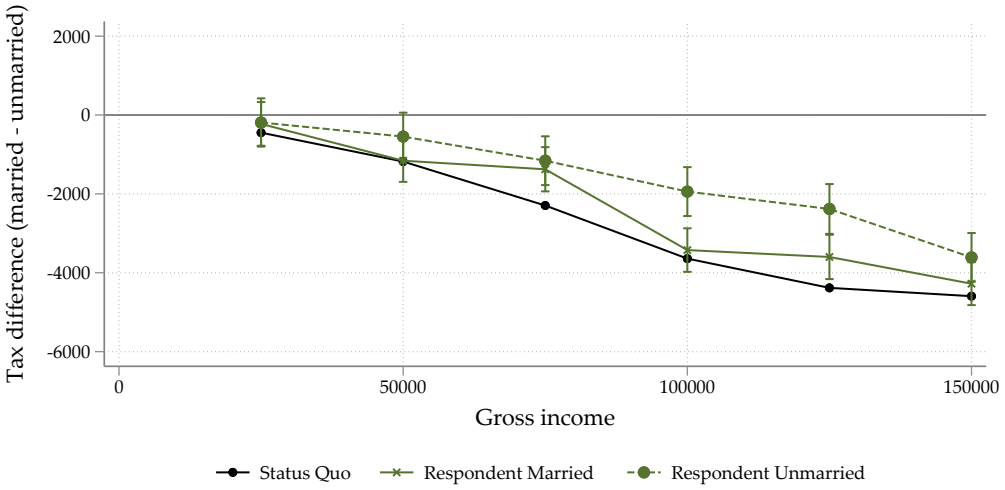
(B) Level Comparison



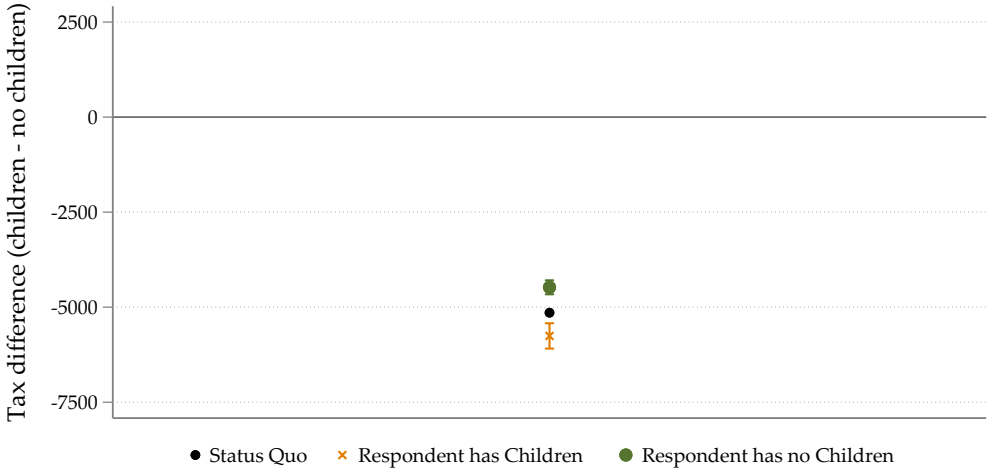
(C) Direction of Jointness

Note: This figure displays the average approval rates for questions about the (in)dependency of marginal tax rates between primary and secondary earner. Results for the level comparison of marginal tax rates between the primary and the secondary earner (D.16b) as well as the results for the preferred direction of jointness (see Figure D.16c) consist of answers from those respondents that think that the tax on additional income should not be independent (see Figure D.16a).
Source: Own calculations based on *forsa* survey module.

Figure D.17. – Vignette Evaluations, Tax Liability, Self-Interested Behavior



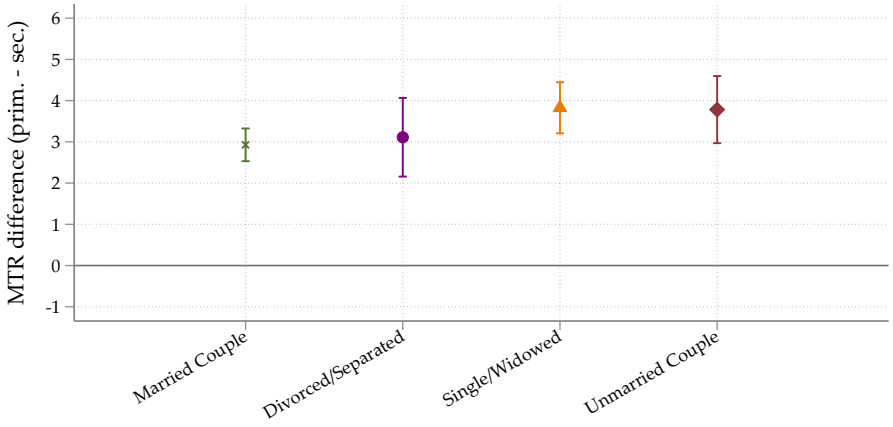
(A) Preferred Marriage Bonus



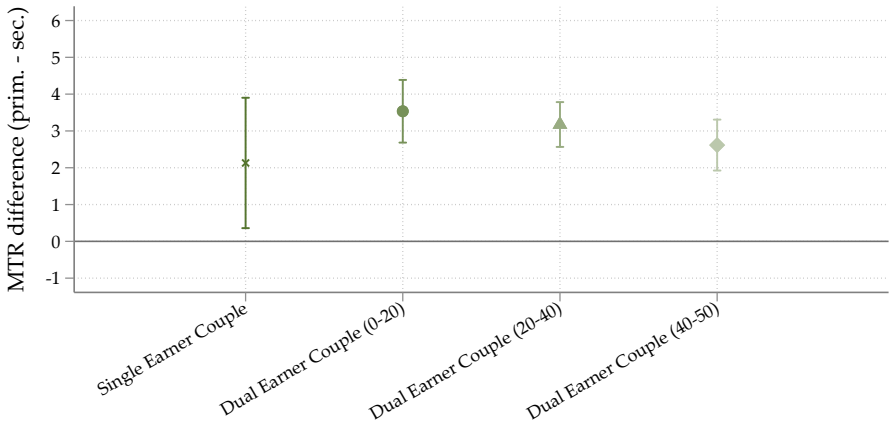
(B) Preferred Tax Treatment of Children Bonus

Note: This figure displays heterogeneity analysis visualizing the self-interested behavior of respondents in the vignette experiment. Figure D.17a displays the preferred marriage bonus across different vignette household income levels separately for married and unmarried respondents. Figure D.17b displays the preferred tax treatment of children in the vignette separately for respondents with and without children. Results are based on the predicted average of a fully saturated linear regression model using all household types and including all interactions of household type, gross income, the primary earner income share and the number of children of the vignette household type (see Table 4.1). We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

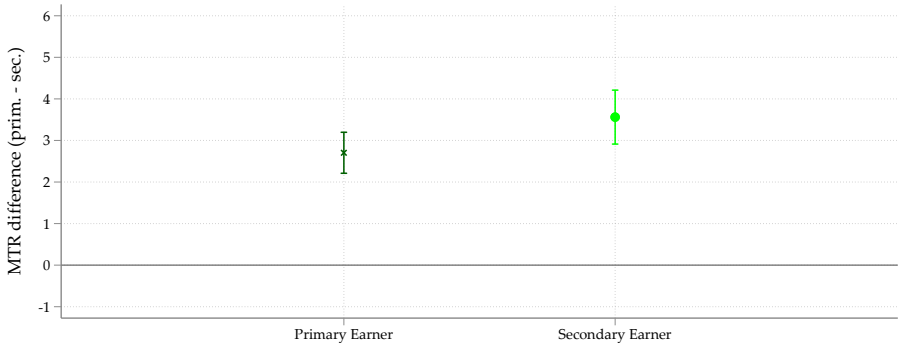
Figure D.18. – Vignette Evaluations, Marginal Tax Rates, Self-Interested Behavior



(A) Respondent Household Type



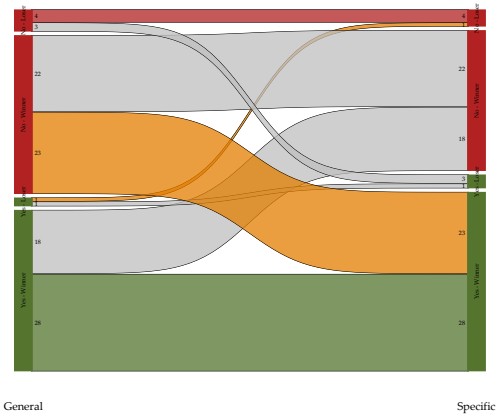
(B) Married Couples, Respondent Couple Type



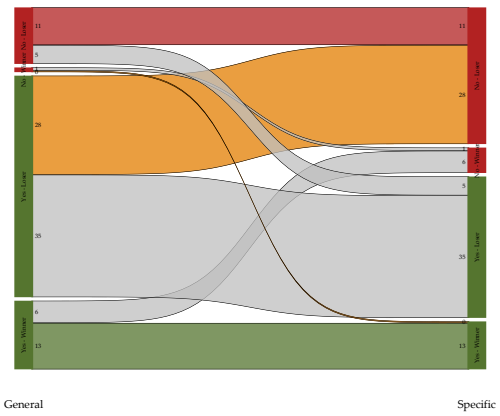
(C) Married Couples, Respondent Earner Type

Note: This figure displays heterogeneity analysis visualizing the self-interested behavior of respondents in the vignette experiment concerning the implemented difference between the primary and secondary earner marginal tax rate in the experiment. Figure D.18a displays the average difference across different respondent household types. Figure D.18b and Figure D.18c display the average marginal tax rate difference focusing on married respondents only, and separating these respondents by couple type and earner type respectively. We use survey weights provided by *forsa*.
Source: Own calculations based on *forsa* survey module.

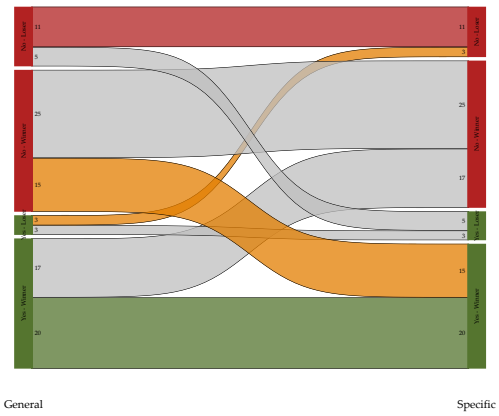
Figure D.19. – Support for Reforms - Changes through Information, Details



(A) Limit Benefit of Married Couples (*Realsplitting*)



(B) Higher Benefit for Families w/ Children (*Familiensplitting*)



(C) Individual Taxation

Note: This figure displays how the support for three reform options concerning the taxation of couples changes after the respondents are provided with detailed information on the status quo and the consequences of specific reforms. The information is provided through three different household pairs. We exclude respondents that answer "Don't know" or "Indifferent" in either the general elicitation or the specific elicitation, since the different response levels do not allow for a consistent comparison. Each respondent is included three times in the data, since specific support was elicited for each of the three household pairs shown to respondents separately. We use survey weights provided by *forsa*. Respondents are further separated by whether they are considered a reform winner or a reform loser. To determine winners and losers, we simulate the reform for each *forsa* household, taking into account the aggregate fiscal effects and redistributing associated revenue gains at the individual level. For details on the aggregate fiscal effects of the reforms based on the ifo microsimulation model, see Appendix D.1.4. Persistent self-interested behavior of reform winners (losers) is marked using green (red) pathways. Orange pathways illustrate those respondents that do not act in line with self-interest in the general question, but shift behavior in line with self-interest after the factorial survey.
Source: Own calculations based on *forsa* survey module.

D.1.4 Simulated Reform Effects, SOEP

This appendix provides information on the effect of each of the three main reforms discussed in the main text based on the ifo microsimulation model and data from the German Socio-Economic Panel (GSOEP). Table D.6 displays the aggregate fiscal effects of the reform and the necessary per-capital lump-sum adjustment. Tables D.7-D.9 illustrate the effect of the reform on disposable household income for different household types.

Table D.6. – Aggregate Fiscal Effects, Change in Billion Euro – Taxation Reform including Lump Sum Transfer

Budgetary Variable	Individual Taxation	Real Tax Splitting	Family Tax Splitting
Income Tax	21.6	4.5	-23.2
Withholding Tax	0.2	-0.3	-0.0
Solidarity Surcharge	1.4	0.2	-1.0
Lump Sum Transfers	-23.1	-4.4	24.1
Per Adult (Euro)	*-341*	*-65*	*355*
Total	0.0	0.0	0.0

Note: Aggregate fiscal effects of a simulated reform. Positive values mean a relief, negative values a burden on the fiscal budget. Own calculations based on GSOEP v39, data year 2018.

Table D.7. – Disposable Household Income – Individual Taxation and Lump Sum Transfer

	Status Quo (Euro)		Change (Euro)				Share Winners (%)	
	Mean	p50	Mean	p25	p50	p75	Households	Adults
Total	34,413	26,723	1	341	341	344	83.36	74.22
By Household Type								
Single	20,766	17,616	341	341	341	341	100.00	100.00
Single Parent	29,006	24,998	341	341	341	341	100.00	100.00
Couple, unmarr. w/o Children	43,741	39,648	681	681	681	681	100.00	100.00
Couple, unmarr., with Children	48,905	44,643	681	681	681	681	100.00	100.00
Couple, married, w/o Children	45,365	37,907	-418	-717	325	673	60.97	60.97
Couple, married, with Children	59,788	51,838	-984	-1,937	-80	562	48.03	47.38
By Number of Children								
No Children	29,858	23,389	136	341	341	341	88.21	82.84
One Child	48,659	43,188	-288	-317	341	630	71.08	68.27
Two Children	56,003	49,015	-735	-1,392	341	594	58.44	56.86
Three Children	58,199	47,447	-947	-1,851	178	633	54.62	53.46
Four or more Children	51,056	45,546	-836	-1,786	161	681	52.95	52.91
By Primary Earner Gross Income Share (Couples)								
[50,60]	44,957	41,444	634	630	681	681	99.58	99.51
(60,90]	51,734	45,271	-167	-591	190	679	57.79	55.40
(90,100]	50,841	36,869	-2,404	-3,995	-2,273	681	30.98	27.96
By Decile of Equivalized Status Quo Household Income								
Decile 1	8,862	8,986	389	341	341	341	98.82	97.15
Decile 2	15,567	13,072	330	341	341	341	94.85	89.46
Decile 3	19,792	16,213	223	341	341	341	89.80	79.75
Decile 4	24,359	19,294	148	341	341	413	86.65	76.59
Decile 5	28,481	29,200	77	341	341	362	83.46	74.17
Decile 6	32,957	33,728	46	337	341	403	81.43	72.58
Decile 7	38,068	38,973	29	341	341	540	81.58	73.97
Decile 8	45,198	45,385	-141	75	341	502	76.16	67.09
Decile 9	54,904	54,928	-382	-425	341	450	69.09	59.25
Decile 10	95,511	78,312	-1,021	-1,567	341	372	61.10	52.13
By Decile of Equivalized Status Quo Household Income (Married Couples)								
Decile 1	17,385	16,691	455	681	681	681	87.53	84.86
Decile 2	25,916	23,468	-78	-526	588	681	66.32	58.54
Decile 3	31,213	27,760	-248	-672	408	680	63.92	57.54
Decile 4	35,785	31,961	-355	-820	244	624	59.89	56.45
Decile 5	40,805	36,300	-416	-823	152	594	56.52	54.88
Decile 6	45,253	40,811	-432	-721	247	570	57.58	56.55
Decile 7	51,520	46,589	-575	-846	160	541	56.03	54.68
Decile 8	59,460	53,928	-919	-1,408	-120	470	45.82	43.32
Decile 9	69,732	65,987	-1,308	-2,085	-508	391	37.58	36.13
Decile 10	118,835	94,639	-2,110	-3,459	-1,050	426	36.24	35.97

Note: The table shows summary statistics for the disposable household income per year. Column changes shows statistics for the differences in household's income compared to the status quo. Own calculations based on GSOEP v39, data year 2018.

Table D.8. – Disposable Household Income – Real Tax Splitting and Lump Sum Transfer

	Status Quo (Euro)		Change (Euro)				Share Winners (%)	
	Mean	p50	Mean	p25	p50	p75	Households	Adults
Total	34,413	26,723	-1	65	65	130	91.85	86.60
By Household Type								
Single	20,766	17,616	65	65	65	65	100.00	100.00
Single Parent	29,006	24,998	65	65	65	65	100.00	100.00
Couple, unmarr. w/o Children	43,741	39,648	130	130	130	130	100.00	100.00
Couple, unmarr., with Children	48,905	44,643	130	130	130	130	100.00	100.00
Couple, married, w/o Children	45,365	37,907	-64	104	130	131	84.03	84.03
Couple, married, with Children	59,788	51,838	-223	-90	130	131	69.10	68.68
By Number of Children								
No Children	29,858	23,389	31	65	65	130	95.18	92.98
One Child	48,659	43,188	-53	65	98	130	83.10	81.46
Two Children	56,003	49,015	-185	-1	114	130	74.96	74.01
Three Children	58,199	47,447	-276	-71	101	130	71.49	70.76
Four or more Children	51,056	45,546	-156	8	128	130	75.41	75.25
By Primary Earner Gross Income Share (Couples)								
[50,60]	44,957	41,444	138	130	130	131	98.79	98.72
(60,90]	51,734	45,271	16	101	130	131	83.84	82.18
(90,100]	50,841	36,869	-619	-585	43	130	55.81	51.13
By Decile of Equivalized Status Quo Household Income								
Decile 1	8,862	8,986	76	65	65	65	99.92	99.82
Decile 2	15,567	13,072	80	65	65	65	99.12	97.73
Decile 3	19,792	16,213	77	65	65	130	97.64	94.32
Decile 4	24,359	19,294	80	65	65	130	96.15	91.89
Decile 5	28,481	29,200	81	65	65	130	96.12	91.99
Decile 6	32,957	33,728	75	65	65	130	94.77	90.42
Decile 7	38,068	38,973	68	65	65	130	93.48	89.19
Decile 8	45,198	45,385	3	65	65	130	88.98	82.67
Decile 9	54,904	54,928	-106	65	65	130	81.33	73.62
Decile 10	95,511	78,312	-533	-481	65	130	63.75	54.33
By Decile of Equivalized Status Quo Household Income (Married Couples)								
Decile 1	17,385	16,691	124	130	130	130	98.67	98.08
Decile 2	25,916	23,468	95	92	130	130	92.38	88.38
Decile 3	31,213	27,760	92	105	130	130	89.37	85.15
Decile 4	35,785	31,961	95	120	130	140	90.72	86.70
Decile 5	40,805	36,300	79	120	130	139	87.90	84.48
Decile 6	45,253	40,811	68	118	131	138	86.45	83.46
Decile 7	51,520	46,589	-18	88	130	131	81.85	78.21
Decile 8	59,460	53,928	-169	-96	130	131	69.75	65.44
Decile 9	69,732	65,987	-403	-464	91	130	58.54	54.96
Decile 10	118,835	94,639	-1,056	-1,676	-227	130	36.92	35.52

Note: The table shows summary statistics for the disposable household income per year. Column changes shows statistics for the differences in household's income compared to the status quo. Own calculations based on GSOEP v39, data year 2018.

Table D.9. – Disposable Household Income – Family Tax Splitting and Lump Sum Transfer

	Status Quo (Euro)		Change (Euro)				Share Winners (%)	
	Mean	p50	Mean	p25	p50	p75	Households	Adults
Total	34,413	26,723	-0	-711	-355	-355	16.13	31.42
By Household Type								
Single	20,766	17,616	-355	-355	-355	-355	0.00	0.00
Single Parent	29,006	24,998	1,465	-355	363	2,771	52.89	51.63
Couple, unmarr. w/o Children	43,741	39,648	-711	-711	-711	-711	0.00	0.00
Couple, unmarr., with Children	48,905	44,643	1,234	302	1,062	1,865	84.02	83.35
Couple, married, w/o Children	45,365	37,907	-711	-711	-711	-711	0.00	0.00
Couple, married, with Children	59,788	51,838	2,165	798	1,740	3,275	86.80	85.63
By Number of Children								
No Children	29,858	23,389	-488	-711	-355	-355	0.00	0.00
One Child	48,659	43,188	1,185	267	963	1,758	79.51	81.95
Two Children	56,003	49,015	2,539	1,148	2,479	3,490	82.71	83.86
Three Children	58,199	47,447	3,276	18	3,174	5,151	75.20	75.90
Four or more Children	51,056	45,546	2,481	-711	1,394	4,848	57.91	57.16
By Primary Earner Gross Income Share (Couples)								
[50,60]	44,957	41,444	60	-711	-711	438	25.70	35.61
(60,90]	51,734	45,271	337	-711	-711	986	32.14	46.02
(90,100]	50,841	36,869	299	-711	-711	735	31.93	44.08
By Decile of Equivalized Status Quo Household Income								
Decile 1	8,862	8,986	-422	-355	-355	-355	0.19	0.61
Decile 2	15,567	13,072	-403	-355	-355	-355	3.20	7.89
Decile 3	19,792	16,213	-270	-641	-355	-355	12.13	26.05
Decile 4	24,359	19,294	-88	-711	-355	-355	18.99	37.60
Decile 5	28,481	29,200	10	-711	-355	-355	20.90	39.96
Decile 6	32,957	33,728	103	-711	-355	-355	23.39	42.90
Decile 7	38,068	38,973	133	-711	-355	-355	23.40	41.46
Decile 8	45,198	45,385	274	-711	-355	565	25.14	43.36
Decile 9	54,904	54,928	329	-711	-355	-355	23.03	39.74
Decile 10	95,511	78,312	648	-711	-711	-355	20.10	34.62
By Decile of Equivalized Status Quo Household Income (Married Couples)								
Decile 1	17,385	16,691	-674	-711	-711	-711	2.10	3.76
Decile 2	25,916	23,468	-309	-711	-711	-519	21.68	33.40
Decile 3	31,213	27,760	63	-711	-711	546	33.28	48.89
Decile 4	35,785	31,961	219	-711	-711	812	37.09	52.73
Decile 5	40,805	36,300	385	-711	-711	990	41.40	56.44
Decile 6	45,253	40,811	388	-711	-711	1,091	39.50	53.84
Decile 7	51,520	46,589	562	-711	-711	1,369	40.18	54.74
Decile 8	59,460	53,928	734	-711	-711	1,638	39.35	54.27
Decile 9	69,732	65,987	751	-711	-711	1,953	33.31	46.72
Decile 10	118,835	94,639	1,126	-711	-711	3,389	28.40	41.08

Note: The table shows summary statistics for the disposable household income per year. Column changes shows statistics for the differences in household's income compared to the status quo. Own calculations based on GSOEP v39, data year 2018.

D.2 QUESTIONNAIRE (GERMAN)

This appendix displays the original questionnaire in German (for an English translation, see Appendix D.3). Text displayed to respondents is shown in black, supplementary information (e.g. about the randomization) are shown in red, answer options are shown in green, and randomized elements of text displayed to respondents is shown in blue.

D.2.1 Allgemeine Reformfragen

[T1-Q1]

Im deutschen Einkommensteuersystem gibt es das Ehegattensplitting. Dies führt unter anderem dazu, dass ein verheiratetes Paar meist weniger Steuern zahlt als ein unverheiratetes Paar mit dem gleichen Gesamteinkommen.

1. Welche der folgenden Reformen befürworten Sie?

Reihenfolge der Reformen randomisieren

a Begrenzung des Steuervorteils für verheiratete Paare (Realsplitting)?
Ja / Nein / Egal / Keine Angabe

b Größerer Steuervorteil für Familien mit Kindern (Familiensplitting)?
Ja / Nein / Egal / Keine Angabe

c Gewährung eines Steuervorteils ausschließlich für Familien mit Kindern (Die steuerliche Belastung von kinderlosen Paaren wäre dann unabhängig davon, ob sie verheiratet sind oder nicht)?
Ja / Nein / Egal / Keine Angabe

d Wechsel zur Individualbesteuerung (Ehepartner werden separat gemäß ihren einzelnen individuellen Einkommen besteuert)?
Ja / Nein / Egal / Keine Angabe

[T1-Q2]

2. Falls das Ehegattensplitting reformiert wird, sollte es einen Bestandsschutz für bestehende Ehen geben und die Neuregelung ausschließlich für neu geschlossene Ehen gelten?

