Towards Gender Equality and Enhanced Innovation: Taxation of Couples, Parental Benefits, and Public R&D

Leonie Koch

Munich 2024

Towards Gender Equality and Enhanced Innovation:

Taxation of Couples, Parental Benefits, and Public R&D

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

2024

vorgelegt von
Leonie Koch

Referent: Prof. Dr. Dr. h.c. Clemens Fuest
Korreferent: Prof. Dr. Dominik Sachs
Promotionsabschlussberatung: 10. Juli 2024

Tag der mündlichen Prüfung: 03. Juli 2024
Namen der Berichterstatter: Clemens Fuest, Dominik Sachs, Andreas Peichl

To Mum

Acknowledgments

Writing this dissertation has been a fun but challenging endeavor, that benefited greatly from the support and interaction with my outstanding colleagues and friends at LMU. First and foremost, I want to thank Clemens Fuest, my first supervisor. Throughout the years, you have granted me the freedom and support to develop and pursue my research agenda, while continuously guiding me with extremely valuable feedback that shaped my path forward. I am equally grateful to my second supervisor and co-author, Dominik Sachs. Your thorough questions and insightful suggestions motivated me significantly. I enjoyed working with you, as you are not only a great researcher but an equally good teacher. I also want to thank Andreas Peichl for his excellent feedback and encouragement as a co-author as well as for serving as the third member of my dissertation committee.

Turning to my co-authors, all of them deserve special credit for this dissertation. Elena and Carina, thank you for our lively discussions and the rewarding teamwork. It has been fun to share the excitement about tax policy and gender equality. Thank you, Amelie, for joining the parental benefit project and the productive last months. I especially enjoyed our shared time at the data center which made the tedious work with the TPP less painful. Thank you, Uli and Emanuel, for providing me with the necessary guidance and continued encouragement to keep me on a productive path. Moreover, I want to thank Martin for taking me on board as a research assistant and for the collaboration on Chapter 4. It was a tremendous joy to work with all of you and I learned a lot. Our collaboration enriched my journey substantially.

My PhD has been shaped by many great colleagues who have become friends. Sharing lunch breaks and working as well as teaching alongside with you, Daniel, Amelie and Manuel, has been a true pleasure. Thank you, Mehmet, for the great atmosphere in our shared office and the mutual support during these last challenging months. I am also thankful to the junior faculty members Jakob and Jonas for providing me with a breadth of comments and feedback on my projects. It was great to be part of such a fun and encouraging group at CES. Martina, Renate, Ronja, Uschi, Uli and Susanne, your sustained support on the administrative side has been outstanding. I am grateful for your always welcoming and warm manner. My projects have also benefited a lot from the stimulating and welcoming research environment of the Public Economics Group at LMU.

Additionally, the first three chapters would not have been possible without the excellent support from Hannes Fauser from the Research Data Center of the German Federal Statistical Office, Anja Hlawatsch, Karen Meyer and Selina Straub from the Bavarian Statistical Office as well as Agnesa Nimanaj, Nicolas Bunde and Sebastian Wichert from the EBDC at ifo.

An important portion of support throughout this time came from friends. Thanks to those I knew before and those that I got to know in Munich. I want to especially thank, Luisa and Sarah. I am grateful for having been on this journey with you. Outside of

Munich, Laura and Mariale have been two extremely valuable voices of encouragement and feedback that have helped me to navigate through these years.

My deepest gratitude goes to my family. My parents, Regina and Joachim, and my sisters, Ann-Christine and Sabrina, have been the backbone of support and encouragement throughout my life. I would also like to thank my brothers-in-law, Lucas and Tom, my family-in-law, Barbara, Michael, Pauline, Nouman, Luis, and Samira as well as my aunt and uncle, Daniela and Steve, for your kindness and enduring support. Finally, I would like to thank my husband Philipp for your unconditional support. Your love and humour have been my source of strength and I am beyond grateful for having you by my side. Mads, thank you for bringing such joy into my life and letting me forget about all the PhD related challenges.

LEONIE KOCH
Munich
March 2024

Contents

List of Figures

List of Tables

Preface

Today, many economies around the world grapple with decelerating and meagre growth rates. This is caused by, among others, demographic challenges such as an aging population and the rapid pace of technological transformation. These structural shifts in labor markets, marked by a notable shortage of skilled labor, pose significant barriers to achieving sustainable economic growth. Amidst this backdrop, supply-side policies that strengthen the overall labor supply, make more efficient use of the workforce and enhance productivity are imperative as potential pathways to spur economic growth.