Ja / Nein / Egal / Keine Angabe

D.2.2 Prinzipien der Ehegattenbesteuerung

[T2-Q1]

1. Wie sehr stimmen Sie folgenden Aussagen zu?
Reihenfolge der Aussagen in Tabelle randomisieren
Antworten jeweils: Gar nicht / Eher nicht / Unentschieden / Eher / Sehr

a Die Höhe der Steuerzahlung eines verheirateten Paares sollte nur von dessen Gesamteinkommen und nicht von der Verteilung des Gesamteinkommens zwischen den Ehepartnern abhängen?

b Alle Personen sollten unabhängig von der Höhe ihres Einkommens einen konstanten Prozentsatz ihres Einkommens an Steuern zahlen.

c Ein verheiratetes Paar sollte genauso viel Steuern zahlen wie ein unverheiratetes Paar, wenn beide Paare das gleiche Gesamteinkommen haben.

d Ehepartner sollten wie zwei Einzelpersonen besteuert werden, die beide jeweils die Hälfte des ehelichen Gesamteinkommens verdienen.

e Der absolute steuerliche Vorteil für Ehepaare im Vergleich zu einem unverheirateten Paar sollte mit dem Einkommen des Ehepaars steigen.

f Eine Eheschließung sollte nicht dazu führen, dass sich die Anreize zur Einkommenserzielung für einen der Partner verändern.

g Die Höhe des steuerlichen Vorteils für Ehepaare sollte in der Einkommensteuer gleichmäßig (50/50) auf die beiden Ehepartner aufgeteilt werden.

h Die steuerliche Behandlung von Ehepartnern sollte die Entscheidung der Ehepartner über die Aufteilung der Heimarbeit nicht beeinflussen.

i Die Veränderung der Steuerlast einer Person bei Eheschließung sollte unabhängig davon sein, welches Einkommen der zukünftige Ehepartner hat.

[T2-Q2]

2. Betrachteten Sie ein Ehepaar mit [einem Hauptverdiener/einer Hauptverdienerin] und [einer Zweitverdienerin / einem Zweitverdiener].

Wenn einer der beiden Partner zusätzliches Einkommen erzielt, sollte die steuerliche Belastung dieses zusätzlichen Einkommens

i unabhängig vom Einkommen des anderen Partners sein?

Ja / Nein / Weiß nicht

Falls Nein (Mehrere Optionen möglich)

ii höher für [den Hauptverdiener/die Hauptverdienerin] sein?

iii höher für [den Zweitverdiener/die Zweitverdienerin] sein?

iv mit dem Einkommen des anderen Partners steigen?

v mit dem Einkommen des anderen Partners fallen?

D.2.3 Haushaltspaare

Erklärung:

1. Jeder Teilnehmer wird zu drei Haushaltspaaren befragt. Zu jedem der drei Haushaltspaare werden 6 Fragen gestellt:

   a) Aufteilung Steuerzahlung zwischen Haushalt A und Haushalt B

   b) Grenzsteuersätze Haushalt A

   c) Grenzsteuersätze Haushalt B

   d) Reform Realsplitting

   e) Reform Familiensplitting

   f) Reform Individualbesteuerung

2. Die gezeigten Haushaltspaare werden „mit Zurücklegen“ aus einem Pool (siehe Excel Tabelle) gezogen.

   a) Wenn möglich sollte innerhalb eines Befragten das exakt gleiche Haushaltspaar nicht mehr als einmal vorkommen.

   b) Forsa erhält eine Tabelle mit allen Informationen zu den 3486 Haushaltskombinationen.

3. Im Folgenden sind die Variablennamen aus der Excel-Tabelle rot markiert. Hier soll jeweils die Info für das gezogene Haushaltspaar eingesetzt werden.

4. Bitte das Geschlecht des Haupt- und Zweitverdieners für den Haushalt randomisieren aber dann für eine Runde konstant halten. Neu ziehen für das zweite und dritte gezeigte Haushaltspaar.

Text und Tabelle mit zwei Haushalten bleibt sichtbar für Fragen 1a – 1f

Betrachten Sie im folgenden zwei Haushalte A und B und deren Steuerzahlung im Jahr 2021:

Haushalt A	
Haushaltstyp	[A_hhtyp]
Anzahl der Kinder	[A_anzahlkinder]
Brutto-Einkommen	[A_brutto]
Einkommen Partner 1	[A_brutto_p1]
Einkommen Partner 2	[A_brutto_p2]
Steuerzahlung	[A_steuer _statusquo]

Haushalt B	
Haushaltstyp	[B_hhtyp]
Anzahl der Kinder	[B_anzahlkinder]
Brutto-Einkommen	[B_brutto]
Einkommen Partner 1	[B_brutto_p1]
Einkommen Partner 2	[B_brutto_p2]
Steuerzahlung	[B_steuer _statusquo]

[T3-Q1-A]

Insgesamt haben die beiden Haushalte [Summe_steuer_statusquo] an Steuern gezahlt. Wenn Sie über die Verteilung dieser gesamten Steuerzahlung zwischen den Haushalten entscheiden könnten, wie viel Steuern sollten die beiden Haushalte ihrer Meinung nach jeweils zahlen? Markieren Sie die gewünschte Aufteilung mit Hilfe des Balkens.

Haushalt A — Haushalt B

€[A Steuer] €[B Steuer]

[T3-Q1-B]

Wenn [der Hauptverdiener, die Hauptverdienerin] in **Haushalt A** 100 Euro mehr im Jahr verdient erhält [er, sie] davon nach Steuern [A_nettoZ_statusquo_p1]. Wenn [die Zweitverdienerin, der Zweitverdiener] 100 Euro brutto mehr im Jahr verdient erhält [sie, er] davon nach Steuern [A_nettoZ_statusquo_p2]. Wenn Sie darüber entscheiden könnten, wie viel sollte [der Hauptverdiener, die Hauptverdienerin] und der [die Zweitverdienerin, der Zweitverdiener] von den zusätzlichen 100 Euro jeweils nach Steuern übrigbehalten?

[Hauptverdienter, Hauptverdienerin]

[Zweitverdiener, Zweitverdienerin]

[T3-Q1-C]

Wenn [der Hauptverdiener, die Hauptverdienerin] in **Haushalt B** 100 Euro mehr im Jahr verdient erhält [er, sie] davon nach Steuern [A_nettoZ_statusquo_p1]. Wenn [die Zweitverdienerin, der Zweitverdiener] 100 Euro brutto mehr im Jahr verdient erhält [sie, er] davon nach Steuern [A_nettoZ_statusquo_p2]. Wenn Sie darüber entscheiden könnten, wie viel sollte [der Hauptverdiener, die Hauptverdienerin] und der [die Zweitverdienerin, der Zweitverdiener] von den zusätzlichen 100 Euro jeweils nach Steuern übrigbehalten?

[Hauptverdienter, Hauptverdienerin]

[Zweitverdiener, Zweitverdienerin]

[T3-Q1-D]

Im Vorfeld der Bundestagswahl wurde diskutiert, ob man das Ehegattensplitting durch ein Realsplitting ersetzen soll. Eine Reform würde die Steuerzahlungen der beiden Haushalte im Vergleich zum aktuellen System wie folgt verändern:

Haushalt A Realsplitting	Haushalt B Realsplitting
[A_diff_steuer_realsplitting]	[B_diff_steuer_realsplitting]

Die Reform würde außerdem dazu führen, dass die Personen im Haushalt von jeweils **100 zusätzlich verdienten Euro** nach Steuern folgenden Betrag mehr oder weniger erhalten als im aktuellen System:

Hauptverdiener, Hauptverdienerin	[A_diff_nettoZ _realsplitting_p1]	Hauptverdiener, Hauptverdienerin	[B_diff_nettoZ _realsplitting_p1]
Zweitverdiener, Zweitverdienerin	[A_diff_nettoZ _realsplitting_p2]	Zweitverdiener, Zweitverdienerin	[B_diff_nettoZ _realsplitting_p2]

Würden Sie dieser Reform zustimmen? (Ja / Nein / Keine Angabe)

[T3-Q1-E]

Im Vorfeld der Bundestagswahl wurde diskutiert, ob man das Ehegattensplitting durch ein Familiensplitting ersetzen soll.

Eine Reform würde die Steuerzahlungen der beiden Haushalte im Vergleich zum aktuellen System wie folgt verändern:

Haushalt A Individualbesteuerung	Haushalt B Individualbesteuerung
[A_diff_steuer_indbest]	[B_diff_steuer_indbest]

Die Reform würde außerdem dazu führen, dass die Personen im Haushalt von jeweils **100 zusätzlich verdienten Euro** nach Steuern folgenden Betrag mehr oder weniger erhalten als im aktuellen System:

Hauptverdiener, Hauptverdienerin	[A_diff_nettoZ_indbest_p1]	Hauptverdiener, Hauptverdienerin	[B_diff_nettoZ_indbest_p1]
Zweitverdiener, Zweitverdienerin	[A_diff_nettoZ_indbest_p2]	Zweitverdiener, Zweitverdienerin	[B_diff_nettoZ_indbest_p2]

Würden Sie dieser Reform zustimmen? (Ja / Nein / Keine Angabe)

[T3-Q1-F]

Im Vorfeld der Bundestagswahl wurde diskutiert, ob man das Ehegattensplitting durch Individualbesteuerung ersetzen soll.

Eine Reform würde die Steuerzahlungen der beiden Haushalte im Vergleich zum aktuellen System wie folgt verändern:

Haushalt A Familiensplitting	Haushalt B Familiensplitting
[A_diff_steuer_famsplitting]	[B_diff_steuer_famsplitting]

Die Reform würde außerdem dazu führen, dass die Personen im Haushalt von jeweils **100 zusätzlich verdienten Euro** nach Steuern folgenden Betrag mehr oder weniger erhalten als im aktuellen System:

Hauptverdiener, Hauptverdienerin	[A_diff_nettoZ_famsplitting_p1]	Hauptverdiener, Hauptverdienerin	[B_diff_nettoZ_famsplitting_p1]
Zweitverdiener, Zweitverdienerin	[A_diff_nettoZ_famsplitting_p2]	Zweitverdiener, Zweitverdienerin	[B_diff_nettoZ_famsplitting_p2]

Würden Sie dieser Reform zustimmen? (Ja / Nein / Keine Angabe)

D.2.4 Erklärungen / Mouseover Texte

Mouseover-Erklärungstexte für Teil 1, Frage 1 und Teil 3, Fragen 1d – 1f (Wörter sind in gestrichelter Linie unterstrichen)

- **Ehegattensplitting:** Beim Ehegattensplitting wird das Gesamteinkommen eines verheirateten Paars für die Besteuerung durch einen Multiplikator von 2 geteilt. Damit wird das Gesamteinkommen des Paars gleichmäßig auf die Ehepartner verteilt.

- **Realsplitting:** Während beim Ehegattensplitting das Einkommen unbegrenzt zwischen den Partnern fiktiv übertragen werden kann, begrenzt das Realsplitting den maximal übertragbaren Betrag auf 13.805 Euro. Somit wird der Splitting-Vorteil in seiner Höhe begrenzt. Liegt das Einkommen der Partner mehr als 27.610 Euro auseinander sinkt der Splittingvorteil im Vergleich zum Ehegattensplitting. Wir nehmen in unseren Berechnungen an, dass die durch Minder- bzw. Mehreinnahmen entstehende Belastung bzw. Entlastung gleichmäßig von allen Erwachsenen getragen wird.

- **Familiensplitting:** Während beim Ehegattensplitting die Anzahl der Kinder keine Rolle spielt, bezieht ein Familiensplitting auch Kinder in den Multiplikator ein. So steigt der Multiplikator pro Kind um 0,5. So hätte zum Beispiel ein verheiratetes Paar mit zwei Kindern einen Splitting-Faktor von 3. Über die Progressivität des Steuersystems würden Familien mit Kindern entlastet. Wir nehmen in unseren Berechnungen an, dass die durch Minder- bzw. Mehreinnahmen entstehende Belastung bzw. Entlastung gleichmäßig von allen Erwachsenen getragen wird.

- **Individualbesteuerung:** Während beim Ehegattensplitting verheiratete Paare gemeinsam mit ihrem gesamten Einkommen veranlagt werden, betrachtet die Individualbesteuerung die Ehepartner und deren Einkommen separat. Im Vergleich zum Ehegattensplitting würden so Ehepaare stärker belastet, falls die Ehepartner unterschiedliche Einkommen erzielen. Wir nehmen in unseren Berechnungen an, dass die durch Minder- bzw. Mehreinnahmen entstehende Belastung bzw. Entlastung gleichmäßig von allen Erwachsenen getragen wird.

D.3 QUESTIONNAIRE (ENGLISH TRANSLATION)

This appendix displays a translation of the original questionnaire in German (see Appendix D.2).[1] Text displayed to respondents is shown in black, supplementary information (e.g. about the randomization) are shown in red, answer options are shown in green, and randomized elements of text displayed to respondents is shown in blue.

D.3.1 General Reform Questions

[T1-Q1]

In the German income tax system, there is the concept of "Ehegattensplitting" (tax splitting for married couples). This often results in a married couple paying less taxes than an unmarried couple with the same total income.

1. Which of the following reforms do you support?

Randomize the order of reforms

a Limitation of tax benefits for married couples (Realsplitting)?
Yes / No / Indifferent / No Answer

b Greater tax benefits for families with children (Familiensplitting)?
Yes / No / Indifferent / No Answer

c Granting tax benefits exclusively to families with children (The tax burden on childless couples would then be independent of whether they are married or not)?
Yes / No / Indifferent / No Answer

d Switch to Individual Taxation (Spouses are taxed separately based on their individual incomes)?
Yes / No / Indifferent / No Answer

[1] The primary translation utilized ChatGPT.

[T1-Q2]

2. If "Ehegattensplitting" is reformed, should there be protection for existing marriages, and the new regulation should apply exclusively to newlywed couples?

Yes / No / Indifferent / No Answer

D.3.2 Principles of Couples Taxation

[T2-Q1]

1. How much do you agree with the following statements?
Randomize the order of statements in the table
Answers for each statement: Not at all / Somewhat not / Undecided / Somewhat / Very

a The tax payment of a married couple should depend only on their total income and not on the distribution of total income between the spouses.

b All individuals, regardless of their income level, should pay a constant percentage of their income in taxes.

c A married couple should pay the same amount of taxes as an unmarried couple when both couples have the same total income.

d Spouses should be taxed as two individual persons, each earning half of the marital total income.

e The absolute tax advantage for married couples compared to an unmarried couple should increase with the income of the married couple.

f Marriage should not change the incentives for income generation for one of the partners.

g The tax advantage for married couples should be evenly split (50/50) in income tax between the two spouses.

h The tax treatment of spouses should not affect the decision of spouses regarding the division of home labor.

i The change in the tax burden of an individual upon marriage should be independent of the income of the future spouse.

[T2-Q2]

2. Consider a married couple with [a primary earner] and [a secondary earner].

When one of the two partners earns additional income, should the tax burden on this additional income

i be independent of the income of the other partner?

Yes / No / Don't know

If No (Multiple options possible)

ii be higher for [the primary earner]?

iii be higher for [the secondary earner]?

iv increase with the income of the other partner?

v decrease with the income of the other partner?

D.3.3 Household Pairs

Explanation:

1. Each participant will be asked about three household pairs. For each of the three household couples, 6 questions will be asked:
    a) Allocation of tax payments between Household A and Household B
    b) Marginal tax rates for Household A
    c) Marginal tax rates for Household B
    d) Real Splitting Reform
    e) Family Splitting Reform
    f) Individual Taxation Reform
2. The displayed household couples are drawn "with replacement" from a pool (see Excel table).
    a) If possible, the exact same household couple should not appear more than once for a participant.
    b) Forsa receives a table with all information about the 3486 household combinations.
3. Below, the variable names from the Excel table are highlighted in red. Here, the information for the drawn household couple should be inserted.
4. Please randomize the gender of the primary and secondary earners but then keep it constant for one round. Draw again for the second and third displayed household couples.

Text and table with two households remain visible for questions 1a - 1f

Consider the following two households, A and B, and their tax payments in the year 2021:

Household A	
Household Type	[A_hhtyp]
Number of Children	[A_anzahlkinder]
Gross Income	[A_brutto]
Income Partner 1	[A_brutto_p1]
Income Partner 2	[A_brutto_p2]
Tax Payment	[A_steuer_statusquo]

Household B	
Household Type	[B_hhtyp]
Number of Children	[B_anzahlkinder]
Gross Income	[B_brutto]
Income Partner 1	[B_brutto_p1]
Income Partner 2	[B_brutto_p2]
Tax Payment	[B_steuer_statusquo]

[T3-Q1-A]

In total, both households paid [Summe_steuer_statusquo] in taxes. If you could decide on the distribution of this total tax payment between the households, how much should each household, in your opinion, pay in taxes? Mark the desired distribution using the slider.

Household A — Household B

€[A Steuer] €[B Steuer]

[T3-Q1-B]

If [the primary earner] in **Household A** earns an additional 100 euros per year, they will receive [A_nettoZ_statusquo_p1] after taxes. If [the secondary earner] earns an additional 100 euros gross per year, they will receive [A_nettoZ_statusquo_p2] after taxes. If you could decide how much [the primary earner] and [the secondary earner] should each keep after taxes from the additional 100 euros, what should it be?

[Primary Earner]

0 100

[Secondary Earner]

0 100

[T3-Q1-C]

If [the primary earner] in **Household B** earns an additional 100 euros per year, they will receive [A_nettoZ_statusquo_p1] after taxes. If [the secondary earner] earns an additional 100 euros gross per year, they will receive [A_nettoZ_statusquo_p2] after taxes. If you could decide how much [the primary earner] and [the secondary earner] should each keep after taxes from the additional 100 euros, what should it be?

[Primary Earner]

0 100

[Secondary Earner]

0 100

[T3-Q1-D]

Prior to the federal election, there was a discussion about replacing the joint taxation system for married couples with a real splitting system. A reform would change the tax payments for both households as follows compared to the current system:

Household A Real Splitting	Household B Real Splitting
[A_diff_steuer_realsplitting]	[B_diff_steuer_realsplitting]

The reform would also lead to individuals in each household receiving the following additional amount more or less after taxes from **an additional 100 euros** earned:

Primary Earner	[A_diff_nettoZ _realsplitting_p1]	Primary Earner	[B_diff_nettoZ _realsplitting_p1]
Secondary Earner	[A_diff_nettoZ _realsplitting_p2]	Secondary Earner	[B_diff_nettoZ _realsplitting_p2]

Would you approve of this reform? (Yes / No / No Answer)

[T3-Q1-E]

Prior to the federal election, there was a discussion about replacing the joint taxation system for married couples with a family splitting system.

A reform would change the tax payments for both households as follows compared to the current system:

Household A Individual Taxation	Household B Individual Taxation
[A_diff_steuer_indbest]	[B_diff_steuer_indbest]

The reform would also lead to individuals in each household receiving the following additional amount more or less after taxes from **an additional 100 euros** earned:

Primary Earner	[A_diff_nettoZ _indbest_p1]	Primary Earner	[B_diff_nettoZ _indbest_p1]
Secondary Earner	[A_diff_nettoZ _indbest_p2]	Secondary Earner	[B_diff_nettoZ _indbest_p2]

Would you approve of this reform? (Yes / No / No Answer)

[T3-Q1-F]

Prior to the federal election, there was a discussion about replacing the joint taxation system for married couples with individual taxation.

A reform would change the tax payments for both households as follows compared to the current system:

Household A Family Splitting	Household B Family Splitting
[A_diff_steuer_famsplitting]	[B_diff_steuer_famsplitting]

The reform would also lead to individuals in each household receiving the following additional amount more or less after taxes from **an additional 100 euros** earned:

Primary Earner	[A_diff_nettoZ _famsplitting_p1]	Primary Earner	[B_diff_nettoZ _famsplitting_p1]
Secondary Earner	[A_diff_nettoZ _famsplitting_p2]	Secondary Earner	[B_diff_nettoZ _famsplitting_p2]

Would you approve of this reform? (Yes / No / No Answer)

D.3.4 Explanations / Mouseover Texts

Mouseover explanation texts for Part 1, Question 1, and Part 3, Questions 1d – 1f (Words are underlined with a dashed line)

- **Ehegattensplitting (Marriage Tax Splitting):** In marriage tax splitting, the total income of a married couple is divided by a multiplier of 2 for taxation purposes. This evenly distributes the couple's total income between the spouses.

- **Realsplitting (Real Tax Splitting):** While in marriage tax splitting, income can be fictitiously transferred between partners without limits, real splitting limits the maximum transferable amount to 13,805 euros. This caps the splitting advantage's magnitude. If the partners' incomes differ by more than 27,610 euros, the splitting advantage decreases compared to marriage tax splitting. In our calculations, we assume that any resulting increase or decrease in income is evenly shared among all adults.

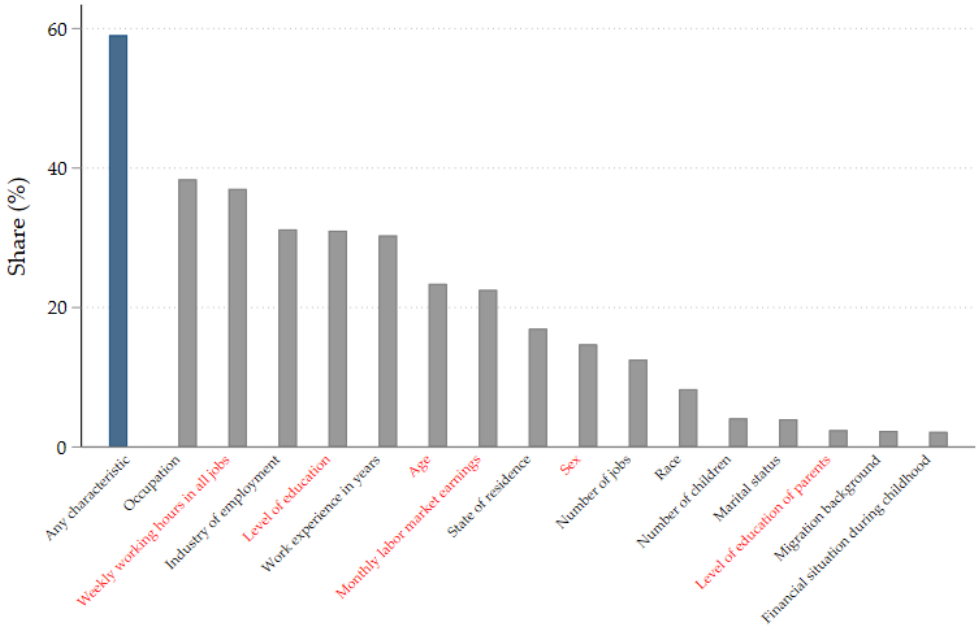
- **Familiensplitting (Family Tax Splitting):** While marriage tax splitting does not consider the number of children, family tax splitting includes children in the multiplier. The multiplier increases by 0.5 per child. For example, a married couple with two children would have a splitting factor of 3. Through the progressiveness of the tax system, families with children receive tax relief. In our calculations, we assume that any resulting increase or decrease in income is evenly shared among all adults.

- **Individualbesteuerung (Individual Taxation):** Unlike marriage tax splitting, where married couples are jointly assessed with their total income, individual taxation considers the spouses and their incomes separately. Compared to marriage tax splitting, this would result in higher taxes for married couples if the spouses have different incomes. In our calculations, we assume that any resulting increase or decrease in income is evenly shared among all adults.