Firstly, the expansion of the labor supply represents a fundamental approach to fueling economic growth. This encompasses strategies to elevate the integration of underrepresented groups, such as women, into the workforce to contribute more substantially to the labor market. A recent study by the World Bank reports that "eliminating the gender gap over the next decade would essentially double the current global growth rate" (World Bank 2024, p.16). Yet, despite notable gains in female labor force participation over recent decades, significant gender employment gaps remain, particularly post-childbirth (Bertrand 2020; Blau and Kahn 2017; Kleven, Landais, and Søgaard 2019). A growing body of literature has identified several factors that influence women's labor market decisions and dynamics: job flexibility and continuity, family policies and gender norms (Bertrand 2020; Goldin 2014; Olivetti and Petrongolo 2017). However, causal evidence on the impact of taxation of couples and the design of parental leave benefits is still scarce. For economists and policymakers, it is therefore crucial to thoroughly understand the constraints and determinants further to be able to design effective public policies that can increase female labor supply. To fill this gap, I pursue two different angles of women's labor supply determinants that highlight the effects of current tax as well as family policies, in this dissertation. The first and second chapters evaluate the role of income taxation of couples on partners' labor supply. The third chapter shifts the focus to parental leave benefits and analyzes its impact on female labor supply decisions prior to childbirth.

Secondly, enhancing productivity gains by e.g. stimulating innovation is another important policy dimension to bolster economic growth (Romer 1990). Innovation is a key driver in improving productivity by creating new technologies to make production more efficiently, reducing costs and increasing outputs. For this, policymakers can stimulate innovation by directly funding investments in research and development (R&D) of public institutions or supporting R&D by private firms. There is, however, limited empirical evidence available on how best to encourage R&D investments by firms, and the spillovers of public R&D to private firms. Therefore, the fourth chapter provides insights into the mechanisms by which public R&D spending can promote private R&D. Furthermore, it investigates the role as well as magnitude of local knowledge spillovers of public R&D.

I employ quasi-experimental methods. By leveraging as-good-as-random variation in the implementation of policies, quasi-experimental methods identify causal effects on

the outcomes of interest. These methods enable the identification of the effects of interest without making explicit assumptions on the structure of the decision problem, as identification stems from variation that is independent of the underlying decision factors. Complementary, I discuss the potential mechanisms behind my findings. Chapter 1, 2 and 3 exploit policy variations of tax and parental leave policies to identify the effects of interest. Chapter 4 uses an instrumental-variable approach to identify local knowledge spillovers of public R&D.

In terms of data, I use administrative data from government and international organization sources. These data sets are distinguished by a large number of observations and high information quality, since the data collection process typically penalizes misreporting. Chapters 1, 2 and 3 are conducted with longitudinal income tax return data from the German Federal Statistical Office. For these chapters, my co-authors and I had access to the universe of German taxpayers who file tax returns. Additionally, for Chapters 1 and 2 my co-authors and I are, to the best of our knowledge, the first to link partners before filing jointly - a great advantage compared to previous studies using the German tax return data. The reliability of the micro tax data is especially beneficial when studying sensitive information such as income, which often suffers from measurement error in survey data. Chapter 4 is conducted with patent application and citation data from the OECD.

The remainder of this section provides an overview of the four chapters in this dissertation. Each chapter is a self-contained paper with supplementary material provided in a corresponding appendix. A consolidated bibliography for all four chapters is provided at the end.

Chapters 1 and 2 focus on the impact of taxation of couples on the labor supply of spouses. Joint taxation is widely perceived as impeding gender equality by discouraging the labor supply of the secondary earner, typically women. The chapters take two different perspectives.

Chapter 1 addresses the following question: How does a move from joint taxation to individual taxation affect the gender earnings gap? I exploit two specific features of the income tax in Germany to study newlywed spouses' labor supply responses to a move to individual taxation. First, married couples can freely choose between individual and joint income taxation for both their withholding income tax and final income tax. Second, the 2013 administrative reform to the German withholding income tax introduced the individual withholding tax schedule as the default for newlyweds. This implied lower average and marginal tax rates for the secondary earner, typically the wife, if couples stayed with the default.