Appendix to Chapter 5

E.1 SURVEY TOOL AND DATA COLLECTION

Figure E.1. – Distribution Task, Chosen Characteristics



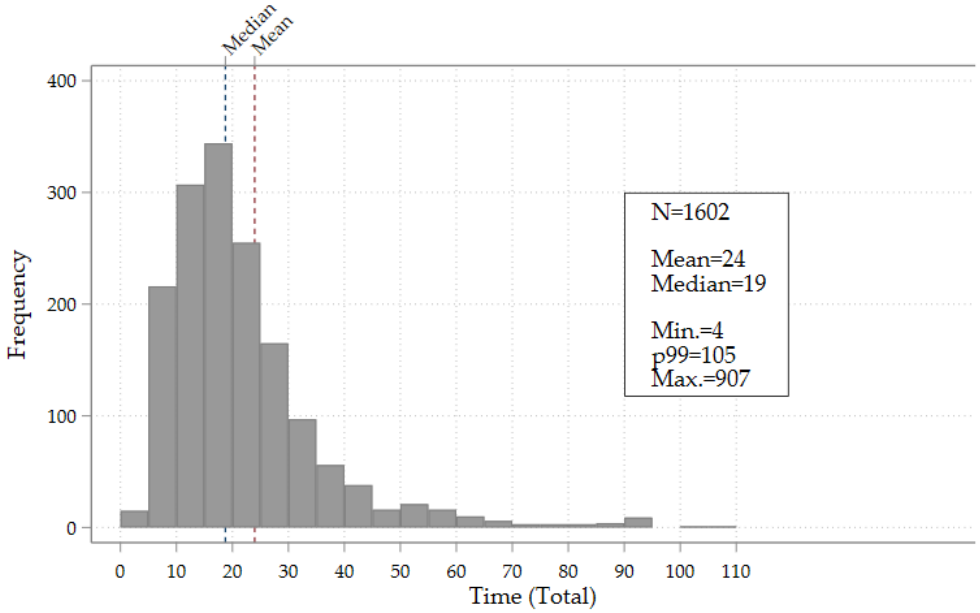
Note: The figure shows results from the fourth module of the baseline survey that refers to other questions about the labor market. The figure illustrates what share of respondents chose to get to know a certain characteristic when asked to decide on the fair earnings split between two persons. In the first step, respondents were informed about the distribution task and asked whether they would like to know anything about the two persons. In the second step, respondents could choose up to five characteristics that they would see for the two persons before redistributing earnings. Characteristics that we chose ex-ante to be included in the main survey module on fairness preferences are highlighted in red.
Source: Own calculations based on survey responses.

Table E.1. – Characteristics of the Real Person

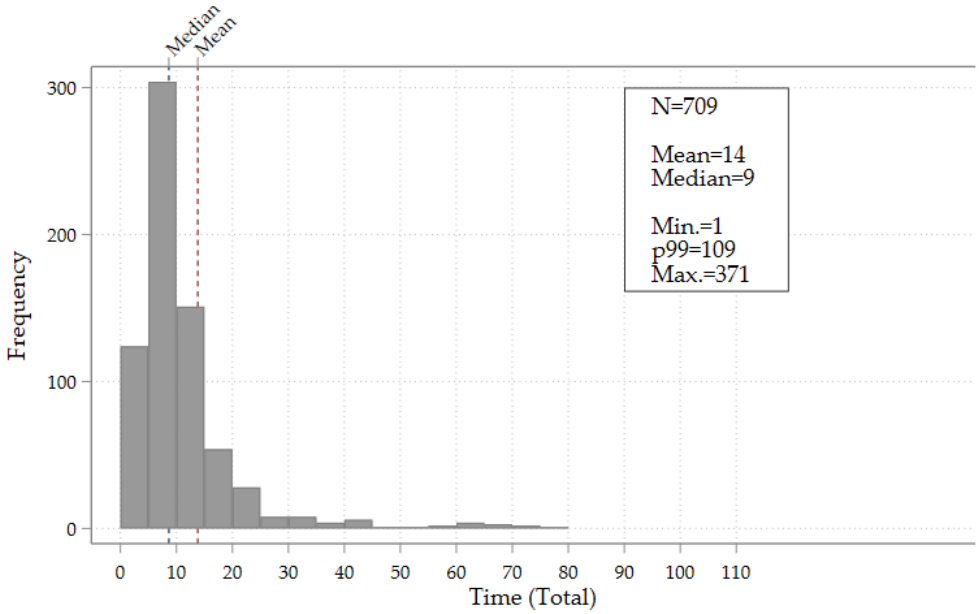
Characteristic	Expression
Gender	Female
Age	26
Education	University
Education of Parents	High School
Working Hours	20
Monthly Earnings	$4,100

Note: This table displays the expressions of characteristics for the real person who was hired for one month. The initial monthly earnings were drawn randomly from the five available values displayed in Table 4.1.

Figure E.2. – Survey Completion Time



(A) Baseline Wave



(B) Follow-Up Wave

Note: This figure displays the frequency distribution of total survey completion time using fixed five-minute bins for the baseline wave (Figure E.2a) and the follow-up wave (Figure E.2b). For better visibility, the figure only shows the distribution for values below the 99th percentile of the respective survey completion time distribution. The vertical blue (red) lines display the median (mean) of the survey completion time. Further information on the respective distribution is shown in the box next to the histogram.
Source: Own calculations based on survey responses.

Table E.2. – Demographics

Variable	Share Baseline (%)	Share Follow-Up (%)
Gender		
Male	48.19	47.95
Female	51.00	51.90
Non-Binary	0.81	0.14
Age		
18 - 24	12.98	5.92
25 - 34	13.98	7.33
35 - 44	16.29	13.68
45 - 54	15.61	10.58
55 - 64	18.04	25.25
Older than 64	23.10	37.24
Education		
Less than Middle School	0.37	0.28
Middle School	3.93	3.10
High School Graduate	39.95	35.68
Completed Some College	16.79	14.95
College Degree	25.03	29.06
Master's Degree	10.11	11.42
Doctoral Degree / Law or Professional Degree	3.81	5.50
Ethnicity		
White / Caucasian	81.09	88.15
Black / African American	9.43	5.92
American Indian	1.12	0.42
Alaska Native	0.19	0.00
Asian / Asian American	3.68	3.24
Native Hawaiian	0.31	0.14
Pacific Islander	0.06	0.00
Other	3.87	2.12
Prefer not to answer	0.25	0.00
Hispanic		
Yes	10.05	6.49
No	89.83	93.51
Prefer not to answer	0.12	0.00
Region		
North-East	19.35	24.68
Mid-West	33.52	32.44
West	23.91	18.62
South	23.22	24.26

Continued on next page.

Table E.2. – Demographics (cont.)

Variable	Share Baseline (%)	Share Follow-Up (%)
Employment Status		
Employed	53.00	47.53
Self-Employed	8.93	8.32
Unemployed	5.74	3.67
Student	3.68	1.13
Retiree	28.65	39.35
Working Hours		
9 hours/week or less	27.53	34.27
10-34 hours/week	16.92	12.41
35-44 hours/week	35.83	35.26
More than 44 hours/week	11.36	9.73
Prefer not to answer	8.36	8.32
Income		
Did not work	22.28	27.79
$0 - $25,000	19.41	12.98
$25,000 - $40,000	17.04	12.55
$40,000 - $59,000	14.17	13.82
$59,000 - $88,000	12.23	14.25
More than $88,000	14.67	18.62
Prefer not to answer	0.19	0.00
Education of Mother		
Less than Middle School	2.00	2.82
Middle School	7.99	8.46
High School Graduate	47.50	48.52
Completed Some College	12.80	8.74
College Degree	19.66	21.02
Master's Degree	7.30	7.19
Doctoral Degree / Law or Professional Degree	1.81	2.26
Prefer not to answer	0.94	0.99
Education of Father		
Less than Middle School	3.43	4.37
Middle School	11.61	14.39
High School Graduate	44.19	42.03
Completed Some College	11.74	9.31
College Degree	17.79	18.90
Master's Degree	5.93	6.21
Doctoral Degree / Law or Professional Degree	3.00	3.24
Prefer not to answer	2.31	1.55

Note: This table presents summary statistics for demographics in the baseline wave (N=1602) and the follow-up wave (N=709). Information on respondent demographics are obtained from the demographics module of the baseline wave of the survey. We also include answers of those respondents that did prefer not to answer specific demographic questions.
Source: Own calculations based on survey responses.

E.2 EFFECTS OF THE "REAL-PERSON TREATMENT"

Table E.3. – Joint Balancing Tests

Demogr.	Binary Split	Dep. Variable	*N*	Point Estimate	Control Mean	RW p-values
Gender	Male vs. Female	TreatB	1589	0.027	0.5006	0.964
		TreatC	1589	-0.007	0.5176	1.000
Age	$<$ 45 vs. $\geq$ 45 years	TreatB	1602	-0.010	0.5725	1.000
		TreatC	1602	-0.050	0.5925	0.418
Ethnicity	White/Caucasian vs. Other	TreatB	1598	0.008	0.1830	1.000
		TreatC	1598	0.006	0.1842	1.000
Education	$\leq$ HS vs. $>$ HS	TreatB	1602	0.005	0.5550	1.000
		TreatC	1602	0.037	0.5387	0.818
Employment	(Self)Employed vs. Other	TreatB	1602	-0.031	0.3962	0.922
		TreatC	1602	-0.018	0.3900	0.996
Hours	$<$ 35 vs. $\geq$ 35 hours	TreatB	1468	-0.003	0.5163	1.000
		TreatC	1468	-0.019	0.5249	0.996
Income	$<$ \$40000 vs. $\geq$ \$40000	TreatB	1599	0.003	0.4098	1.000
		TreatC	1599	-0.003	0.4128	1.000

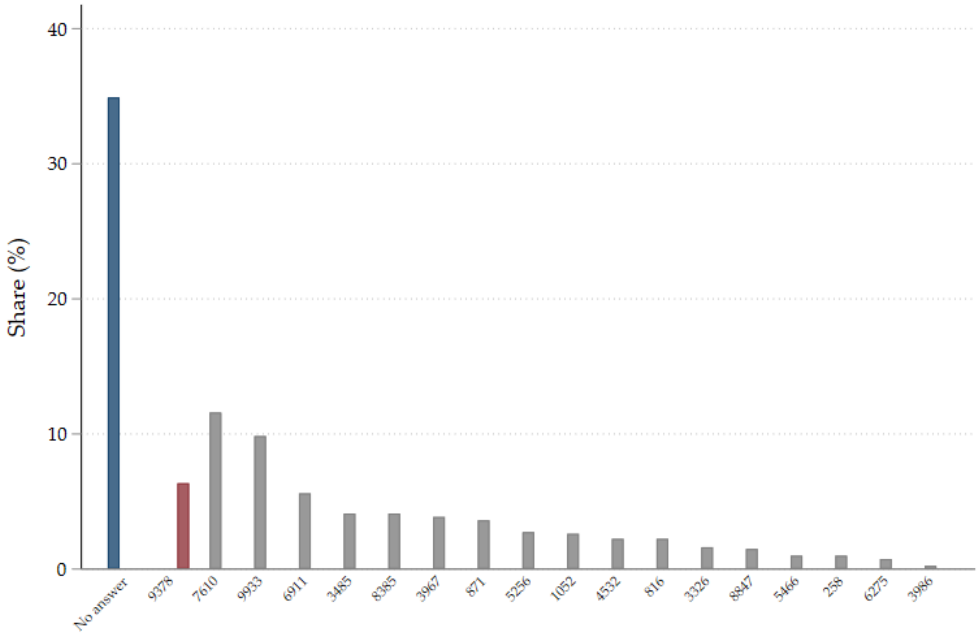
Note: This table presents results of the joint balancing test outlined in the pre-analysis plan. The table shows point estimates for the coefficient of interest, the mean of the control group (first group in binary split), and associated heteroskedasticity robust p-values adjusted for multiple hypothesis testing using the Romano and M. Wolf (2016) step-down procedure with 1000 bootstrap draws. *TreatB* refers to the treatment in the module on inequality perceptions. *TreatC* refers to the treatment in the module on fairness preferences. Sample sizes differ across demographic characteristics due to the exclusion of responses such as "Prefer not to answer" (see Table 5.2).
Source: Own calculation based on survey responses.

Table E.4. – Allocated Shares, "Real-Person Treatment"

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	1602	1.06	59.611	0.324	0.322	0.841
Q2	1602	0.08	27.904	0.931	0.932	0.932
Q3	1602	0.68	56.569	0.591	0.598	0.841
Q4	1602	-0.83	36.864	0.321	0.325	0.841
Q5	1602	-0.85	69.950	0.343	0.337	0.841

Note: This table presents results of the regression analysis outlined in equation (5.2) where the dependent variable is the share allocated to person A. We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculation based on survey responses.

Figure E.3. – Guess of the Real Person ID, Shares



Note: This figure displays the shares of different guesses for the real person for those respondents in the preference treatment group with financial incentives ($N = 802$). The blue bar indicates the share of respondents who did not provide any answer, while the red bar represents the share of respondents who guessed the real person correctly. Other guesses are displayed as gray bars. A correct guess was incentivized with $5.
Source: Own calculations based on survey responses.

Table E.5. – Distribution of Allocations, Kolmogorov–Smirnov Test

Question	N	N (unique)	D	p-value (approx.)
Q1	1602	1039	0.04	0.399
Q2	1602	927	0.05	0.313
Q3	1602	1196	0.04	0.414
Q4	1602	1066	0.03	0.853
Q5	1602	1003	0.04	0.443

Note: This table presents results from two-sample Kolmogorov–Smirnov (KS) test to determine if there are any differences in the distribution of the allocations between the treatment and control groups. The KS test is based on a test statistic that measures the maximum absolute (vertical) difference between the cumulative distribution functions of the two groups. The table shows for each question the maximum absolute difference D. The p-values of this test statistic are based on a five-term approximation of the asymptotic distributions. Given that the KS test is designed for continuous distributions without value ties, the table also presents the number of unique values for each question.
Source: Own calculation based on survey responses.

Table E.6. – Balancing Tests, Alternative Sample Restrictions

Panel A: Directly Passed Understanding Question

Demogr.	Binary Split	*N*	Point Estimate	Control Mean	RW p-values
Gender	Male vs. Female	1144	0.007	0.5017	0.993
Age	$<$ 45 years vs. $\geq$ 45 years	1154	-0.039	0.6478	0.709
Ethnicity	White/Caucasian vs. Other	1151	0.003	0.1583	0.993
Education	$\leq$ HS vs. $>$ HS	1154	0.022	0.5773	0.945
Employment	(Self)Employed vs. Other	1154	-0.031	0.4141	0.837
Hours	$<$ 35 hours vs. $\geq$ 35 hours	1060	-0.008	0.5028	0.993
Income	$<$ \$40,000 vs. $\geq$ \$40,000	1152	0.013	0.4172	0.987

Panel B: Exclude Response Time Outliers

Demogr.	Binary Split	*N*	Point Estimate	Control Mean	RW p-values
Gender	Male vs. Female	915	0.010	0.4957	0.993
Age	$<$ 45 years vs. $\geq$ 45 years	924	-0.070	0.6858	0.171
Ethnicity	White/Caucasian vs. Other	923	-0.018	0.1592	0.948
Education	$\leq$ HS vs. $>$ HS	924	0.004	0.5987	0.993
Employment	(Self)Employed vs. Other	924	-0.054	0.4289	0.441
Hours	$<$ 35 hours vs. $\geq$ 35 hours	851	-0.004	0.5000	0.993
Income	$<$ \$40,000 vs. $\geq$ \$40,000	922	0.011	0.4286	0.993

Note: This table presents results from the balancing test outlined in equation (5.1) for different restricted samples. Panel A displays results for the sample of respondents that passed the understanding question of the fairness preference module at the first try. Panel B presents results for a sample that excludes respondents with high (above p90) and low (below p10) response times of the fairness preference module. The table presents point estimates for the coefficient of interest, the mean of the control group (first group in binary split), and associated heteroskedasticity robust p-values adjusted for multiple hypothesis testing using the Romano and M. Wolf (2016) step-down procedure with 1000 bootstrap draws. Sample sizes differ across demographic characteristics due to the exclusion of responses such as "Prefer not to answer" (see Table 5.2).
Source: Own calculation based on survey responses.

Table E.7. – Distribution of Allocations, Kolmogorov–Smirnov Test, Alternative Samples

Panel A: Directly Passed Training Question

Question	N	N (unique)	D	p-value (approx.)
Q1	1154	825	0.05	0.399
Q2	1154	693	0.06	0.306
Q3	1154	904	0.04	0.691
Q4	1154	789	0.03	0.941
Q5	1154	742	0.05	0.385

Panel B: Exclude Response Time Outliers

Question	N	N (unique)	D	p-value (approx.)
Q1	924	700	0.08	0.126
Q2	924	583	0.08	0.103
Q3	924	756	0.07	0.257
Q4	924	655	0.05	0.581
Q5	924	628	0.07	0.253

Note: This table presents results from two-sample Kolmogorov–Smirnov (KS) test for different restricted samples. Panel A displays results for the sample of respondents that passed the training question of the fairness preference module at the first try. Panel B presents results for a sample that excludes respondents with high (above p90) and low (below p10) response times of the fairness preference module. The test determines if there are any differences in the distribution of the allocations between the treatment and control groups. The KS test is based on a test statistic that measures the maximum absolute (vertical) difference between the cumulative distribution functions of the two groups. The table shows for each question the maximum absolute difference D. The p-values of this test statistic are based on a five-term approximation of the asymptotic distributions. Given that the KS test is designed for continuous distributions without value ties, the table also presents the number of unique values for each question.
Source: Own calculation based on survey responses.

Table E.8. – Allocations, "Real-Person Treatment", Male

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	772	19.77	1139.31	0.909	0.907	0.945
Q2	772	97.57	-2355.06	0.490	0.485	0.857
Q3	772	-478.21	2666.53	0.407	0.396	0.846
Q4	772	-108.54	-4311.73	0.766	0.764	0.945
Q5	772	-691.37	6579.08	0.095	0.104	0.356

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.9. – Allocations, "Real-Person Treatment", Female

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	817	176.46	943.63	0.260	0.249	0.652
Q2	817	-54.76	-2262.16	0.684	0.685	0.791
Q3	817	812.18	1483.75	0.135	0.144	0.471
Q4	817	-384.54	-3893.72	0.296	0.286	0.652
Q5	817	219.52	5812.89	0.560	0.563	0.791

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.10. – Allocations, "Real-Person Treatment", Age below 45

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	693	-120.46	917.07	0.456	0.469	0.893
Q2	693	81.05	-2148.02	0.560	0.556	0.908
Q3	693	132.49	826.70	0.812	0.805	0.915
Q4	693	142.14	-3988.26	0.712	0.692	0.915
Q5	693	-667.37	5687.35	0.103	0.119	0.389

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.11. – Allocations, "Real-Person Treatment", Age 45 and above

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	909	331.63	1121.19	0.041	0.053	0.169
Q2	909	-76.33	-2401.16	0.569	0.559	0.810
Q3	909	467.27	2868.35	0.389	0.383	0.735
Q4	909	-605.29	-4129.73	0.082	0.083	0.273
Q5	909	153.77	6526.30	0.682	0.688	0.810

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.12. – Allocations, "Real-Person Treatment", White / Caucasian Ethnicity

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	1299	136.48	1091.64	0.305	0.278	0.739
Q2	1299	-50.15	-2337.88	0.646	0.648	0.739
Q3	1299	443.18	2294.39	0.320	0.313	0.739
Q4	1299	-509.04	-4020.58	0.075	0.077	0.298
Q5	1299	-228.71	6350.19	0.463	0.465	0.739

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.13. – Allocations, "Real-Person Treatment", Non White / Non Caucasian Ethnicity

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	299	58.40	798.94	0.797	0.782	0.839
Q2	299	260.73	-2121.51	0.221	0.223	0.665
Q3	299	-698.25	915.12	0.392	0.397	0.759
Q4	299	761.15	-4262.26	0.206	0.209	0.665
Q5	299	-323.56	5420.10	0.600	0.623	0.839

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.14. – Allocations, "Real-Person Treatment", High School Education and Below

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	709	148.21	736.75	0.372	0.374	0.820
Q2	709	-29.74	-2207.45	0.838	0.836	0.878
Q3	709	338.06	787.28	0.566	0.564	0.878
Q4	709	-798.36	-3754.04	0.042	0.047	0.184
Q5	709	-268.66	5817.89	0.520	0.508	0.878

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.15. – Allocations, "Real-Person Treatment", More Than High School Education

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	893	53.52	1295.94	0.738	0.723	0.971
Q2	893	47.31	-2375.54	0.718	0.725	0.971
Q3	893	-31.33	3105.79	0.952	0.952	0.971
Q4	893	181.08	-4344.37	0.597	0.605	0.966
Q5	893	-304.67	6498.24	0.414	0.410	0.912

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.16. – Allocations, "Real-Person Treatment", (Self-) Employed

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	992	104.00	1063.81	0.468	0.487	0.903
Q2	992	-26.56	-2184.63	0.825	0.826	0.903
Q3	992	577.65	1482.95	0.238	0.228	0.678
Q4	992	-228.93	-3958.98	0.489	0.500	0.903
Q5	992	-217.43	5883.71	0.538	0.564	0.903

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.17. – Allocations, "Real-Person Treatment", Unemployed/Student/Retiree

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	610	129.04	997.67	0.511	0.510	0.939
Q2	610	53.03	-2475.35	0.746	0.740	0.939
Q3	610	-336.65	2901.99	0.609	0.600	0.939
Q4	610	-316.62	-4248.97	0.444	0.484	0.939
Q5	610	-304.58	6654.78	0.500	0.507	0.939

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.18. – Allocations, "Real-Person Treatment", Working Less Than 35 Hours/Week

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	712	-51.47	1160.55	0.774	0.778	0.928
Q2	712	66.53	-2362.80	0.649	0.634	0.928
Q3	712	-388.03	2460.73	0.521	0.524	0.928
Q4	712	-243.60	-4156.77	0.522	0.528	0.928
Q5	712	-466.87	6390.95	0.266	0.249	0.747

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.19. – Allocations, "Real-Person Treatment", Working 35 Hours/Week or More

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	756	214.37	1078.48	0.198	0.179	0.580
Q2	756	12.33	-2261.43	0.930	0.930	0.930
Q3	756	564.42	1923.83	0.320	0.286	0.740
Q4	756	-178.92	-3953.36	0.637	0.633	0.920
Q5	756	-216.74	6108.81	0.589	0.573	0.920

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.20. – Allocations, "Real-Person Treatment", Earnings Less Than 40000

Question	N	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	941	46.18	997.80	0.753	0.727	0.977
Q2	941	-5.75	-2314.82	0.963	0.965	0.995
Q3	941	43.72	1641.26	0.932	0.939	0.995
Q4	941	-260.73	-4050.91	0.450	0.426	0.927
Q5	941	-171.49	6122.00	0.630	0.630	0.970

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.21. – Allocations, "Real-Person Treatment", Earnings 40000 or More

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	658	213.33	1095.28	0.258	0.257	0.732
Q2	658	36.73	-2281.88	0.813	0.828	0.859
Q3	658	391.29	2666.52	0.521	0.520	0.859
Q4	658	-266.91	-4085.23	0.495	0.496	0.859
Q5	658	-403.93	6281.06	0.368	0.381	0.791

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculations based on survey responses.

Table E.22. – Allocations, "Real-Person Treatment", 5 Questions

Question	*N*	Point Estimate	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	534	248.68	0.211	0.203	0.562
Q2	534	156.33	0.377	0.389	0.717
Q3	534	-543.47	0.434	0.432	0.717
Q4	534	-102.52	0.823	0.803	0.803
Q5	534	-964.28	0.053	0.058	0.223

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, and associated heteroskedasticity robust model p-values, as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications. We focus on those respondents that answer exactly five questions.
Source: Own calculations based on survey responses.