I use novel data on the universe of German newlywed taxpayers filing tax returns based on administrative micro data from German tax authorities and link spouses before and after filing jointly. My sample are dual earner newlyweds who both report positive labor income and no income from self-employment in the pre-marriage year, so that the withholding income tax is relevant. First, I investigate the impact of the default introduction on the choice of withholding income tax schedules for newlyweds by leveraging the default introduction as an exogenous change to the share of newlyweds choosing individual taxation. Employing a stacked Difference-in-Differences approach, I find that the share of newlyweds choosing the individual withholding tax schedule increases from

57% to 70% for newlywed couples for which the primary earner is male and the secondary earner is female after the default introduction. Thus, the default matters for individuals' responses to taxation.

I then turn to my primary outcome of interest, examining the labor income responses of both spouses to the change in marginal withholding tax rates resulting from the move from joint to individual tax schedules. Using newlywed cohorts before the default introduction as the control group and those after as the treatment group, I employ marriage as the event. The findings indicate that female labor force participation increases by 1 percentage point, while male labor force participation remains unchanged. Studying the intensive margin of labor supply, I find that secondary-earning newlywed women (constituting approximately 75% of newlywed women) experience a 3% increase in their labor income in the first year following the default introduction, with no discernible effect for newlywed men. This effect for women persists for two years post-marriage, gradually diminishing thereafter. Notably, the effects are more substantial and persistent when the sample is restricted to newlywed women with children. The partner pay gap, representing the difference between labor incomes of primary earner and secondary relative to the income of the couple, narrows by 2 percentage points after the default introduction. This demonstrates how changes in marginal withholding tax rates can enhance gender equality by equalizing spouses' labor income and raise overall labor supply. Finally, I estimate an elasticity of labor income with respect to the net-of-withholding-tax rate of 1.7 for newlywed women following the standard approach to estimate the elasticity of taxable income. These results document a substantial response to the reform and provide evidence that not only final income taxes but also withholding income taxes matter for labor supply.

Chapter 2, which is joint work with Elena Herold and Carina Neisser, shifts the focus from the withholding income tax to the final income tax. We exploit the introduction of joint taxation for existing same-sex civil partnerships in Germany in 2013. Three features of this reform allow us to isolate the incentive part of joint income taxation on partners' labor supply. First, the reform offers a quasi-experimental setting that exogenously changes partners' tax schedule. Due to kinks in the German tax code, changes in marginal tax rates are heterogeneous depending on a household's overall and relative income. Second, civil partnership incentives do not change for couples who already are in a civil partnership as the reform only changed partners' tax burden. Third, to disentangle gender norms and the role of the tax system in place, same-sex couples are a valuable comparison group as they are arguably less affected by traditional within-couple gender norms.

We use novel data on the universe of all same-sex couples who file jointly based on administrative micro data from German tax authorities. We uniquely link partners before and after filing jointly. Employing a dynamic Difference-in-Differences approach with different-sex couples as the control group, we find that a same-sex secondary earner experiences a significant decline in his income of up to 17% after filing taxes jointly relative to a different-sex secondary earner and prior to the introduction of joint taxation for same-sex couples. The partner pay gap significantly widens by up to 14 percentage points three years after the introduction of joint filing for same-sex couples relative to the partner pay gap of different-sex couples. Therefore, joint taxation has substantial effects on partners' income and intra-household inequality.

Additionally, we explore whether same-sex couples engage in tax planning in the form of shifting the withholding tax burden among partners. We find that same-sex civil partners are, irrespective of the size of their partner pay gap, less likely to choose a withholding income tax schedule which results in a higher marginal tax burden for the secondary earner compared to different-sex spouses with a male primary earner.

Chapter 3, which is joint work with Ulrich Glogowsky, Amelie Grosenick, Emanuel Hansen, Andreas Peichl and Dominik Sachs, focuses on the impact of parental leave benefits on pre-birth earnings. The introduction of a new German parental leave benefit in 2007 provides a unique quasi-experimental setting. As the amount of the benefit is linked to net earnings in the 12 months prior to childbirth, the effective marginal tax rates of expectant parents are substantially reduced. The benefit thus implies strong incentives to increase pre-birth earnings. Further, net-earnings can also be affected by tax planning in the form of adjustments in withholding tax schedules for married couples in Germany.