Table E.23. – Allocations, "Real-Person Treatment", 7 Questions

Question	*N*	Point Estimate	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	536	-42.11	0.830	0.835	0.991
Q2	536	115.69	0.479	0.462	0.942
Q3	536	-464.22	0.480	0.468	0.942
Q4	536	-454.02	0.299	0.315	0.864
Q5	536	-114.90	0.803	0.811	0.991
Q6	536	-17.16	0.965	0.957	0.991
Q7	536	-109.54	0.434	0.433	0.942

Note: This table presents results of the regression analysis outlined in equation (5.2). We present point estimates for the coefficient of interest β^j, and associated heteroskedasticity robust model p-values, as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications. We focus on those respondents that answer exactly seven questions.
Source: Own calculations based on survey responses.

Table E.24. – Allocations, "Real-Person Treatment", 9 Questions

Question	*N*	Point Estimate	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	532	137.73	0.510	0.514	0.961
Q2	532	-248.76	0.131	0.133	0.571
Q3	532	1653.34	0.017	0.013	0.121
Q4	532	-209.97	0.639	0.636	0.961
Q5	532	291.45	0.550	0.560	0.961
Q6	532	43.65	0.913	0.911	0.961
Q7	532	122.00	0.362	0.381	0.913
Q8	532	397.68	0.537	0.518	0.961
Q9	532	832.08	0.189	0.192	0.684

Note: This table presents results of the regression analysis outlined in equation (5.2). We present the number of respondents, the number of rejected hypotheses according to heteroskedasticity robust model p-values, resample p-values, and p-values adjusted for multiple hypothesis testing. We focus on those respondents that answer exactly nine questions.
Source: Own calculations based on survey responses.

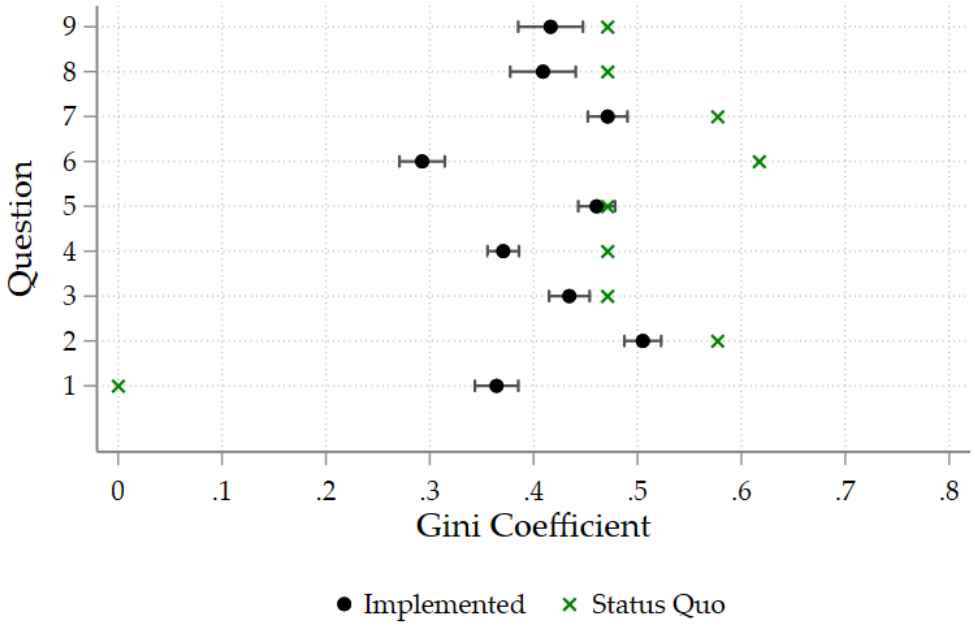
Table E.25. – Preferences, Response Time (Min.), "Real-Person Treatment", No Limit

Question	*N*	Point Estimate	Control Mean	Model p-value	Resample p-value	Romano-Wolf p-value
Q1	1602	-0.28	1.00	0.246	0.276	0.758
Q2	1602	0.03	0.46	0.490	0.513	0.758
Q3	1602	-0.03	0.50	0.476	0.483	0.758
Q4	1602	-0.02	0.37	0.344	0.386	0.758
Q5	1602	-0.04	0.36	0.242	0.280	0.758

Note: This table presents results of the regression analysis outlined in equation (5.2) using the response time in minutes as the dependent variable. We present point estimates for the coefficient of interest β^j, the mean of the control group, and heteroskedasticity robust uncorrected analytical p-values (model p-values), uncorrected bootstrapped p-values (resample p-values), as well as p-values adjusted for multiple hypothesis testing using 1000 bootstrap replications.
Source: Own calculation based on survey responses.

E.3 NATURE OF FAIRNESS PREFERENCES

Figure E.4. – Gini Coefficient, Treatment Group



Note: This figure compares implemented Gini coefficients (gray dots) to initial Gini coefficients (green crosses) in each vignette. Gray bars indicate 95 percent confidence intervals. Gini coefficients are calculated at the respondent level as $\frac{|x-y|}{x+y}$ where x (y) is the amount allocated to vignette person A (B). We focus exclusively on respondents in the treatment group, i.e., those respondents who were informed about the real-world consequences of their choices.
Source: Own calculations based on survey responses.

Table E.26. – Preference Types, Treatment Group

	Share Egalitarians (%)			Share Libertarians (%)		
Max. abs. difference (pp)	2	5	10	2	5	10
Allow for 0 inconsistent answers	0.72	1.07	2.32	5.74	6.27	8.13
Allow for 1 inconsistent answers	1.07	**2.17**	5.58	7.66	**9.04**	12.95
Allow for 2 inconsistent answers	1.49	3.82	10.49	10.36	13.28	22.37

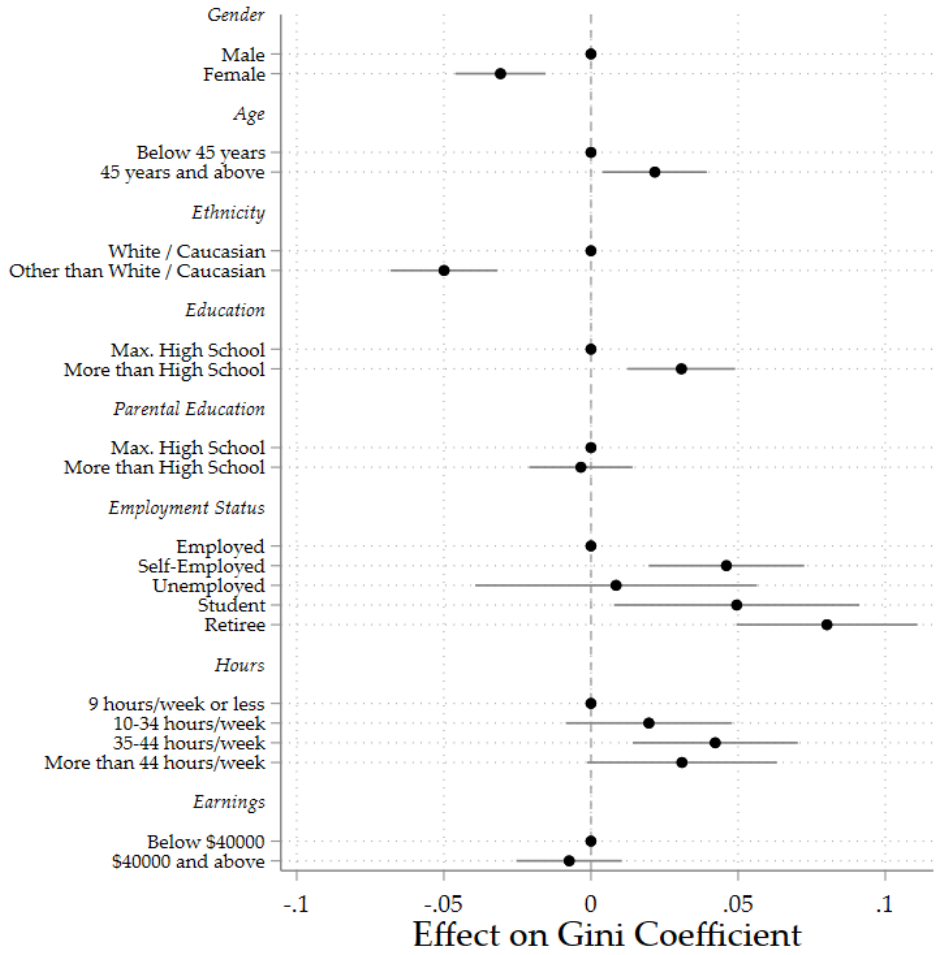
Note: This table presents shares of egalitarians (libertarians) according to consistent choices in all questions of the baseline wave. We also vary the leniency of the classification by allowing for 0, 1, 2 answers that are inconsistent with egalitarian (libertarian) choices. In the baseline (highlighted estimates), we allow for a deviation of +/-5 percentage points and inconsistent choices in one question only. We focus exclusively on respondents in the treatment group, i.e., those respondents who were informed about the real world consequences of their decisions.
Source: Own calculations based on survey responses.

Table E.27. – Fairness Preferences, Regression Table

Vignette Person Char.	Log (Preferred Earnings)		Preferred Share (%)		Log(Preferred Ratio)	
Female (binary)	0.009	0.010	-0.149	-0.136	-0.009	-0.008
	(0.511)	(0.456)	(0.738)	(0.761)	(0.736)	(0.756)
Age (binary)	0.154	0.152	3.409	3.390	0.202	0.201
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (cat. 1,2,3)	0.091	0.089	2.411	2.394	0.138	0.138
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Parental Education (cat. 1,2,3)	-0.047	-0.045	-3.872	-3.851	-0.212	-0.211
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Working Hours (cat. 1,2,3,4)	0.233	0.234	7.950	7.956	0.436	0.437
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(Earnings)	0.380	0.379	.	.	.	.
	(0.000)	(0.000)				
Earnings Share	.	.	4.136	4.132	.	.
			(0.000)	(0.000)		
Log(Ratio)	.	.	.	.	0.428	0.427
					(0.000)	(0.000)
Constant	4.149	4.163	17.008	17.025	-0.694	-0.695
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sum of Earnings FE	Yes	Yes	Yes	Yes	Yes	Yes
Respondent FE	No	Yes	No	Yes	No	Yes
Number of Observations	10980	10980	10980	10980	10980	10980

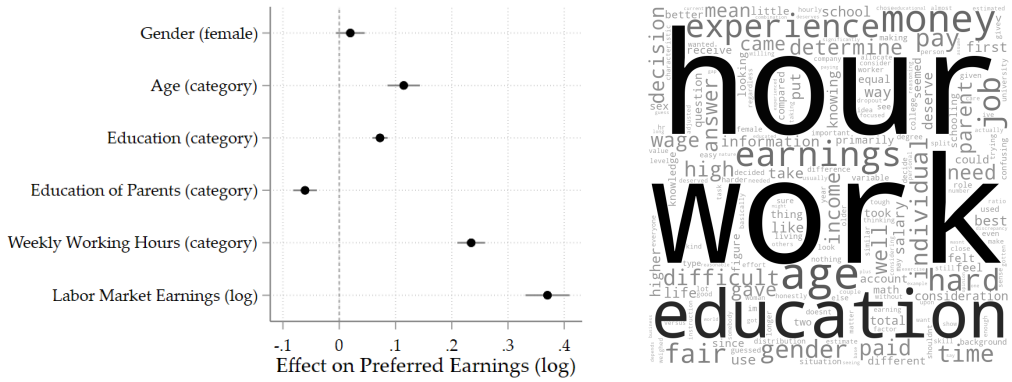
Note: This table presents regression results from equation 5.4 where we regress the outcome variable of interest on linearized categorical information on different vignette person characteristics (column 1). The regression further controls for the total amount of earnings in a person pair and, as a robustness test, for respondent fixed effects. The table provides regression results for three outcome variables of interest, i.e., the log value of preferred earnings (column 2-3), the preferred earnings share (column 4-5), and the log value of the preferred ratio (column 5-6). We focus exclusively on respondents in the control group, i.e., those respondents who faced hypothetical scenarios. Point estimates are shown with respective p-values based on heteroskedasticity robust standard errors clustered at the respondent level in brackets. To keep the number of observations constant across columns, we focus on non-zero earnings allocations across all specifications.
Source: Own calculations based on survey responses.

Figure E.5. – Gini Coefficient and Respondent Demographics



Note: This figure displays the coefficients from a regression that regresses the implemented Gini coefficient on various demographic characteristics, i.e., gender, age, ethnicity, education, parental education, employment status, working hours, and labor market earnings of the respondent. The graph includes the base level, the point estimate for the other level(s), as well as 95 percent confidence intervals. We focus exclusively on respondents in the control group, i.e., those respondents who faced hypothetical scenarios.
Source: Own calculations based on survey responses.

Figure E.6. – Fairness Preferences, Importance of Profile Characteristics, Treatment Group



(A) Regression Analysis (B) Text Analysis

Note: This figure displays how respondents take the displayed characteristics in a profile into account for their distributional choices. Figure E.6a displays the coefficients from a regression analysis following equation (5.4) that regresses the preferred earnings (log) on linearized categorical information on the person's characteristics. The regression further controls for the total amount of earnings in a person pair. The figure shows point estimates and heteroskedasticity robust 95 percent confidence intervals. Figure E.6b displays a word cloud from a text analysis of the open-ended question in the fairness preference module. At the end of the module, people were asked about how they came up with their distributional choices and to describe their reasoning in their own words. Based on the text corpus from the open-ended answers, we used natural language processing techniques to rank the frequency of specific terms. We focus exclusively on respondents in the treatment group, i.e., those respondents who were informed about the real-world consequences of their choices.
Source: Own calculation based on survey responses.

Appendix to Chapter 6

F.1 INSTITUTIONAL REGULATIONS

Table F.1. – Part-Time Regulations Around the World

Country	Description	Sources
Austria	There are two cases, in which there exists specific rights to part-time work in Austria. First, in case of a conditional illness, employees are entitled to work part-time for a period between one and six months, during which they can reduce their hours by 25-50 percent. Second, parents are entitled to reduce their hours worked after completing their parental time-off until the seventh birthday of the child. The entitlement is granted to employees in companies with more than thirty employees who have worked in the respective firm for at least three years without interruption. Another form of part-time work is partial retirement, which employees do not have an entitlement to but which is strongly subsidized by the government	Arbeitszeitgesetz, Gleichbehandlungsgesetz
Canada	There is no legal difference between part-time and full-time workers, i.e. any minimum employee benefits provided to full-time workers, must also be provided to part-time workers (at a pro-rata basis where applicable). However, any entitlements to request part-time work are negotiated between trade unions and employers.	Federal Canada Labour Code
Denmark	Part-time work refers to less than regular hours at the specific workplace. The coverage of part-time workers under the Danish Salaried Employees Act refers to individuals working more than eight hours per week. Part-time workers cannot be treated differentially unless justified by objective reasons. Part-time workers have the same number of holiday days. Part-time salaries are calculated proportionally	Danish Salaried Employees Act, Danish government
France	In France, part-time work is defined by weekly working hours between 24 and 35 hours. Usually, collective agreements organize the conditions for part-time work and corresponding rights to request it. There are, however, some general rights to part-time work for specific reasons (family reasons, parental leave, seriously ill loved one, creation of a company, arduous job). Part-time workers have the same rights as full-time employees, i.e. all salaries, severance pay, and retirement pay are allocated on a proportional basis while number of paid leave days and the calculation of seniority are the same for part-time and full-time workers.	Le Code du travail

Note: continued on next page.

Table F.2. – Part-Time Regulations Around the World (cont.)

Country	Description	Sources
Netherlands	Usually, industry specific agreements take precedent over the labour laws that govern the right to part-time work. In the absence of such agreements, firms with fewer than 10 employees must consider any requests to reduce working hours and must not discriminate on the basis of a request. Firms with 10 or more employees are required to grant applications for part-time work if the employee has been employed for at least six months (26 weeks). Any rejection must be justified on the basis of objective business-related reasons that could negatively impact the firm.	
United Kingdom	Specific rights to request part-time in case of child care were introduced in 2002, and extended to specific rights in case of elderly care. in 2006. Since 2014, there is a general right to request part-time work. Part-time workers have the same rights than comparable full-time worker. Both specific and general rights are applicable to employees with a minimum employment length of 26 weeks and can make one request in any twelve-month period. By default, the reduction of work hours is permanent if not negotiated otherwise. Employers can reject requests on a number of business reasons and there is no right of appeal. In July 2023, parliament passed an extension of the right to request part-time work. In particular, employees can now make two requests in any twelve-month period and do not need to specify themselves the potential effects on the company. In addition, the employer is required to consult with other employees before refusing the request.	Part-time Workers Regulations (2000), Employment Act (2002), Work and Families Act (2006), The Flexible Working Regulations (2009), Children and Families Act (2014), Employment Relations Act (2023).
United States	In the United States, there are few regulations governing the relation between part-time and full-time work. There is no right to request part-time work. Moreover, some government benefits are only available for workers with certain amount of hours (e.g. hours requirement for basic health insurance under the affordable care act, unpaid leave after child birth or care for a family member). There are no federal laws that ensure equal treatment of full-time and part-time employees, for instance in hourly wages or sick pay.	Fair Labor Standards Act (1938), Family and Medical Leave Act (1993), Affordable Care Act (2010).

Note: This table describes part-time regulations and policies related to an entitlement for part-time work for different countries across the world.

Table F.3. – Entitlement to Part-Time Work in Germany over Time

Enactment	Name	Type	Length	Min. firm size (employees)	Min. employment length (months)	Work requirement (hours)
01.01.2001	Permanent part-time entitlement	General	Non-temporary	15	6	N/A
01.01.2007	Part-time entitlement during parental leave	Specific (parental leave)	Temporary (2 months - 3 years)	15	N/A	15-32
01.01.2008	Part-time entitlement during home care provision	Specific (home care)	Temporary (up to 6 months)	15	N/A	N/A
01.01.2015	Part-time entitlement during home care provision	Specific (home care)	Temporary (up to 2 years)	25	N/A	≥ 15
01.01.2019	Temporary part-time entitlement	General	Temporary (1-5 years)	45 (200)	6	N/A

Note: This table shows the evolution of part-time entitlement policies over time in Germany.
Source: TzBfG, BEEG, PflegeZG, FPfZG.

Table F.4. – General Entitlements to Part-Time Work, Germany

	Temporary part-time entitlement (*Brückenteilzeit*)	Permanent part-time entitlement (*zeitlich nicht begrenzte Teilzeitarbeit*)
Enactment	01.01.2019	01.01.2001
Length	Pre-specified, between one and five years.	Unspecified.
Requirements	- No justification needed, - Six month empl. in firm, - More than 45 employees	- No justification needed, - Six month empl. in firm, - More than 15 employees
Application	Three months in advance in written form, including desired work time and distribution of work time.	
Return to full-time	Automatic return to regular working hours after the pre-specified period. Within the pre-specified period, no entitlement to request further reductions or an extension of work time.	No general entitlement to request a return to regular working hours. However, after the request for an extension of hours is expressed, any new full-time position must give the already part-time employed the benefit in case of equal qualifications. This is also true for employees hired as part-time.
General Exceptions	The request can be declined by the employer in case of a significant impairment of the organization, the work process or safety in the company or the incurring of disproportionately high costs for the employer. The decline has to be issued in written form at least one month prior to the desired start of the work time reduction.	
Detailed Exceptions	For medium-sized firms with 46-200 employees, employers can decline the request if for each 15 employees, there is already one employee working in temporary part-time (e.g. for a firm with 140 employees, there need to be already 10 working in temp. part-time work.)	N/A
Repeated Application	Embargo period for two years (one year for declined request based on detailed exception)	Embargo period for two years
Consecutive Application	Embargo period for one year after return to regular hours	N/A

Note: The table describes regulations providing an entitlement to part-time work according to the TzBfG. In all regulations, employees exclude trainees, interns, or apprentices. Accompanying regulations include the explicit prohibition of discrimination, equal treatment in terms of pay and other employer benefits, prohibition of termination.
Source: Gesetz über Teilzeitarbeit und befristete Arbeitsverträge (TzBfG).

Table F.5. – Specific Entitlements to Part-Time Work, Germany

Type	Description	Regulatory framework
Partial retirement	Possibility for employees above age 55 to reduce their base working time by 50 percent. It requires employers to top-up the reduced earnings and social security contributions. Mutual agreement between employer and employee necessary.	Altersteilzeitgesetz (AltTZG)
Parental leave	Persons eligible for parental leave are allowed to work up to 32 hours per week (monthly average). There is an entitlement for two reductions of hours during parental leave to between 15 and 32 hours for at least two months for employees with (i) six month employment in the firm in (ii) firms with more than 15 employees. Similar exceptions as in Table F.4 apply.	Gesetz zum Elterngeld und zur Elternzeit (Bundeselterngeld- und Elternzeitgesetz - BEEG)
Home Care	Employees that provide home care for close relatives (at least care level 1) can request leave/part-time employment for up to 6 months. The entitlement only applies to individuals in firms with more than 15 employees.	Gesetz über die Pflegezeit (Pflegezeitgesetz - PflegeZG)
	Employees that provide home care for close relatives (at least care level 1) can request part-time employment under at least 15 hours per week for up to 24 months. The entitlement only applies to individuals in firms with more than 25 employees.	Gesetz über die Familienpflegezeit (Familienpflegezeitgesetz - FPfZG)

Note: The table provides an overview over different types of regulations concerning part-time work that go beyond the TzBfG in Table F.4.
Source: AltTZG, BEEG, PflegeZG, FPfZG.

F.2 SUPPLEMENTARY TABLES AND GRAPHS

Table F.6. – Summary Statistics, Mikrozensus (2020)

Variable	Share (%)
Region	
West Germany (incl. Berlin)	85.05
East Germany	14.95
Sex	
Male	49.54
Female	50.46
Age	
<15	13.77
15-24	10.36
25-54	39.63
55-64	15.16
>65	21.07
Completed Education	
In Education / Missing	13.97
Primary	4.57
Secondary I	12.97
Secondary II	37.09
Post Sec.	9.11
Bachelor	12.56
Master	8.30
Doctorate	1.43
Labor Force Status	
Employed	50.81
Unemployed	2.04
Job-seeking non-employed individuals	0.52
Other non-employed individuals	46.63
Civil Status	
Single	41.90
Married	45.23
Widowed	6.30
Divorced	6.57
Number of Individuals in Sample (Population):	469,678 (82,932,119)

Notes: This table presents summary statistics for key demographics. Data does not include individuals in communal accommodations.
Source: Own calculations based on Mikrozensus (1996-2020).

Table F.7. – Summary Statistics, BHP (2018)

Panel A

Variable	Full Sample				By Elig. (Mean)	
	Mean	P50	P10	P90	$E = 0$	$E = 1$
Age of Establishment	14.59	12.00	1.00	35.00	14.46	25.27
Number of Empl.	12.04	3.00	1.00	19.00	7.85	360.31
Part-Time Share	68.58	87.50	0.00	100.00	69.15	21.26
Share of Female Empl.	58.73	66.67	0.00	100.00	59.10	34.03
Daily Wage P50, FT (EUR)	88.47	81.10	49.05	137.54	87.83	115.48
Share of Empl., low qual.	13.65	.	.	.	13.48	12.47
Share of Empl., mid qual.	72.65	.	.	.	72.86	68.96
Share of Empl., high qual.	13.70	.	.	.	13.67	18.57
Share of Empl., age 15-29	16.43	.	.	.	16.15	19.88
Share of Empl., age 30-44	27.79	.	.	.	27.80	32.51
Share of Empl., age 45-59	38.70	.	.	.	38.81	39.09
Share of Empl., age above 60	17.08	.	.	.	17.23	8.52
Eligible	1.20	0.00	0.00	0.00	0.00	100.00

Panel B

Category	Full Sample	$E = 0$	$E = 1$
Industry ($N = 1,539,243$)			
Agriculture, Forestry, and Fishing	2.71	2.69	0.70
Energy and Water Supply, Mining	1.76	1.71	5.66
Manufacturing	3.18	2.94	23.17
Construction Industry	1.94	1.90	4.84
Trade	25.92	26.02	16.87
Transportation and Comunication	8.43	8.47	6.92
Banking and Insurance	14.00	14.10	6.08
Services (businesses and liberal professions)	7.20	7.13	12.67
Non-profit Organizations and Private Households	15.39	15.35	20.88
Public Authorities and Social Security	19.46	19.70	2.19
German State of Establishment ($N = 1,540,429$)			
Schleswig-Holstein	3.73	3.74	2.53
Hamburg	2.37	2.36	3.14
Lower Saxony	9.44	9.45	8.31
Bremen	0.75	0.74	1.09
North Rhine-Westphalia	21.25	21.23	21.81
Hesse	7.55	7.55	7.70
Rhineland-Palatinate	5.19	5.21	3.87
Baden-Wuerttemberg	13.58	13.58	13.28
Bavaria	17.56	17.58	15.24
Saarland	1.19	1.19	1.18
Berlin	4.06	4.06	4.36
Brandenburg	2.67	2.66	3.33
Mecklenburg-Vorpommern	1.85	1.85	1.85
Saxony	4.37	4.36	5.94
Saxony-Anhalt	2.19	2.18	3.20
Thuringia	2.27	2.26	3.16

Note: This table presents summary statistics for BHP data in 2018. Panel A describes distributional properties of variables for the full sample, and separately for establishments that are (not) eligible to the part-time policy, i.e. $E = 1$ ($E = 0$). The number of observations used in this panel is constant and equals $N = 1,540,438$. For the age and qualification distribution of employees, only mean estimates are available. Panel B displays the distribution of industry, and German state of establishment location, both for the full sample and separated by eligibility to the part-time employment policy in percent.
Source: Own calculations based on BHP (2018).

Table F.8. – Summary Statistics, SIAB (2018)

Panel A

Variable	Full Sample				By Elig. (Mean)	
	Mean	P50	P10	P90	$E = 0$	$E = 1$
Sex (women=1)	0.48	0.00	0.00	1.00	0.53	0.41
Age	45.40	47.00	30.00	59.00	45.49	45.26
Daily Gross Wage (EUR)	100.12	93.59	23.82	190.76	83.21	125.26
Experience	18.29	17.83	4.98	32.29	17.69	19.18
Part-Time Indicator	33.60	0.00	0.00	100.00	39.91	24.21
Switch ($PT \to FT$)	1.40	0.00	0.00	0.00	1.41	1.38
Switch ($FT \to PT$)	1.36	0.00	0.00	0.00	1.43	1.24
Switch ($FT \to PT$, in est.)	0.81	0.00	0.00	0.00	0.71	0.95
Switch ($PT \to FT$, in est.)	0.74	0.00	0.00	0.00	0.68	0.84
Total Employees	1020.26	90.00	6.00	1519.00	48.01	2465.09
Full-Time Employees	776.26	51.00	2.00	1073.00	28.57	1887.37
Part-Time Employees	244.00	20.00	2.00	446.00	19.43	577.72
Eligible	40.22	0.00	0.00	100.00	0.00	100.00

Panel B

Category	Full Sample	$E = 0$	$E = 1$
Education ($N = 393,014$)			
No schooling information, *Hauptschule*	6.47	6.61	6.28
Mittlere Reife, no voc. degree	2.53	2.72	2.28
Mittlere Reife, voc. degree	55.55	59.25	50.73
Fachhochschulreife/Abitur, no voc. degree	2.54	2.55	2.53
Fachhochschulreife/Abitur, voc. degree	12.99	12.98	13.00
University	19.92	15.88	25.17
Industry ($N = 538,114$)			
Agriculture, Forestry, and Fishing	0.96	1.32	0.43
Energy and Water Supply, Mining	3.89	3.49	4.48
Manufacturing	15.68	7.77	27.43
Construction Industry	3.70	3.04	4.68
Trade	21.21	28.63	10.19
Transportation and Comunication	7.41	8.82	5.31
Banking and Insurance	8.14	8.55	7.54
Services (businesses and liberal professions)	7.42	7.40	7.46
Non-profit Organizations and Private Households	27.61	25.72	30.42
Public Authorities and Social Security	3.97	5.26	2.06
German State of Residence ($N = 535,051$)			
Schleswig-Holstein	3.29	3.67	2.73
Hamburg	2.25	2.07	2.53
Lower Saxony	9.58	9.89	9.13
Bremen	0.73	0.69	0.79
North Rhine-Westphalia	20.89	20.55	21.40
Hesse	7.58	7.30	8.00
Rhineland-Palatinate	4.78	4.94	4.55
Baden-Wuerttemberg	13.88	13.38	14.62
Bavaria	16.71	16.39	17.18
Saarland	1.17	1.12	1.24
Berlin	3.86	3.55	4.32
Brandenburg	3.07	3.26	2.79
Mecklenburg-Vorpommern	1.85	2.17	1.37
Saxony	4.99	5.32	4.51
Saxony-Anhalt	2.68	2.81	2.49
Thuringia	2.68	2.91	2.35

Note: This table presents summary statistics for SIAB data in 2018. Panel A describes distributional properties of variables for the full sample, and separately for individuals working in firms that are (not) eligible to the part-time policy, i.e. $E = 1$ ($E = 0$). The number of observations used in this panel is constant and equals $N = 538,115$. Panel B displays the distribution of education, industry, and German state of location, both for the full sample and separated by eligibility to the part-time employment policy. Highlighted gray cells refer to estimates used in the welfare analysis in Section 6.5.4 (see also Table 6.6).
Source: Own calculations based on SIAB (2018).

Table F.9. – Hourly Wages and Working Hours, Summary Statistics (2014)

Panel A

Variable	Full Sample						Part-Time	
	Mean	P50	P10	P90	Min.	Max.	No	Yes
Weekly Work Hours	33.29	38.31	14.58	41.98	3.13	64.89	38.88	21.24
Part-Time (0,1)	0.30	0.00	0.00	1.00	0.00	1.00	0.00	1.00
Daily Gross Wage	88.97	82.23	14.80	178.05	0.03	325.27	100.81	37.70
Hourly Gross Wage	13.40	11.83	5.29	24.01	0.02	318.73	13.26	8.70
Experience (years)	15.81	15.21	2.46	30.42	0.00	39.52	15.51	14.66

Panel B

Category	Full Sample	No	Yes
Education ($N = 377,884$)			
No schooling information, *Hauptschule*	7.95	10.31	6.94
Mittlere Reife, no voc. degree	4.06	4.91	3.70
Mittlere Reife, voc. degree	56.55	56.28	56.66
Fachhochschulreife/Abitur, no voc. degree	4.54	6.84	3.54
Fachhochschulreife/Abitur, voc. degree	11.24	10.71	11.47
University	15.66	10.95	17.69
Industry ($N = 377,884$)			
Up to 5 years	18.78	18.41	19.63
More than 5 years, up to 10 years	15.34	15.44	15.12
More than 10 years, up to 15 years	15.17	14.58	16.55
More than 15 years, up to 20 years	14.90	13.86	17.31
More than 20 years, up to 25 years	17.15	17.87	15.49
More than 25 years, up to 30 years	8.11	8.43	7.37
More than 30 years, up to 35 years	5.61	6.09	4.51
More than 35 years, up to 40 years	4.93	5.32	4.03

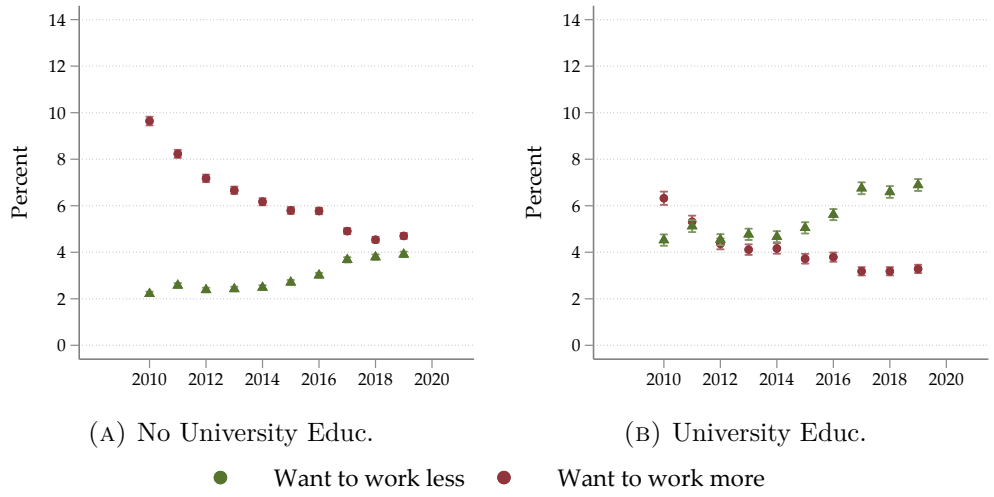
Note: This table presents summary statistics for the sample used in the analysis of the relationship between hourly wages and weekly working hours. Panel A describes distributional properties of variables. Panel B displays the distribution of education and experience. The number of observations lies at 377,884. Highlighted gray cells refer to estimates used in the welfare analysis in Section 6.5.4 (see also Table 6.6).
Source: Own calculations based on SIAB (2014).

Table F.10. – Hourly Wages and Working Hours, Heterogeneity

Variable	Category	δ (s.e.)	β (s.e.)	C.-D. F-Stat.	N_T (N_S)
Sex	Men	-0.72 (0.002)	0.32 (0.003)	512,586	963,491 (450,905)
	Women	-0.57 (0.001)	0.11 (0.003)	165,243	814,006 (291,673)
Year	2010-2012	-0.60 (0.001)	0.17 (0.003)	188,535	1,028,215 (430,900)
	2013-2014	-0.60 (0.002)	0.17 (0.003)	137,564	749,282 (311,678)
Education	Lower Education	-0.64 (0.001)	0.17 (0.002)	313,599	1,338,958 (418,888)
	Higher Education	-0.54 (0.002)	0.18 (0.004)	81,179	438,539 (194,214)

Note: This table displays for different sub-samples, how the hourly wage rate varies with working hours. The analysis is based on equations (6.3) and (6.4). The table displays three sample splits based on sex, year, and education. Sub-samples based on education compare the effect for individuals of lower education (up to university entrance qualification without a vocational degree) and higher education (university entrance qualification with vocational degree or some university education). All specifications include the following fully interacted controls: year (5 categories), experience (9 categories), occupation (339 categories), education (7 categories), industry (272 categories), and state (16 categories). For each specification, the table shows the first stage coefficient δ and standard error (column 3). Column 4 displays the second stage coefficient β and its standard error. Column 5 displays the Cragg-Donald Wald F-Statistic providing information about the strength of the instrument in the first stage. Column 6 displays the total number of observations N and the number of singleton observations N_S that are dropped in the respective specification.
Source: Own calculations based on SIAB (2010-2014).

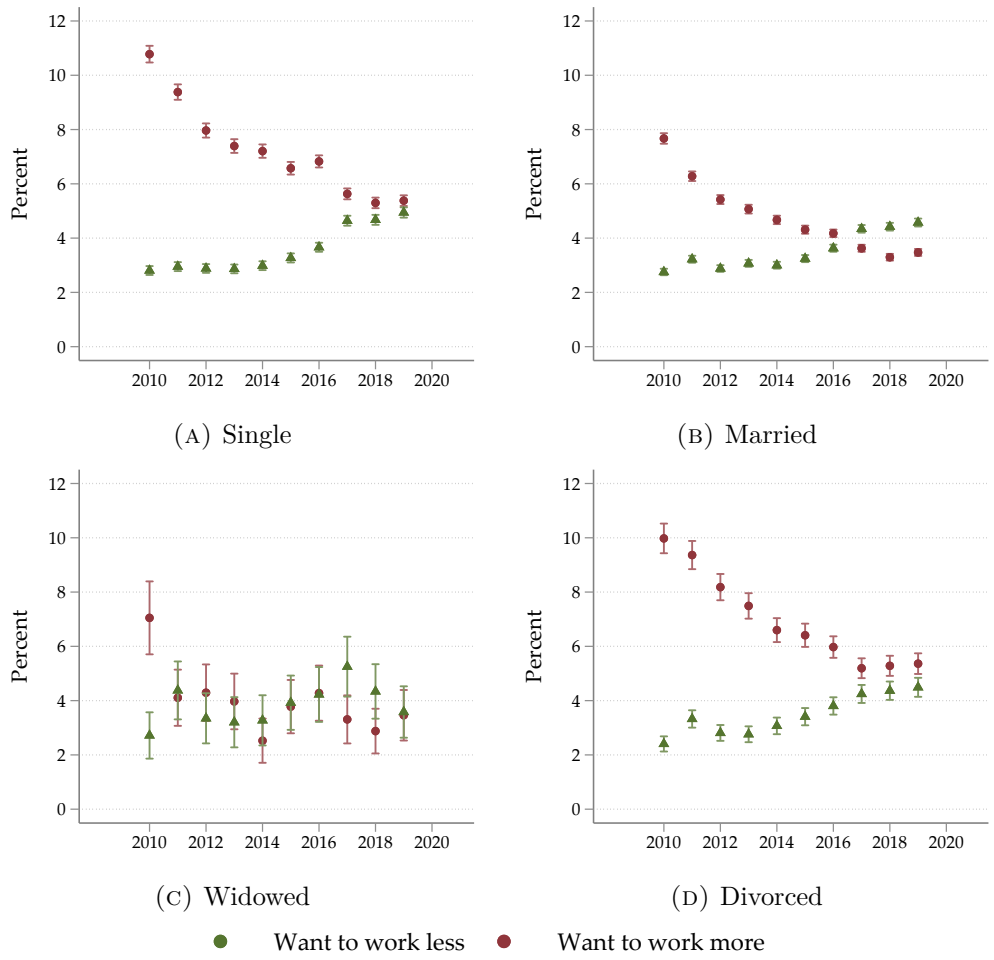
Figure F.1. – Desired Change in Working Hours, by Education



(A) No University Educ. (B) University Educ.

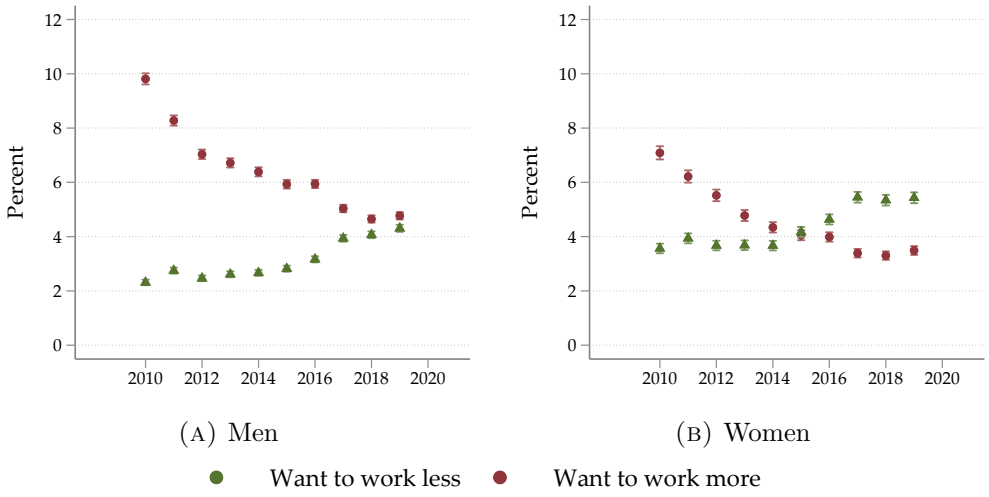
Note: This figure displays for each year between 2010 and 2019 the share of full-time employed individuals who state a desire to increase (decrease) their regular weekly working hours considering a correspondingly higher (lower) salary. The figure distinguishes between individuals of different education. The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

Figure F.2. – Desired Change in Working Hours, by Civil Status



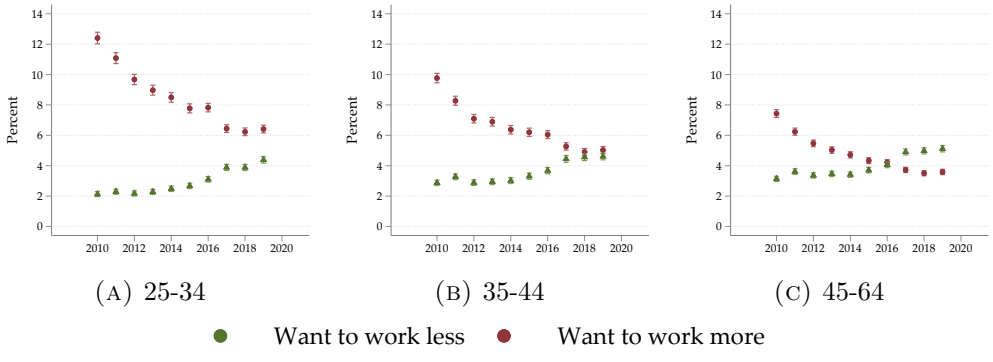
Note: This figure displays for each year between 2010 and 2019 the share of full-time employed individuals who state a desire to increase (decrease) their regular weekly working hours considering a correspondingly higher (lower) salary. The figure distinguishes between individuals of different education. The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

Figure F.3. – Desired Change in Working Hours, by Sex



Note: This figure displays for each year between 2010 and 2019 the share of full-time employed individuals who state a desire to increase (decrease) their regular weekly working hours considering a correspondingly higher (lower) salary. The figure distinguishes between different civil statuses. The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

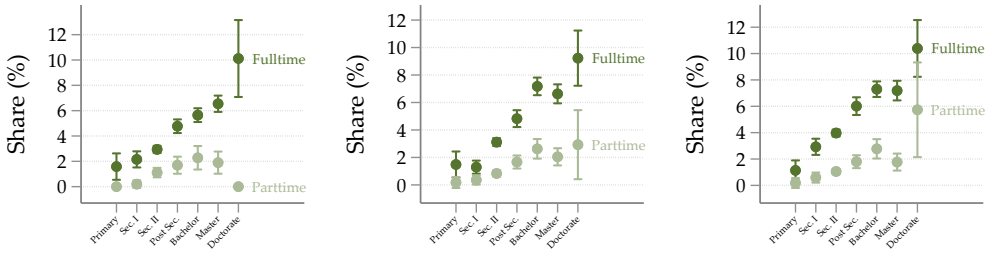
Figure F.4. – Desired Change in Working Hours, by Age



Note: This figure displays for each year between 2010 and 2019 the share of full-time employed individuals who state a desire to increase (decrease) their regular weekly working hours considering a correspondingly higher (lower) salary. The figure distinguishes between individuals of different ages. The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

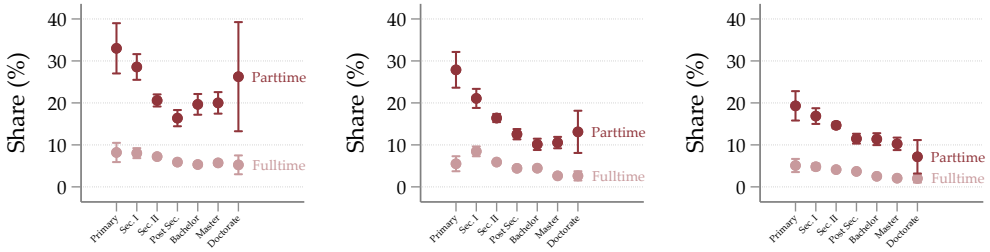
Figure F.5. – Desire for Work-Life Balance across Education, by Age

Desire to Work Less



(A) 25-34 (B) 35-44 (C) 45-64

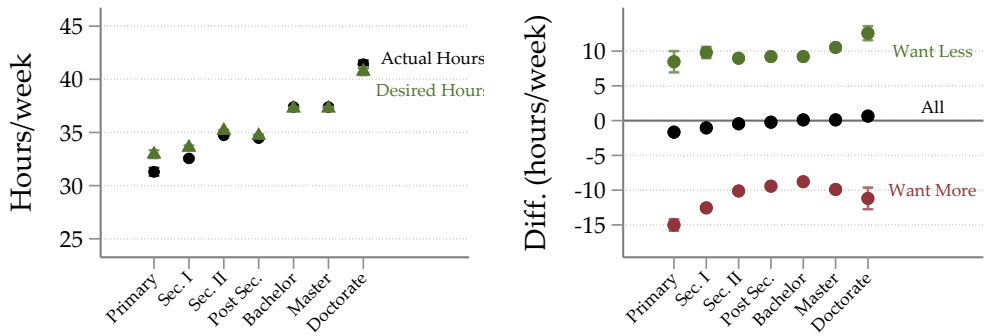
Desire to Work More



(D) 25-34 (E) 35-44 (F) 45-64

Note: This figure displays for 2019 and different age groups, how the share of individuals who state a desire to increase (decrease) their regular weekly working hours considering a correspondingly higher (lower) salary varies with ability. The figure shows point estimates and associated 95 percent confidence intervals. The upper (lower) panel shows the desire for lower (higher) working hours across different levels of education. Education information has been harmonized based on the International Standard Classification of Education (ISCED-2011). The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

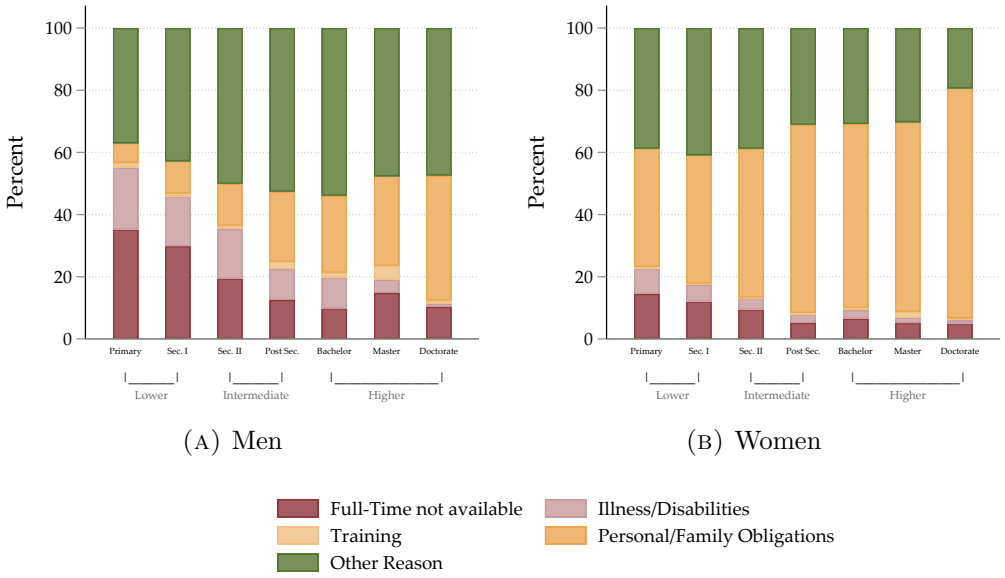
Figure F.6. – Desired Working Hours across Education (2019)



(A) Actual and Desired Hours (B) Difference

Note: This figure visualizes information on actual and desired weekly working hours. Figure F.6a displays the average hours per week, while Figure F.6b visualizes the difference between actual and desired working time, both in the aggregate, and separately for those individuals that state a desire to change their working time in a specific direction. The figure shows point estimates and associated 95 percent confidence intervals. All analyses focus on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

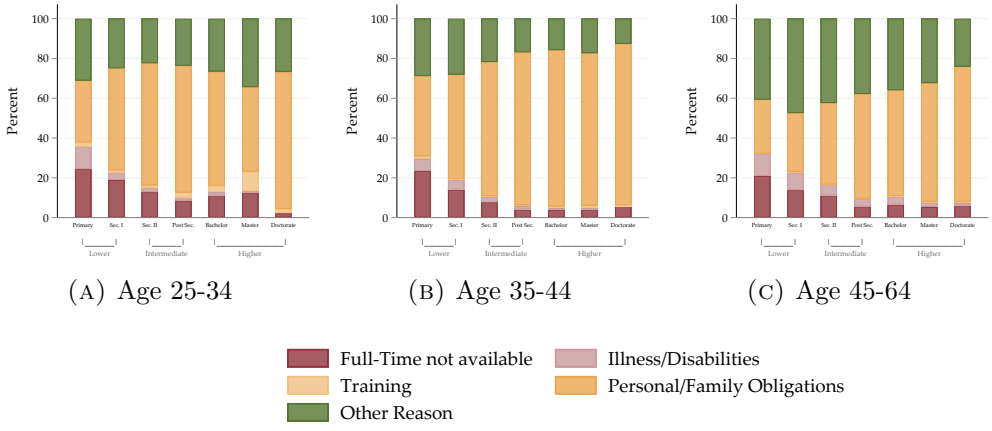
Figure F.7. – Reason for Part-Time Work, by Sex



(A) Men (B) Women

Note: This figure displays how the distribution of reasons for part-time work varies across education for those individuals that are employed part-time in 2019. Figure F.7a displays the detailed breakdown of stated reasons for part-time employment for men. Figure F.7b repeats the exercise for women. The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

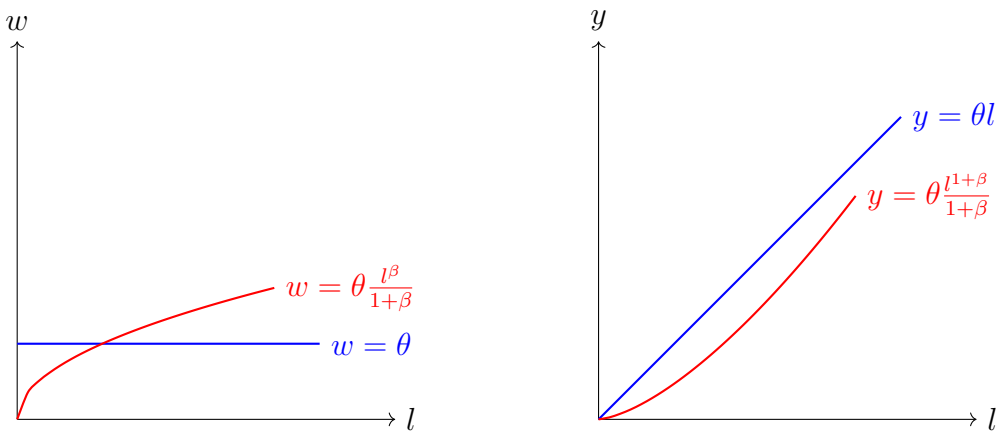
Figure F.8. – Reason for Part-Time Work, by Age



(A) Age 25-34 (B) Age 35-44 (C) Age 45-64

Note: This figure displays how the distribution of reasons for part-time work varies across education for those individuals that are employed part-time in 2019. The figure distinguishes between individuals of different age. The analysis focuses on employed individuals with positive earnings that are not in education and are between 25 and 64 years old.
Source: Own calculations based on Mikrozensus (1996-2020).

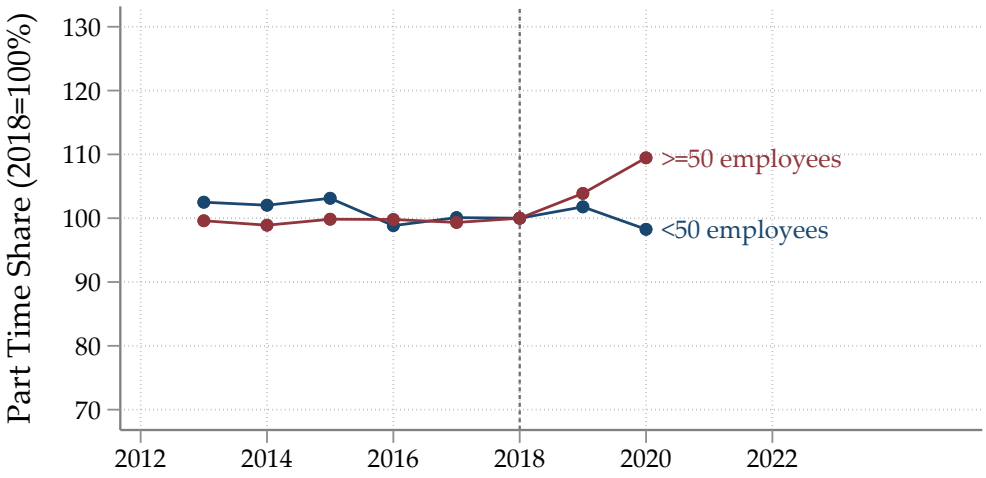
Figure F.9. – Convexity of Earnings, Illustration



(A) Hourly Wage (B) Earnings

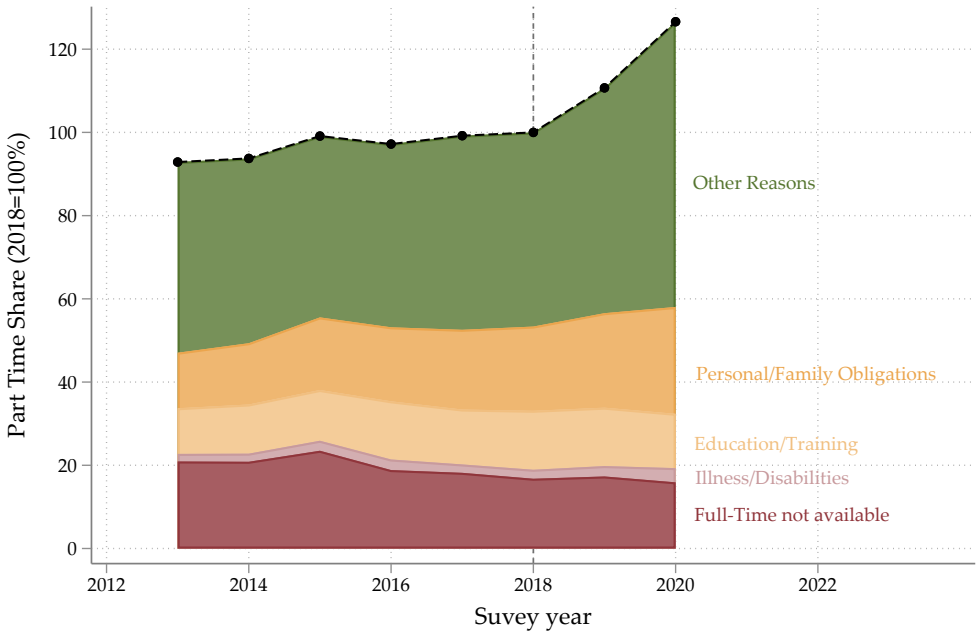
Note: This figure illustrates the difference between a linear and a convex earnings schedule as blue (red) lines. Figure F.9a displays how the hourly wage varies with working hours while Figure F.9b shows how earnings vary with working hours. In the illustration, $\beta = 0.5$ and $\theta = 1$.

Figure F.10. – Work-Life Balance Policy and Part-Time Employment, Microcensus

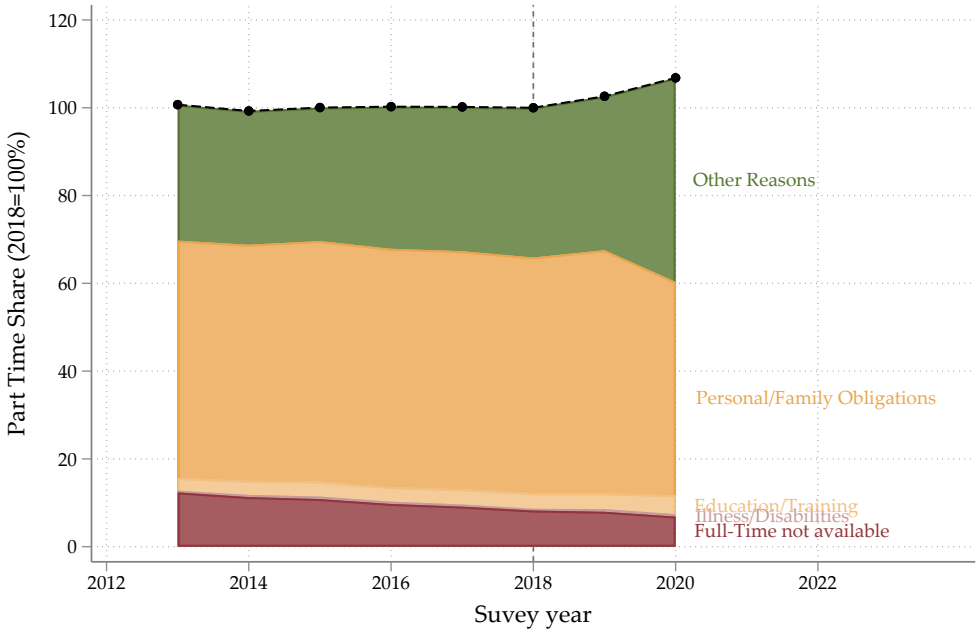


Note: This figure displays how the share of part-time employees evolved for individuals that work in establishments of different employment size. Since Microcensus data only contains coarse categories of the number of employees in the local establishment, the figure sets the closest possible cut to the threshold relevant for the policy (see Figure 6.9) and distinguishes between individuals working in establishments with less than 50 employees, and up to 50 employees. The part-time share among both groups is normalized to 100 percent in 2018, the year before the policy came into effect.
Source: Own calculations based on Mikrozensus (1996-2020).

Figure F.11. – Work-Life Balance Policy and the Nature of Part-Time Employment, by Sex



(A) Men



(B) Women

Note: This figure displays the normalized part-time share for individuals that are employed part-time in establishments with 50 employees and more, separately for men and women. Since Microcensus data only contains coarse categories of the number of employees in the local establishment, this employment size cut-off is based on the closest possible cut to the threshold relevant for the policy (see Figure 6.9). The part-time share is normalized to 100 percent in 2018, the year before the policy came into effect. The figure further shows the distribution of stated reasons of part-time employment. The analysis focuses on part-time employed individuals with positive earnings that are not in education and are between 25 and 64 years old. Civil servants are excluded.
Source: Own calculations based on Mikrozensus (1996-2020).

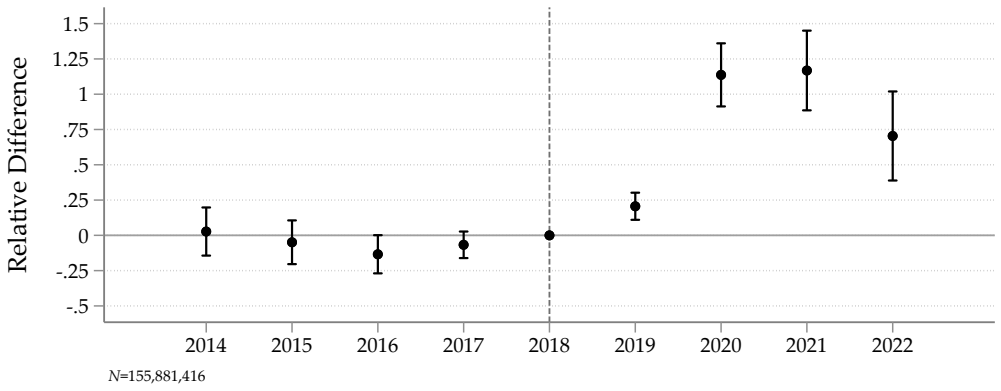
Effect of WLB Policy, Establishment-Level Specification

The analysis in Section 6.4 uses employees as the unit of observation, since reactions to the policy occur at this level. A disadvantage of the individual-level analysis is that the data structure does not allow to control for time-constant unobserved heterogeneity across establishments. For this purpose, however, one can turn to the BHP data and estimate the effect of the policy on the part-time share $Y_{e,t}$ in establishment e at time t based on the regression equation

$$Y_{et} = \sum_{t \neq 2018} \beta^t d_t \times E_{e,t} + \mu_t + \gamma_e + \epsilon_{et} \tag{F.1}$$

where γ_e are establishment fixed effects that control for time-invariant unobserved characteristics of establishments and μ_t are time fixed effects that control for time specific effects that affect all establishments equally. $\epsilon_{e,t}$ is the idiosyncratic error term. To make estimates comparable to the individual-level analysis, the regression is weighted by the total pre-reform employment size of establishments. This allows us to obtain an individual-level interpretation of the effects while using establishment level outcomes. Figure F.12 shows the results for the employment-weighted part-time share at the establishment level.

Figure F.12. – Effect of Work-Life Balance Policy, Establishment-Level



Note: This figure shows the dynamic difference-in-difference estimation of the policy effect based on equation (F.1). Treatment assignment is based on the establishment size and the part-time share of the establishment according to the relevant thresholds displayed in Figure 6.9. The figure displays results using a switching indicator from part-time to full-time (0,100) as outcome variable of interest and distinguishes between estimates using all switches (black) and estimates using only within-establishment switches (grey). All analyses are weighted by the pre-reform (2018) employment size to simulate an individual level analyses using establishment data. Standard errors are clustered at the level of establishments.
Source: Own calculations based on BHP (2014-2022).

F.3 SUPPLEMENTARY DERIVATIONS

Individual Optimization, Second Order Conditions

In Section 6.5.3, I present first order conditions for the individual utility maximization problem. To show that the critical point defined by these first order conditions is indeed a local maximum, I show below that the second order conditions based on the Hessian matrix. Remember that the first partial derivatives of the utility function V from equation (6.11) read

$$\frac{\partial V}{\partial e_F} = u_F - u_U - \psi'(e_F)\,, \tag{F.2}$$

$$\frac{\partial V}{\partial e_P} = u_P - u_U - \psi'(e_P)\,. \tag{F.3}$$

The corresponding Hessian matrix collecting the second partial derivatives of the utility function read

$$H = \begin{pmatrix} V_{e_F e_F} & V_{e_F e_P} \\ V_{e_P e_F} & V_{e_P e_P} \end{pmatrix} = \begin{pmatrix} -\psi''(e_F) & 0 \\ 0 & -\psi''(e_P) \end{pmatrix}. \tag{F.4}$$

The Hessian matrix is negative definite, since the principal minors are alternating in sign and the first principal minor is negative, i.e.,

$$\text{First principal minor:} \quad \Delta_1 \det\left[V_{e_F e_F}\right] = -\psi''(e_F) < 0$$

$$\text{Second principal minor:} \quad \Delta_2 = \det\begin{bmatrix} V_{e_F e_F} & V_{e_F e_P} \\ V_{e_P e_F} & V_{e_P e_P} \end{bmatrix} = \psi''(e_F)\psi''(e_P) > 0$$

Given that the Hessian matrix is negative definite, the critical point defined via the first order conditions is indeed a local maximum.

Individual Optimization — Comparative Statics

In Section 6.5.3, I define elasticities that are based on comparative statics at the individual optimum. In particular, the total differential of the first order conditions (6.12)-(6.13) yield the following partial derivatives that identify reactions of individual choices to changes in the policy parameters:

$$\begin{aligned}
\frac{\partial e_F}{\partial b_U} &= \frac{-u'_U}{\psi''_F} < 0 & \qquad \frac{\partial e_P}{\partial b_U} &= \frac{-u'_U}{\psi''_P} < 0 \\
\frac{\partial e_F}{\partial b_P} &= \frac{-u'_P}{\psi''_F} < 0 & \qquad \frac{\partial e_P}{\partial b_P} &= \frac{u'_P}{\psi''_P} > 0 \\
\frac{\partial e_F}{\partial \tau_U} &= \frac{-u'_F}{\psi''_F} < 0 & \qquad \frac{\partial e_P}{\partial \tau_U} &= \frac{-u'_P}{\psi''_P} < 0 \\
\frac{\partial e_F}{\partial \tau_P} &= \frac{-u'_F}{\psi''_F} < 0 & \qquad \frac{\partial e_P}{\partial \tau_P} &= \frac{u'_F}{\psi''_P} > 0
\end{aligned} \tag{F.5}$$

The sign of these partial derivatives define the sign of elasticities defined in equations (6.14)-(6.17).

Fiscal Effects of Benefit Changes

In Section 6.5.3, I show in equations (6.18) and (6.19) how changes in the benefit levels b_P and b_U affect the resources of the government by implicit changes in the tax rates τ_P and τ_U. In the following, I derive these implicit changes based on the total differential of the government constraints, and taking endogenous choices of individuals into account.

Fiscal Effects of b_P on τ_P. The effect of an isolated increase in b_P on τ_P can be derived by using the total differential on the government constraint imposed by the actuarially fair part-time employment insurance given by

$$I_2 = e_F \tau_P - e_P b_P = 0. \tag{F.6}$$

Looking at the total differential

$$\begin{aligned} dI_2 &= \frac{\partial I_2}{\partial e_F} de_F + \frac{\partial I_2}{\partial e_P} de_P + \frac{\partial I_2}{\partial b_P} db_P + \frac{\partial I_2}{\partial \tau_P} d\tau_P = 0 \\ &= \tau_P de_F - b_P de_P + e_F d\tau_P - e_P db_P \end{aligned} \tag{F.7}$$

we obtain the required change in τ_P taking the endogenous changes in choice variables e_F and e_P into account:

$$\begin{aligned} \frac{d\tau_P}{db_P} &= -\frac{\tau_P}{e_F}\frac{de_F}{db_P} + \frac{b_P}{e_F}\frac{de_P}{db_P} + \frac{e_P}{e_F} \\ &= \frac{e_P}{e_F}\left(1 + \mathcal{E}_{e_P,b_P} - \mathcal{E}_{e_F,b_P}\right) \\ &= \frac{e_P}{e_F}\left(1 + \frac{e_T}{e_F}\mathcal{E}_{\frac{e_P}{e_T},b_P}\right) \end{aligned} \tag{F.8}$$

Fiscal Effects of b_P on τ_U. Even absent any changes in b_U, a change in b_P requires the government to make changes in τ_U to keep the insurance constraints satisfied. Intuitively, when increasing b_P, there is a spill-over to the insurance provided for unemployed individuals that are driven by changes in full-time and part-time employment. We can assess the required change by looking at the total differential of the combined insurance constraint given by

$$I_3 = I_1 + I_2 = (e_F + e_P)\,\tau_U - (1 - e_F - e_P)\,b_U + e_F\tau_P - e_P b_P = 0 \tag{F.9}$$

to derive the required change via the total differential as

$$\begin{aligned} \frac{d\tau_U}{db_P} &= -\left(\frac{b_U}{e_T^2}\right)\frac{de_T}{db_P} - \frac{b_P}{e_T}\left(\frac{e_P}{e_F}\frac{de_F}{db_P} - \frac{de_P}{db_P}\right) + \frac{(1-e_T)}{e_T}\frac{db_U}{db_P} + \frac{e_P}{e_T} - \frac{e_F}{e_T}\frac{d\tau_P}{db_P} \\ &= -\left(\frac{b_U}{e_T^2}\right)\frac{de_T}{db_P} - \frac{b_P}{e_T}\left(\frac{e_P}{e_F}\frac{de_F}{db_P} - \frac{de_P}{db_P}\right) + \frac{(1-e_T)}{e_T}\frac{db_U}{db_P} + \frac{e_P}{e_T} - \frac{e_F}{e_T}\frac{e_P}{e_F} - \frac{e_F}{e_T}\frac{e_P}{e_F}\frac{e_T}{e_F}\mathcal{E}_{\frac{e_P}{e_T},b_P} \\ &= -\left(\frac{b_U}{e_T^2}\right)\frac{de_T}{db_P} + \frac{e_P}{e_T}\left(\mathcal{E}_{\frac{e_P}{e_T},b_P}\frac{e_T}{e_F}\right) + \frac{(1-e_T)}{e_T}\frac{db_U}{db_P} - \frac{e_P}{e_F}\mathcal{E}_{\frac{e_P}{e_T},b_P} \\ &= -\left(\frac{b_U}{e_T^2}\right)\frac{de_T}{db_P} + \frac{(1-e_T)}{e_T}\frac{db_U}{db_P} \\ &= \left(\frac{b_U(1-e_T)}{b_P e_T^2}\right)\frac{d(1-e_T)}{db_P}\frac{b_P}{1-e_T} + \frac{(1-e_T)}{e_T}\frac{db_U}{db_P} \\ &= \frac{b_U}{b_P}\frac{1-e_T}{e_T}\frac{\mathcal{E}_{1-e_T,b_U}}{e_T} + \frac{(1-e_T)}{e_T}\frac{db_U}{db_P}. \end{aligned} \tag{F.10}$$

Since we look at an isolated change in b_P and do not modify b_U, the equation simplifies to

$$\frac{d\tau_U}{db_P} = \frac{b_U}{b_P}\frac{1-e_T}{e_T}\frac{\mathcal{E}_{1-e_T,b_U}}{e_T}. \tag{F.11}$$

Fiscal Effects of b_U on τ_U. The effect of an isolated increase in b_U on τ_U can be derived by using the total differential on the government constraint imposed by the actuarially fair unemployment insurance given by

$$I_1 = (e_F + e_P)\,\tau_U - (1 - e_F - e_P)\,b_U = 0. \tag{F.12}$$

Looking at the total differential

$$\begin{aligned} dI_1 &= \frac{\partial I_1}{\partial e_F}de_F + \frac{\partial I_1}{\partial e_P}de_P + \frac{\partial I_1}{\partial b_U}db_U + \frac{\partial I_1}{\partial \tau_U}d\tau_U = 0 \\ &= (\tau_U + b_U)\,(de_F + de_P) - (1 - e_F - e_P)db_U + (e_F + e_F)d\tau_U = 0 \end{aligned} \tag{F.13}$$

we obtain the required change in τ_U taking the changes in choice variables into account:

$$\begin{aligned} \frac{d\tau_U}{db_U} &= \frac{1-e_T}{e_T} - \frac{\tau_U + b_U}{e_T}\frac{de_T}{db_U} \\ &= \frac{1-e_T}{e_T} - \frac{\frac{1-e_T}{e_T}b_U + b_U}{e_T}\frac{de_T}{db_U} \\ &= \frac{1-e_T}{e_T} - \frac{b_U}{e_T^2}\frac{de_T}{db_U} \\ &= \frac{1-e_T}{e_T} + \frac{1-e_T}{e_T^2}\frac{d(1-e_T)}{db_U}\frac{b_U}{1-e_T} \\ &= \frac{1-e_T}{e_T}\left(1 + \frac{\mathcal{E}_{1-e_T,b_U}}{e_T}\right) \end{aligned} \tag{F.14}$$

Fiscal Effects of b_U on τ_P. Even absent any changes in b_P, a change in b_U requires the government to make changes in τ_P to keep the insurance constraints satisfied. Intuitively, when increasing b_U, there is a spill-over to the insurance provided for part-time employed

individuals that are driven by changes in full-time and part-time employment. We can assess the required change by looking at the total differential of the combined insurance constraint, i.e. equation (F.9). Looking at the total differential

$$\begin{aligned}
dI_3 &= \left(\frac{\partial I_1}{\partial e_F} + \frac{\partial I_2}{\partial e_F}\right) de_F + \left(\frac{\partial I_1}{\partial e_P} + \frac{\partial I_2}{\partial e_P}\right) de_P + \frac{\partial I_1}{\partial b_U} db_U + \frac{\partial I_1}{\partial \tau U} d\tau_U + \frac{\partial I_2}{\partial b_P} db_P + \frac{\partial I_2}{\partial \tau_P} d\tau_P = 0 \\
&= (\tau_U + b_U + \tau_P)\, de_F + (\tau_U + b_U - b_P)\, de_P - (1 - e_F - e_P) db_U + (e_F + e_P) d\tau_U - e_P db_P + e_F d\tau_P = 0 \\
&= (\tau_U + b_U)\, de_T + \tau_P de_F - b_P de_P - (1 - e_T) db_U + e_T d\tau_U - e_P db_P + e_F d\tau_P = 0 \\
&= \left(\frac{1 - e_T}{e_T} b_U + b_U\right) de_T + \tau_P de_F - b_P de_P - (1 - e_T) db_U + e_T d\tau_U - e_P db_P + e_F d\tau_P = 0 \\
&= \left(\frac{b_U}{e_T}\right) de_T + b_P \frac{e_P}{e_F} de_F - b_P de_P - (1 - e_T) db_U + e_T d\tau_U - e_P db_P + e_F d\tau_P = 0 \\
&= \left(\frac{b_U}{e_T}\right) de_T + b_P \left(\frac{e_P}{e_F} de_F - de_P\right) - (1 - e_T) db_U + e_T d\tau_U - e_P db_P + e_F d\tau_P = 0
\end{aligned} \tag{F.15}$$

we obtain the required change in τ_P taking the endogenous changes in choice variables e_F and e_P into account, i.e.,

$$\begin{aligned}
\frac{d\tau_P}{db_U} &= -\frac{b_U}{e_T e_F} \frac{de_T}{db_U} - \frac{b_P}{e_F} \left(\frac{e_P}{e_F} \frac{de_F}{db_U} - \frac{de_P}{db_U}\right) + \frac{1 - e_T}{e_F} + \frac{e_P}{e_F} \frac{db_P}{db_U} - \frac{d\tau_U}{db_U} \frac{e_T}{e_F} \\
&= -\frac{1}{e_F} \frac{de_T}{db_U} \frac{b_U}{e_T} - \frac{b_P}{e_F} \left(\frac{e_P}{b_U} \frac{de_F}{db_U} \frac{b_U}{e_F} - \frac{e_P}{b_U} \frac{de_P}{db_U} \frac{e_P}{b_U}\right) + \frac{1 - e_T}{e_F} + \frac{e_P}{e_F} \frac{db_P}{db_U} - \frac{d\tau_U}{db_U} \frac{e_T}{e_F} \\
&= \frac{d(1 - e_T)}{db_U} \frac{b_U}{1 - e_T} \frac{1 - e_T}{e_T e_F} + \frac{b_P}{b_U} \frac{e_P}{e_F} \left(\mathcal{E}_{e_P, b_U} - \mathcal{E}_{e_F, b_U}\right) + \frac{1 - e_T}{e_F} + \frac{e_P}{e_F} \frac{db_P}{db_U} - \frac{e_T}{e_F} \frac{1 - e_T}{e_T} \left(1 + \frac{\mathcal{E}_{1 - e_T, b_U}}{e_T}\right) \\
&= \mathcal{E}_{1 - e_T, b_U} \frac{1 - e_T}{e_T e_F} + \frac{b_P}{b_U} \frac{e_P}{e_F} \left(\mathcal{E}_{e_P, b_U} - \mathcal{E}_{e_F, b_U}\right) + \frac{1 - e_T}{e_F} + \frac{e_P}{e_F} \frac{db_P}{db_U} - \frac{1 - e_T}{e_F} + \frac{1 - e_T}{e_T e_F} \mathcal{E}_{1 - e_T, b_U} \\
&= \frac{b_P}{b_U} \frac{e_P}{e_F} \left(\mathcal{E}_{e_P, b_U} - \mathcal{E}_{e_F, b_U}\right) + \frac{e_P}{e_F} \frac{db_P}{db_U} \\
&= \frac{b_P}{b_U} \frac{e_P}{e_F} \frac{e_T}{e_F} \mathcal{E}_{\frac{e_P}{e_T}, b_U} + \frac{e_P}{e_F} \frac{db_P}{db_U}.
\end{aligned} \tag{F.16}$$

Since we look at an isolated change in b_U and do not modify b_P, the equation simplifies to

$$\frac{d\tau_P}{db_U} = \frac{b_P}{b_U} \frac{e_P}{e_F} \frac{e_T}{e_F} \mathcal{E}_{\frac{e_P}{e_T}, b_U}. \tag{F.17}$$

Combination of Marginal Welfare Effects

In Section 6.5.3, equation (6.24) presents sufficient statistics for a local welfare assessment of an increase in the part-time employment benefit under the assumption that the unemployment benefit is set optimally. To derive this equation, I evaluate equation (6.22) at zero, since any locally optimal unemployment benefit level must satisfy

$$\frac{\partial W}{\partial b_U} = (1-e_T)\left\{u'_U - \left(1+\frac{\mathcal{E}_{1-e_T,b_U}}{e_T}\right)\overline{u'_T} - \frac{e_T}{1-e_T}\,\frac{e_P}{e_F}\,\frac{b_P}{b_U}\,u'_F\,\mathcal{E}_{\frac{e_P}{e_T},b_U}\right\} = 0$$

which can be reformulated as

$$\frac{\overline{u'_T}}{u'_F}\frac{(1-e_T)}{e_T}\frac{b_U}{b_P}\frac{1}{e_P} = \frac{\frac{\mathcal{E}_{\frac{e_P}{e_T},b_U}}{e_F}}{\frac{(u'_U-\overline{u'_T})}{\overline{u'_T}} - \frac{\mathcal{E}_{1-e_T,b_U}}{e_T}}.$$

Plugging this equation into equation (6.21) yields

$$M_W(b_P|b_U^*) = \underbrace{\frac{(u'_P-u'_F)}{u'_F}}_{\text{Insurance Value}} - \underbrace{\frac{e_T}{e_F}\mathcal{E}_{\frac{e_P}{e_T},b_P}}_{\text{Direct Fiscal Costs}} \underbrace{- \frac{e_T}{e_F}\left(\frac{\mathcal{E}_{\frac{e_P}{e_T},b_U}}{\frac{e_T(u'_U-\overline{u'_T})}{\overline{u'_T}} - \mathcal{E}_{1-e_T,b_U}}\right)\mathcal{E}_{1-e_T,b_P}}_{\text{Spillover to Unemployment Insurance}}. \tag{F.18}$$

Elasticity Estimate

A key ingredient to the fiscal costs of the policy in Section 6.5.4 is the elasticity of the part-time share with respect to changes in the benefit level, i.e., $\mathcal{E}_{\frac{e_P}{e_T},b_P}$. In Section 6.4, I estimate how the policy expansion changed the part-time share, i.e. $d(\frac{e_P}{e_T})\frac{e_T}{e_P} \approx \frac{0.0044}{0.34} \approx 0.013$. The second part of the elasticity, i.e., the change in the benefit level can be assessed by analyzing, how the elimination of the hourly wage penalty (reduction in β to zero) changed the benefit associated with part-time employment b_P. This benefit is not directly observable, but is a reduced-form measure of the difference between taxes under full-time and part-time employment before the reform, i.e.,

$$\begin{aligned} b_P &= T(y(l_F)) - T(y(l_P)) \\ &= \left[\theta l_F - (\theta l_F)^{1-\rho}\right] - \left[\theta l_P - \left(\theta l_F k^{1+\beta}\right)^{1-\rho}\right] \end{aligned} \tag{F.19}$$

A change in β from β_1 to $\beta_2 = 0$ thus has the following relative effect on the benefit level:

$$\begin{aligned} \frac{db_P}{b_P} &= \frac{b_{P,2} - b_{P,1}}{b_{P,1}} \\ &= \frac{k^{(1+\beta_2)(1-\rho)} - k^{(1+\beta_1)(1-\rho)}}{(\theta l_F)^{\rho}(1-k) - \left[1 - k^{(1+\beta_1)(1-\rho)}\right]} \\ &= \frac{k^{(1-\rho)} - k^{(1+\beta_1)(1-\rho)}}{(\theta l_F)^{\rho}(1-k) - \left[1 - k^{(1+\beta_1)(1-\rho)}\right]} \end{aligned} \tag{F.20}$$

Based on the baseline parameter values in Table 6.6, we obtain that the reduction in β is an implicit relative reduction of the benefit level corresponding to $\frac{db_P}{b_P} \approx 0.053$. Equipped with both the change in the part-time share and the change in the benefit level, the baseline elasticity estimate reads

$$\mathcal{E}_{\frac{e_P}{e_T}, b_P} = \frac{\frac{d\left(\frac{e_P}{e_T}\right)}{\frac{e_P}{e_T}}}{\frac{db_P}{b_P}} \approx \frac{0.013}{0.053} \approx 0.24 \tag{F.21}$$

Sensitivity of Elasticity. To assess, how sensitive the elasticity estimate is, I use variation in the empirical estimates of percentage point changes in the part-time share $d\frac{e_P}{e_T}$ from Section 6.4, and also vary the baseline part-time share that serves as the basis for estimating the relative effect of the policy change. The variation of the elasticity estimate inform the discussions about how robust the estimated welfare effect is (see, e.g., Figure 6.14).

Marginal Value of Public Funds

Section 6.5.4 shows that the 2019 work-life balance policy reform was locally welfare-improving according to a comparison of the marginal benefits and costs of the policy. It

Table F.11. – Elasticity Estimate, Sensitivity

Policy Effect (p.p.)	Baseline Part-Time Share		
	Eligible 24.21	Average 33.60	Non-Eligible 39.91
CI% Lower Bound (0.32)	0.25	0.18	0.15
Estimate (0.44)	0.34	**0.24**	0.21
95% CI Upper Bound (0.62)	0.43	0.31	0.26

Note: This table shows, how sensitive the elasticity is to variations in the pre-reform part-time share (horizontal variation), and the measured policy effect (vertical variation). In the baseline, I take the average population-wide part-time share. However, it is also conceivable to use the part-time share among the population that is eligible or non-eligible for the policy. For the variations in the policy effect, I focus on the 95% confidence intervals of the baseline estimate. The baseline elasticity choice is highlighted in gray. Highlighted gray cells refer to estimates used in the welfare analysis in Section 6.5.4 (see also Table 6.6).
Source: Own calculations based on estimates in Table 6.4 and Appendix Table F.8.

is also possible to express the marginal impact of the reform in terms of the marginal value of public funds (MVPF) that relates the willingness to pay for a one-dollar increase in the part-time insurance policy to the net costs of the government. Following the analysis of Hendren and Sprung-Keyser (2020) for unemployment insurance, the MVPF of the policy reads

$$\begin{aligned} MVPF &= \frac{\text{Willingness to Pay}}{\text{Net Cost}} \\ &= \frac{1+\gamma\frac{c_F-c_P}{c_F}}{1+\text{Fiscal Externality}} = \frac{1+\left(1-k^{(1+\beta)(1-\rho)}\right)}{1+\frac{1}{1-s_{PT}}\mathcal{E}_{s_{PT},b_P}} = \frac{1.83}{1.37} = 1.34\,, \end{aligned} \tag{F.22}$$

where the willingness to pay reflects the value of consumption smoothing through the policy. Intuitively, before the risk has materialized, individuals are willing to pay 1.83 EUR for each additional Euro in the part-time employment state. From the perspective of the government, increasing the benefit in the part-time state by one additional Euro comes with costs of 1.37 EUR. This implies a MVPF of 1.34 which is slightly higher than the MVPF for unemployment insurance policies that have been estimated to range from 0.43 to 1.03 (Hendren and Sprung-Keyser, 2020). However, it is important to note that the estimated MVPF is not directly comparable to those of other insurance programs like unemployment insurance, since the consumption smoothing value is not estimated based on consumption data, but focuses on the implicit reduction in earnings.

Bibliography

Acemoglu, D., M. Golosov, and A. Tsyvinski (2008). "Political Economy of Mechanisms." *Econometrica* 76 (3), pp. 619–641.

Acemoglu, D., M. Golosov, and A. Tsyvinski (2010). "Dynamic Mirrlees Taxation under Political Economy Constraints." *Review of Economic Studies* 77 (3), pp. 841–881.

Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2015). "Democracy, Redistribution, and Inequality." *Handbook of Income Distribution*. Vol. 2. Elsevier, pp. 1885–1966.

Adriaans, J. (2023). "Fairness of Earnings in Europe: The Consequences of Unfair under- and Overreward for Life Satisfaction." *European Sociological Review* 39 (1), pp. 118–131.

Ales, L. and C. Sleet (2022). "Optimal Taxation of Income-Generating Choice." *Econometrica* 90 (5), pp. 2397–2436.

Alesina, A., A. Ichino, and L. Karabarbounis (2011). "Gender-Based Taxation and the Division of Family Chores." *American Economic Journal: Economic Policy* 3 (2), pp. 1–40.

Alesina, A., A. Miano, and S. Stantcheva (2023). "Immigration and Redistribution." *Review of Economic Studies* 90 (1), pp. 1–39.

Alesina, A., S. Stantcheva, and E. Teso (2018). "Intergenerational Mobility and Preferences for Redistribution." *American Economic Review* 108 (2), pp. 521–554.

Almås, I., A. W. Cappelen, E. Ø. Sørensen, and B. Tungodden (2024a). *Fairness Across the World: Preferences and Beliefs*. Mimeo.

Almås, I., A. W. Cappelen, and B. Tungodden (2020). "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy* 128 (5), pp. 1753–1788.

Almås, I., P. Hufe, and D. Weishaar (2023). "Equality of Opportunity: Fairness Preferences and Beliefs About Inequality." *Handbook of Equality of Opportunity*. Ed. by M. Sardoč. Cham: Springer International Publishing, pp. 1–23.

– (2024b). *Experimental Evidence on Attitudes Toward Inequality and Fairness*. CEPR Discussion Paper 19620.

Alstadsæter, A., E. Casi, J. Miethe, and B. Stage (2023). *Lost in Information: National Implementation of Global Tax Agreements*. Skatteforsk Working Paper 10.

Alstadsæter, A., N. Johannesen, and G. Zucman (2018). "Who Owns the Wealth in Tax Havens? Macro Evidence and Implications for Global Inequality." *Journal of Public Economics* 162, pp. 89–100.

– (2019). "Tax Evasion and Inequality." *American Economic Review* 109 (6), pp. 2073–2103.

Alstadsæter, A., B. Planterose, G. Zucman, and A. Økland (2022). *Who Owns Offshore Real Estate? Evidence from Dubai*. EU Tax Observatory Working Paper 1.

Alves, C., C. da Costa, F. Lobel, and H. Moreira (2024). "Intrahousehold Inequality and the Joint Taxation of Household Earnings." *Journal of Public Economics* 239, p. 105208.

Andersen, A. G., S. Franklin, T. Getahun, A. Kotsadam, V. Somville, and E. Villanger (2023). "Does Wealth Reduce Support for Redistribution? Evidence from an Ethiopian Housing Lottery." *Journal of Public Economics* 224, p. 104939.

Andre, P. (forthcoming). "Shallow Meritocracy." *Review of Economic Studies*.

Andreoni, J. and J. Miller (2002). "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70 (2), pp. 737–753.

Armour, P., R. V. Burkhauser, and J. Larrimore (2016). "Using the Pareto Distribution to Improve Estimates of Topcoded Earnings." *Economic Inquiry* 54 (2), pp. 1263–1273.

Arneson, R. J. (1989). "Equality and Equal Opportunity for Welfare." *Philosophical Studies* 56 (1), pp. 77–93.

Auspurg, K. and T. Hinz (2014). *Factorial Survey Experiments*. Thousand Oaks: SAGE Publications.

Auspurg, K., T. Hinz, S. Liebig, and C. Sauer (2015). "The Factorial Survey as a Method for Measuring Sensitive Issues." *Improving Survey Methods: Lessons from Recent Research.* New York: Routledge/Taylor & Francis Group, pp. 137–149.

Auspurg, K., T. Hinz, and C. Sauer (2017). "Why Should Women Get Less? Evidence on the Gender Pay Gap from Multifactorial Survey Experiments." *American Sociological Review* 82 (1), pp. 179–210.

Badarinza, C. and T. Ramadorai (2018). "Home Away from Home? Foreign Demand and London House Prices." *Journal of Financial Economics* 130 (3), pp. 532–555.

Baily, N. (1978). "Some Aspects of Optimal Unemployment Insurance." *Journal of Public Economics* 10 (3), pp. 379–402.

Ballard-Rosa, C., L. Martin, and K. Scheve (2017). "The Structure of American Income Tax Policy Preferences." *The Journal of Politics* 79 (1), pp. 1–16.

Bargain, O., M. Dolls, H. Immervoll, D. Neumann, A. Peichl, N. Pestel, and S. Siegloch (2015). "Tax Policy and Income Inequality in the United States, 1979–2007." *Economic Inquiry* 53 (2), pp. 1061–1085.

Bargain, O., K. Orsini, and A. Peichl (2014). "Comparing Labor Supply Elasticities in Europe and the United States New Results." *Journal of Human Resources* 49 (3), pp. 723–838.

Bauer, M., J. Chytilová, and E. Miguel (2020). "Using Survey Questions to Measure Preferences: Lessons from an Experimental Validation in Kenya." *European Economic Review* 127, pp. 1–8.

Bénabou, R. and E. Ok (2001). "Social Mobility and the Demand for Redistribution: The Poum Hypothesis." *The Quarterly Journal of Economics* 116 (2), pp. 447–487.

Bénabou, R. (2000). "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* 90 (1), pp. 96–129.

Bergstrom, K. and W. Dodds (2021). "Optimal Taxation with Multiple Dimensions of Heterogeneity." *Journal of Public Economics* 200, p. 104442.

Berliant, M. and P. C. Boyer (2024). "Politics and Income Taxes: Progress and Progressivity." *Journal of Public Economic Theory* 26 (4), pp. 1–47.

Bick, A. and N. Fuchs-Schündeln (2017). "Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis." *Review of Economic Studies* 85 (3), pp. 1543–1576.

Bick, A., A. Blandin, and R. Rogerson (2022). "Hours and Wages." *The Quarterly Journal of Economics* 137 (3), pp. 1901–1962.

Bierbrauer, F. J. (2024). *Is a Uniform Price on Carbon Desirable? A Public Finance Perspective*. CESifo Working Paper.

Bierbrauer, F. J., P. Boyer, A. Peichl, and D. Weishaar (2025). *The Political Economy of Joint Taxation*. Working Paper.

Bierbrauer, F. J. and P. C. Boyer (2015). "Efficiency, Welfare, and Political Competition." *The Quarterly Journal of Economics* 131 (1), pp. 461–518.

Bierbrauer, F. J., P. C. Boyer, and E. Hansen (2023a). "Pareto-Improving Tax Reforms and the Earned Income Tax Credit." *Econometrica* 91 (3), pp. 1077–1103.

Bierbrauer, F. J., P. C. Boyer, and A. Peichl (2021). "Politically Feasible Reforms of Nonlinear Tax Systems." *American Economic Review* 111 (1), pp. 153–91.

Bierbrauer, F. J., P. C. Boyer, A. Peichl, and D. Weishaar (2023b). *The Taxation of Couples*. CEPR Discussion Paper 18138.

– (2024). *Pareto-Improvements, Welfare Trade-Offs and the Taxation of Couples*. CEPR Discussion Paper 19750.

Bierbrauer, F. J., A. Tsyvinski, and N. Werquin (2022). "Taxes and Turnout: When the Decisive Voter Stays at Home." *American Economic Review* 112 (2), pp. 689–719.

Binetti, A., F. Nuzzi, and S. Stantcheva (2024). "People's Understanding of Inflation." *Journal of Monetary Economics* 148, p. 103652.

Bittker, B. I. (1975). "Federal Income Taxation and the Family." *Stanford Law Review* 27 (6), pp. 1389–1463.

Blanchet, T., J. Fournier, and T. Piketty (2022). "Generalized Pareto Curves: Theory and Applications." *Review of Income and Wealth* 68 (1), pp. 263–288.

Blanchet, T., B. Garbinti, J. Goupille-Lebret, and C. Martínez-Toledano (2018). "Applying Generalized Pareto Curves to Inequality Analysis." *AEA Papers and Proceedings* 108, pp. 114–18.

Blau, F. D. and L. M. Kahn (2013). "Female Labor Supply: Why Is the United States Falling Behind?" *American Economic Review* 103 (3), pp. 251–256.

– (2017). "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3), pp. 789–865.

– (2007). "Changes in the Labor Supply Behavior of Married Women: 1980–2000." *Journal of Labor Economics* 25 (3), pp. 393–438.

Blömer, M. and A. Peichl (2020). *The Ifo Tax and Transfer Behavioral Microsimulation Model*. Ifo Working Paper 335.

Blundell, R., A. Duncan, and C. Meghir (1998). "Estimating Labor Supply Responses Using Tax Reforms." *Econometrica* 66 (4), pp. 827–861.

Blundell, R. and T. Macurdy (1999). "Labor Supply: A Review of Alternative Approaches." *Handbook of Labor Economics*. Vol. 3. Elsevier, pp. 1559–1695.

Boas, H., N. Johannesen, C. Kreiner, L. Larsen, and G. Zucman (2024). *Taxing Capital in a Globalized World: The Effects of Automatic Information Exchange*. NBER Working Paper 32714.

Boerma, J., A. Tsyvinski, and A. Zimin (2022). *Bunching and Taxing Multidimensional Skills*. NBER Working Paper 30015.

Bomare, J. (2019). *Tax Evasion and British Real Estate: An Estimation of the Offshore Wealth Held through British Real Estate*. Master Thesis.

Bomare, J. and S. Le Guern Herry (2022). *Will We Ever Be Able to Track Offshore Wealth? Evidence Fromthe Offshore Real Estate Market in the UK*. Working Paper.

Borella, M., M. De Nardi, and F. Yang (2023a). "Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labour Supply?" *Review of Economic Studies* 90 (1), pp. 102–131.

Borella, M., M. De Nardi, M. Pak, N. Russo, and F. Yang (2023b). "The Importance of Modeling Income Taxes over Time: U.S. Reforms and Outcomes." *Journal of the European Economic Association* 21 (6), pp. 2237–2286.

Borjas, G. J. (1980). "The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias." *Journal of Human Resources* 15 (3), pp. 409–423.

Boskin, M. J. and E. Sheshinski (1983). "Optimal Tax Treatment of the Family: Married Couples." *Journal of Public Economics* 20 (3), pp. 281–297.

Bourne, J., A. Ingianni, and R. McKenzie (2018). "What's in the Laundromat? Mapping and Characterising Offshore Owned Domestic Property in London." *EPB: Urban Analytics and City Science* 50 (9), pp. 2430–2451.

Brett, C. (2009). "Optimal Nonlinear Taxes for Families." *International Tax and Public Finance* 14 (3), pp. 225–261.

Brett, C. and J. A. Weymark (2017). "Voting over Selfishly Optimal Nonlinear Income Tax Schedules." *Games and Economic Behavior* 101, pp. 172–188.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden (2007). "The Pluralism of Fairness Ideals: An Experimental Approach." *American Economic Review* 97 (3), pp. 818–827.

Cappelen, A. W., J. Konow, E. Ø. Sørensen, and B. Tungodden (2013). "Just Luck: An Experimental Study of Risk-Taking and Fairness." *American Economic Review* 103 (4), pp. 1398–1413.

Cappelen, A. W., E. Ø. Sørensen, and B. Tungodden (2010). "Responsibility for What? Fairness and Individual Responsibility." *European Economic Review* 54 (3), pp. 429–441.

Carlier, G., X. Dupuis, J.-C. Rochet, and J. Thanassoulis (2024). "A General Solution to the Quasi Linear Screening Problem." *Journal of Mathematical Economics* 114, p. 103025.

Casi, E., C. Spengel, and B. Stage (2020). "Cross-Border Tax Evasion after the Common Reporting Standard: Game Over?" *Journal of Public Economics* 190, p. 104240.

Chetty, R. (2006a). "A New Method of Estimating Risk Aversion." *American Economic Review* 96 (5), pp. 1821–1834.

– (2006b). "A General Formula for the Optimal Level of Social Insurance." *Journal of Public Economics* 90 (10-11), pp. 1879–1901.

– (2008). "Moral Hazard versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy* 116 (2), pp. 173–234.

Chetty, R. and A. Finkelstein (2013). "Social Insurance: Connecting Theory to Data." *Handbook of Public Economics*. Vol. 5. Elsevier, pp. 111–193.

Choi, S., R. Fisman, D. Gale, and S. Kariv (2007). "Consistency and Heterogeneity of Individual Behavior under Uncertainty." *American Economic Review* 97 (5), pp. 1921–1938.

Cohen, G. A. (1989). "On the Currency of Egalitarian Justice." *Ethics* 99 (4), pp. 906–944.

Collin, M. (2021). *What Lies beneath: Evidence from Leaked Account Data on How Elites Use Offshore Banking*. Global Economy and Development Working Paper 156. Brookings Institution.

Collin, M., F. M. Hollenbach, and D. Szakonyi (2022). *The Impact of Beneficial Ownership Transparency on Illicit Purchases of U.S. Property*. Global Economy and Development Working Paper 170. Brookings Institution.

Coppola, A., M. Maggiori, B. Neiman, and J. Schreger (2021). "Redrawing the Map of Global Capital Flows: The Role of Cross-Border Financing and Tax Havens." *The Quarterly Journal of Economics* 136 (3), pp. 1499–1556.

Costa-Font, J. and F. Cowell (2014). "Social Identity and Redistributive Preferences: A Survey." *Journal of Economic Surveys* 29 (2), pp. 357–374.

Cremer, H., J.-M. Lozachmeur, and P. Pestieau (2012). "Income Taxation of Couples and the Tax Unit Choice." *Journal of Population Economics* 25 (2), pp. 763–778.

Crossley, T. F. and S.-H. Jeon (2007). "Joint Taxation and the Labour Supply of Married Women: Evidence from the Canadian Tax Reform of 1988." *Fiscal Studies* 28 (3), pp. 343–365.

Cvijanović, D. and C. Spaenjers (2021). ""We'll Always Have Paris": Out-of-country Buyers in the Housing Market." *Management Science* 67 (7), pp. 4120–4138.

Damgaard, J., T. Elkjaer, and N. Johannesen (2024). "What Is Real and What Is Not in the Global FDI Network?" *Journal of International Money and Finance* 140, p. 102971.

Day, B., I. J. Bateman, R. T. Carson, D. Dupont, J. J. Louviere, S. Morimoto, R. Scarpa, and P. Wang (2012). "Ordering Effects and Choice Set Awareness in Repeat-Response Stated Preference Studies." *Journal of Environmental Economics and Management* 63 (1), pp. 73–91.

De Simone, L., R. Lester, and K. Markle (2020). "Transparency and Tax Evasion: Evidence from the Foreign Account Tax Compliance Act (FATCA)." *Journal of Accounting Research* 58 (1), pp. 105–153.

Dechezlepretre, A., A. Fabre, T. Kruse, B. Planterose, A. S. Chico, and S. Stantcheva (forthcoming). "Fighting Climate Change: International Attitudes Toward Climate Policies." *American Economic Review*.

Del Rey, E., J. Naval, and J. I. Silva (2022). "Hours and Wages: A Bargaining Approach." *Economics Letters* 217, p. 110652.

Diamond, P. A. (1998). "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review* 88 (1), pp. 83–95.

Dworkin, R. (1981a). "What Is Equality? Part 1: Equality of Welfare." *Philosophy & Public Affairs* 10 (3), pp. 185–246.

– (1981b). "What Is Equality? Part 2: Equality of Resources." *Philosophy & Public Affairs* 10 (4), pp. 283–345.

Eissa, N. and H. W. Hoynes (2004). "Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit." *Journal of Public Economics* 88 (9-10), pp. 1931–1958.

Eissa, N., H. Kleven, and C. Kreiner (2008). "Evaluation of Four Tax Reforms in the United States: Labor Supply and Welfare Effects for Single Mothers." *Journal of Public Economics* 92 (3), pp. 795–816.

Engelmann, D., E. Janeba, L. Mechtenberg, and N. Wehrhöfer (2023). "Preferences over Taxation of High-Income Individuals: Evidence from a Survey Experiment." *European Economic Review* 157, p. 104505.

Enke, B., R. Rodríguez-Padilla, and F. Zimmermann (2022). "Moral Universalism: Measurement and Economic Relevance." *Management Science* 68 (5), pp. 3590–3603.

European Union (2019). "Directive (EU) 2019/1158 of the European Parliament and of the Council of 20 June 2019 on Work-Life Balance for Parents and Carers and Repealing Council Directive 2010/18/EU." *Official Journal of the European Union* L 188, pp. 79–93.

Evans, D. J. (2005). "The Elasticity of Marginal Utility of Consumption: Estimates for 20 OECD Countries." *Fiscal Studies* 26 (2), pp. 197–224.

Falk, A., A. Becker, T. Dohmen, D. Huffman, and U. Sunde (2023). "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences." *Management Science* 69 (4), pp. 1935–1950.

Fallucchi, F., D. Nosenzo, and E. Reuben (2020). "Measuring Preferences for Competition with Experimentally-Validated Survey Questions." *Journal of Economic Behavior & Organization* 178, pp. 402–423.

Farhi, E., C. Sleet, I. Werning, and S. Yeltekin (2012). "Non-Linear Capital Taxation without Commitment." *Review of Economic Studies* 79 (4), pp. 1469–1493.

Feenberg, D. and E. Coutts (1993). "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12(1), pp. 189–194.

Fehr, E., T. Epper, and J. Senn (forthcoming). "Social Preferences and Redistributive Politics." *The Review of Economics and Statistics*.

Ferey, A., B. Lockwood, and D. Taubinsky (2024). "Sufficient Statistics for Nonlinear Tax Systems with General Across-Income Heterogeneity." *American Economic Review* 114 (10), pp. 3206–3249.

Fernández-Kranz, D. and N. Rodríguez-Planas (2021). "Too Family Friendly? The Consequences of Parent Part-Time Working Rights." *Journal of Public Economics* 197, p. 104407.

Ferrario, B. and S. Stantcheva (2022). "Eliciting People's First-Order Concerns: Text Analysis of Open-Ended Survey Questions." *AEA Papers and Proceedings* 112, pp. 163–169.

Fisman, R., K. Gladstone, I. Kuziemko, and S. Naidu (2020). "Do Americans Want to Tax Wealth? Evidence from Online Surveys." *Journal of Public Economics* 188, p. 104207.

Fisman, R., P. Jakiela, S. Kariv, and S. Vannutelli (2023). "The Distributional Preferences of Americans, 2013–2016." *Experimental Economics* 26 (4), pp. 727–748.

Fisman, R., S. Kariv, and D. Markovits (2007). "Individual Preferences for Giving." *American Economic Review* 97 (5).

Fitzenberger, B. and A. Seidlitz (2020). "The 2011 Break in the Part-Time Indicator and the Evolution of Wage Inequality in Germany." *Journal for Labour Market Research* 54 (1), pp. 1–14.

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, and M. Westberry (2021). *Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [Dataset]*. https://doi.org/10.18128/D030.V9.0.

Fuenmayor, A., R. Granell, and M. Mediavilla (2018). "The Effects of Separate Taxation on Labor Participation of Married Couples. An Empirical Analysis Using Propensity Score." *Review of Economics of the Household* 16, pp. 541–561.

Gabaix, X. and R. S. J. Koijen (2021). *In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis*. NBER Working Paper 28967.

Gaertner, W. and E. Schokkaert (2012). *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice*. Cambridge: Cambridge University Press.

Gaertner, W. and L. Schwettmann (2007). "Equity, Responsibility and the Cultural Dimension." *Economica* 74 (296), pp. 627–649.

Ganzer, A., A. Schmucker, J. Stegmaier, and S. Wolter (2022). *Establishment History Panel 1975–2020*. FDZ Datenreport 12/2022. Insitute for Employment Research.

Gayle, G.-L. and A. Shephard (2019). "Optimal Taxation, Marriage, Home Production, and Family Labor Supply." *Econometrica* 87 (1), pp. 291–326.

Gillen, B., E. Snowberg, and L. Yariv (2019). "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." *Journal of Political Economy* 127 (4), pp. 1826–1863.

Goldin, C. (2014). "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4), pp. 1091–1119.

Goldin, C. and L. F. Katz (2011). "The Cost of Workplace Flexibility for High-Powered Professionals." *The Annals of the American Academy of Political and Social Science* 638 (1), pp. 45–67.

Golosov, M. and I. Krasikov (2025). *The Optimal Taxation of Couples*. NBER Working Paper 31140.

Golosov, M., A. Tsyvinski, and N. Werquin (2014). *A Variational Approach to the Analysis of Tax Systems*. NBER Working Paper 20780.

Gorback, C. S. and B. J. Keys (2020). *Global Capital and Local Assets: House Prices, Quantities, and Elasticities*. NBER Working Paper 27370.

Gravelle, J. G. (2015). *Tax Havens: International Tax Avoidance and Evasion*. Congressional Research Service Report R40623.

Guner, N., R. Kaygusuz, and G. Ventura (2012). "Taxation and Household Labour Supply." *Review of Economic Studies* 79 (3), pp. 1113–1149.

– (2014). "Income Taxation of U.S. Households: Facts and Parametric Estimates." *Review of Economic Dynamics* 17 (4), pp. 559–581.

Gunnarsson, Å. and M. Eriksson (2017). "Eliminating the Secondary Earner Bias. Policy Lessons from the Introduction of Partial Individual Taxation in Sweden in 1971." *Nordic Tax Journal* 2017 (1), pp. 89–99.

Gustafsson, S. (1992). "Separate Taxation and Married Women's Labor Supply: A Comparison of West Germany and Sweden." *Journal of Population Economics* 5 (1), pp. 61–85.

Guyton, J., P. Langetieg, D. Reck, M. Risch, and G. Zucman (2021). *Tax Evasion at the Top of the Income Distribution: Theory and Evidence*. NBER Working Paper 28542.

Haaland, I., C. Roth, and J. Wohlfart (2023). "Designing Information Provision Experiments." *Journal of Economic Literature* 61 (1), pp. 3–40.

Hanlon, M., E. L. Maydew, and J. R. Thornock (2015). "Taking the Long Way Home: US Tax Evasion and Offshore Investments in US Equity and Debt Markets." *The Journal of Finance* 70 (1), pp. 257–287.

Harmon, C., H. Oosterbeek, and I. Walker (2003). "The Returns to Education: Microeconomics." *Journal of Economic Surveys* 17 (2), pp. 115–156.

Heathcote, J., K. Storesletten, and G. L. Violante (2020). "Presidential Address 2019: How Should Tax Progressivity Respond to Rising Income Inequality?" *Journal of the European Economic Association* 18 (6), pp. 2715–2754.

– (2017). "Optimal Tax Progressivity: An Analytical Framework." *The Quarterly Journal of Economics* 132 (4), pp. 1693–1754.

Hendren, N. and B. Sprung-Keyser (2020). "A Unified Welfare Analysis of Government Policies." *The Quarterly Journal of Economics* 135 (3), pp. 1209–1318.

Henrich, J., S. J. Heine, and A. Norenzayan (2010). "The Weirdest People in the World?" *Behavioral and Brain Sciences* 33 (2-3), pp. 61–83.

Hirsch, B. T. (2005). "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills." *ILR Review* 58 (4), pp. 525–551.

Holter, H., D. Krüger, and S. Stepancuk (2023). *Until the IRS Do Us Part: (Optimal) Taxation of Households*. Working Paper.

Hufe, P. and D. Weishaar (2025). *Just Cheap Talk? Investigating Fairness Preferences in Hypothetical Scenarios*. CESifo Working Paper 11647.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and E. Saez (2007). "Welfare Reform in European Countries: A Microsimulation Analysis." *Economic Journal* 117 (516), pp. 1–44.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and N. Verdelin (2011). "Optimal Tax and Transfer Programs for Couples with Extensive Labor Supply Responses." *Journal of Public Economics* 95 (11-12), pp. 1485–1500.

Jacquet, L. and E. Lehmann (2021a). *How to Tax Different Incomes?* CEPR Discussion Paper 16571.

– (2021b). "Optimal Income Taxation with Composition Effects." *Journal of the European Economic Association* 19 (2), pp. 1299–1341.

Jasso, G. and M. Webster (1997). "Double Standards in Just Earnings for Male and Female Workers." *Social Psychology Quarterly* 60 (1), pp. 66–78.

– (1999). "Assessing the Gender Gap in Just Earnings and Its Underlying Mechanisms." *Social Psychology Quarterly* 62 (4), pp. 367–380.

Johannesen, N. (2014). "Tax Evasion and Swiss Bank Deposits." *Journal of Public Economics* 111, pp. 46–62.

Johannesen, N., J. Miethe, and D. Weishaar (2024a). *Homes Incorporated: Offshore Ownership of Real Estate in the U.K.* Oxford CBT Working Paper 24/14.

Johannesen, N., D. Reck, M. Risch, J. Slemrod, J. Guyton, and P. Langetieg (2024b). "The Offshore World According to FATCA: New Evidence on the Foreign Wealth of U.S. Households." *Tax Policy and the Economy* 38, pp. 61–99.

Johannesen, N. and G. Zucman (2014). "The End of Bank Secrecy? An Evaluation of the G20 Tax Haven Crackdown." *American Economic Journal: Economic Policy* 6 (1), pp. 65–91.

Kalíšková, K. (2014). "Labor Supply Consequences of Family Taxation: Evidence from the Czech Republic." *Labour Economics* 30, pp. 234–244.

Kaygusuz, R. (2010). "Taxes and Female Labor Supply." *Review of Economic Dynamics* 13 (4), pp. 725–741.

Kleven, H. J., C. T. Kreiner, and E. Saez (2009). "The Optimal Income Taxation of Couples." *Econometrica* 77 (2), pp. 537–560.

Knoll, K., M. Schularick, and T. Steger (2017). "No Price like Home: Global House Prices, 1870–2012." *American Economic Review* 107 (2), pp. 331–53.

Konow, J. (1996). "A Positive Theory of Economic Fairness." *Journal of Economic Behavior & Organization* 31 (1), pp. 13–35.

– (2000). "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." *American Economic Review* 90 (4), pp. 1072–1092.

Konow, J., T. Saijo, and K. Akai (2020). "Equity versus Equality: Spectators, Stakeholders and Groups." *Journal of Economic Psychology* 77, p. 102171.

Kuhn, P. and F. Lozano (2008). "The Expanding Workweek? Understanding Trends in Long Work Hours among U.S. Men, 1979–2006." *Journal of Labor Economics* 26 (2), pp. 311–343.

LaLumia, S. (2008). "The Effects of Joint Taxation of Married Couples on Labor Supply and Non-Wage Income." *Journal of Public Economics* 92 (7), pp. 1698–1719.

Larrimore, J., R. V. Burkhauser, S. Feng, and L. Zayatz (2008). "Consistent Cell Means for Topcoded Incomes in the Public Use March CPS (1976–2007)." *Journal of Economic and Social Measurement* 33 (2-3), pp. 89–128.

Layard, R., G. Mayraz, and S. Nickell (2008). "The Marginal Utility of Income." *Journal of Public Economics* 92 (8-9), pp. 1846–1857.

Leenders, W., A. Lejour, S. Rabaté, and M. van 't Riet (2023). "Offshore Tax Evasion and Wealth Inequality: Evidence from a Tax Amnesty in the Netherlands." *Journal of Public Economics* 217, p. 104785.

Lemieux, T. (2006). "Postsecondary Education and Increasing Wage Inequality." *American Economic Review* 96 (2), pp. 195–199.

Londoño-Vélez, J. and J. Ávila-Mahecha (2021). "Enforcing Wealth Taxes in the Developing World: Quasi-Experimental Evidence from Colombia." *American Economic Review: Insights* 3 (2), pp. 131–148.

Londoño-Vélez, J. and D. Tortarolo (2022). *Revealing 21% of GDP in Hidden Assets: Evidence from Argentina's Tax Amnesties*. EU Tax Observatory Working Paper 6.

Lorenz, N. and D. Sachs (2016). "Identifying Laffer Bounds: A Sufficient-Statistics Approach with an Application to Germany." *Scandinavian Journal of Economics* 118 (4), pp. 646–665.

Maestas, N., K. J. Mullen, D. Powell, T. Von Wachter, and J. B. Wenger (2023). "The Value of Working Conditions in the United States and Implications for the Structure of Wages." *American Economic Review* 113 (7), pp. 2007–2047.

Magnac, T. and S. Roux (2021). "Heterogeneity and Wage Inequalities over the Life Cycle." *European Economic Review* 134, p. 103715.

Malkov, E. (2021). *Optimal Income Taxation of Singles and Couples*. Working Paper.

Mazumder, B. (2005). "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *The Review of Economics and Statistics* 87 (2), pp. 235–255.

Meltzer, A. H. and S. F. Richard (1981). "A Rational Theory of the Size of Government." *Journal of Political Economy* 89 (5), pp. 914–927.

Menkhoff, L. and J. Miethe (2019). "Tax Evasion in New Disguise? Examining Tax Havens' International Bank Deposits." *Journal of Public Economics* 176, pp. 53–78.

Milgrom, P. and I. Segal (2002). "Envelope Theorems for Arbitrary Choice Sets." *Econometrica* 70 (2), pp. 583–601.

Mollerstrom, J., B.-A. Reme, and E. Ø. Sørensen (2015). "Luck, Choice and Responsibility: An Experimental Study of Fairness Views." *Journal of Public Economics* 131, pp. 33–40.

Moore, R. (2022). "How London Became the Place to Be for Putin's Oligarchs." *Guardian*.

National Council of German Women's Organizations (2021). *Gleichstellungs-Check Des Deutschen Frauenrats*. Auswertung Der Wahlprogramme Zur Bundestagswahl 2021.

Neisser, C. (2021). "The Elasticity of Taxable Income: A Meta-Regression Analysis." *The Economic Journal* 131 (640), pp. 3365–3391.

Nozick, R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books.

OECD (2020). *How's Life? 2020: Measuring Well-being*. Paris: Organisation for Economic Co-operation and Development.

Olivetti, C. and B. Petrongolo (2017). "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries." *Journal of Economic Perspectives* 31 (1), pp. 205–230.

Olsson, J. (2025). "Singles, Couples, and Their Labor Supply: Long-Run Trends and Short-Run Fluctuations." *American Economic Journal: Macroeconomics* 17 (1), pp. 1–34.

Piketty, T. (1995). "Social Mobility and Redistributive Politics." *The Quarterly Journal of Economics* 110 (3), pp. 551–584.

– (1997). "La Redistribution Fiscale Face Au Chômage." *Revue française d'économie* 12, pp. 157–201.

– (2003). "Income Inequality in France, 1901-1998." *Journal of Political Economy* 111 (5), pp. 1004–1042.

Piketty, T. and E. Saez (2003). "Income Inequality in the United States, 1913-1998." *The Quarterly Journal of Economics* 118 (1), pp. 1–41.

Risch, M. (2024). "Does Taxing Business Owners Affect Employees? Evidence From A Change in the Top Marginal Tax Rate." *The Quarterly Journal of Economics* 139 (1), pp. 637–692.

Roberts, K. W. (1977). "Voting over Income Tax Schedules." *Journal of Public Economics* 8 (3), pp. 329–340.

Rochet, J.-C. (2024). "Multidimensional Screening after 37 Years." *Journal of Mathematical Economics* 113, p. 103010.

Roemer, J. E. and A. Trannoy (2016). "Equality of Opportunity: Theory and Measurement." *Journal of Economic Literature* 54 (4), pp. 1288–1332.

Romano, J. P. and M. Wolf (2005). "Stepwise Multiple Testing as Formalized Data Snooping." *Econometrica* 73 (4), pp. 1237–1282.

– (2016). "Efficient Computation of Adjusted p -Values for Resampling-Based Stepdown Multiple Testing." *Statistics & Probability Letters* 113, pp. 38–40.

Sá, F. (2014). "Immigration and House Prices in the UK." *The Economic Journal* 125 (587), pp. 1393–1424.

– (forthcoming). "The Effect of Foreign Investors on Local Housing Markets: Evidence from the UK." *Journal of Economic Geography*.

Sachs, D., A. Tsyvinski, and N. Werquin (2020). "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium." *Econometrica* 88 (2), pp. 469–493.

Saez, E. (2002). "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *The Quarterly Journal of Economics* 117 (3), pp. 1039–1073.

Saez, E. and S. Stantcheva (2016). "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." *American Economic Review* 106 (1), pp. 24–45.

Saez, E. (2001). "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68 (1), pp. 205–229.

Saez, E., J. B. Slemrod, and S. H. Giertz (2012). "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50 (1), pp. 3–50.

Scheuer, F. and A. Wolitzky (2016). "Capital Taxation under Political Constraints." *American Economic Review* 106 (8), pp. 2304–2328.

Schmucker, A., S. Seth, and P. vom Berge (2023). *Sample of Integrated Labour Market Biographies (SIAB) 1975-2021*. FDZ Datenreport 02/2023. Institute for Employment Research.

Schroyen, F. (2003). "Redistributive Taxation and the Household: The Case of Individual Filings." *Journal of Public Economics* 87 (11), pp. 2527–2547.

Schuetz, J., S. Uebelmesser, R. Baginski, and C. Aprea (2023). "Pension Reform Preferences in Germany: Does Information Matter?" *European Journal of Political Economy* 80, p. 102476.

Selin, H. (2014). "The Rise in Female Employment and the Role of Tax Incentives. An Empirical Analysis of the Swedish Individual Tax Reform of 1971." *International Tax and Public Finance* 21 (5), pp. 894–922.

Shaxson, N. (2013). "A Tale of Two Londons." *Vanity Fair*.

Spiritus, K., E. Lehmann, F. T. Zoutman, and S. Renes (forthcoming). "Optimal Taxation with Multiple Incomes and Types." *Theoretical Economics*.

Stantcheva, S. (2021). "Understanding Tax Policy: How Do People Reason?" *The Quarterly Journal of Economics* 136 (4), pp. 2309–2369.

– (2022). *Understanding of Trade*. NBER Working Paper 30040, w30040.

– (2023). "How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible." *Annual Review of Economics* 15, pp. 205–35.

Starmans, C., M. Sheskin, and P. Bloom (2017). "Why People Prefer Unequal Societies." *Nature Human Behaviour* 1 (4), pp. 1–7.

Stephens, M. and J. Ward-Batts (2004). "The Impact of Separate Taxation on the Intra-Household Allocation of Assets: Evidence from the UK." *Journal of Public Economics* 88 (9), pp. 1989–2007.

Tørsløv, T., L. Wier, and G. Zucman (2023). "The Missing Profits of Nations." *Review of Economic Studies* 90 (3), pp. 1499–1534.

van de Ven, J. and N. Hérault (2022). "The Evolution of Tax Implicit Value Judgements in the UK: 1968–2018." *Oxford Economic Papers* 74 (2), pp. 594–609.

vom Berge, P. and A. Schmucker (2023). *Creating Cross-Sectional Data and Biographical Variables with the Sample of Integrated Labour Market Biographies 1975-2021 - Programming Examples for Stata*. FDZ Methodenreport 03/2023. Institute for Employment Research.

vom Berge, P., M. Umkehrer, and S. Wanger (2023). "A Correction Procedure for the Working Hours Variable in the IAB Employee History." *Journal for Labour Market Research* 57 (1), pp. 1–16.

Wissenschaftlicher Beirat beim Bundesministerium der Finanzen (2018). *Zur Reform Der Besteuerung von Ehegatten*. Gutachten 02/2018. Bundesministerium der Finanzen.

Wiswall, M. and B. Zafar (2018). "Preference for the Workplace, Investment in Human Capital, and Gender." *The Quarterly Journal of Economics* 133 (1), pp. 457–507.

Woerner, A., T. Imai, D. D. Pace, and K. M. Schmidt (2024). "How to Increase Public Support for Carbon Pricing with Revenue Recycling." *Nature Sustainability* 7 (12), pp. 1633–1641.

Wolf, E. (2002). "Lower Wage Rates for Fewer Hours? A Simultaneous Wage-Hours Model for Germany." *Labour Economics* 9 (5), pp. 643–663.

Zucman, G. (2013). "The Missing Wealth of Nations: Are Europe and the U.S. Net Debtors or Net Creditors?" *The Quarterly Journal of Economics* 128 (3), pp. 1321–1364.

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

München, 10. März 2025

Daniel Weishaar