We leverage administrative micro data from German tax authorities. We have access to the universe of married couples filing tax returns and expecting their first child. We employ a Difference-in-Income Trends estimator, similar to a Difference-in-Differences approach, by comparing the trend differential in expecting mothers' net income during the post-reform period within a validation region where marginal tax rates remain stable in both periods to that in an identification region where significant changes in marginal tax rates occur post-reform. This comparison is juxtaposed with the corresponding differential in the pre-reform period. We show three main results: we first document novel stylized facts on the labor market behavior of women in the pre-birth year. Expecting mothers reduce their earnings in the pre-birth year, especially at the bottom of the income distribution. Second, the introduction of parental leave benefits leads women at the bottom to decrease their pre-birth earnings less and this effect is largely driven by fewer women exiting the labor force in the year preceding childbirth. Third, when descriptively looking at the alternative channel to maximize benefits (i.e. tax planning) we find that a minority of expecting parents exploit this option. The benefit's complex incentive structure is likely to play a role here. These findings suggest that parental leave benefits may have effects beyond the post-birth labor market outcomes of parents that have been studied so far.

Chapter 4, which is joint work with Martin Simmler, focuses on the magnitude of local knowledge spillovers of public R&D in Germany and its determinants. Empirical evidence on the magnitude of local knowledge spillovers of public on firm R&D and its determinants is limited and largely based on universities. Universities engage, however, in both, knowledge as well as "degree" production. We exploit the following unique features of Germany to overcome these challenges and to identify the transmission channels at work. First, Germany relies substantially on public R&D carried out by independent research institutes. Second, the German firm R&D support strategy mainly consisted of funding public R&D during our sample period instead of direct subsidies to firms. Third, German firms have comparably high R&D expenditures.

We use the OECD RegPat and Citation database that covers all patents that are filed with the European Patent Office and under the Patent Co-operation Treaty. Additionally, we employ firm and county level data for Germany. In the first part, we aim to quantify the relative importance of three potential local public knowledge spillover

channels. The first channel is firms' use of specific knowledge that is produced by a local public institution (i.e. public patent). The second channel is the collaboration (or joint ventures) of firms with (local) public institutions. The third channel is the use of non-specific public knowledge by local firms. To analyze the importance of the different spillover channels, we exploit regional variation in public R&D in Germany between 1995 and 2015. We proxy firm R&D with the number of patent applications and technological spillovers by using patent citation data. Collaboration spillovers are proxied by joint patent applications of firms and public institutions. Lastly, we proxy non-specific knowledge spillovers with the number of public patent applications within a region. To assess the relevance of the three channels, we estimate count models of patent applications on the applicant-region level. We find evidence that is consistent with the existence of all three types of spillovers. Our results suggest that the non-specific knowledge spillovers are most important as they account for around 2/3 of the overall local knowledge spillovers of public R&D.

In the second part of the empirical analysis, we aim to quantify the overall magnitude of local knowledge spillovers of public R&D on local firm R&D within a county. We proxy public and firm R&D with the number of patent applications and use an instrumental variable strategy to account for a potential bias in the Ordinary Least Squares estimator due to measurement error, omitted variables and reverse causality. The excluded instrument we employ is the 4-year lagged institutional funding for research institutes. We find evidence for substantial local knowledge spillovers of public R&D. For the median county, our estimates imply that one additional public patent generates around 3 additional firm patents. These are driven by non-specific knowledge spillovers. Our results suggest that public R&D leads to more firm patent applications at the same public costs than using R&D tax credits to stimulate firm R&D.

Chapter 1

Income Taxation of Couples and Gender (In)equality

1.1 Introduction

Promoting gender equality between men and women and ensuring the economic participation of women are key policy objectives for many governments and international organizations (see, e.g. the Sustainable Development Goals by the UN, the UN Declaration of Human Rights, the Treaty on European Union or the G20 Leaders' commitments). Despite these efforts, gender gaps in pay and employment persist globally (see, e.g. Bertrand 2020; Blau and Kahn 2003; Olivetti and Petrongolo 2016). Income taxation can affect the labor supply of both spouses differentially through joint taxation. In progressive tax systems, joint taxation equalizes the marginal tax rates of spouses, resulting in a higher marginal tax rate for the secondary earner, typically the wife, compared to a single with the same income. Consequently, joint taxation is widely perceived as impeding gender equality by discouraging the labor supply of married women (Bick and Fuchs-Schündeln 2017). Spouses are taxed jointly in various countries, including the US, France, and Germany.[1] However, high-quality data and causal evidence of the impact of joint taxation on spouses' labor supply is scarce due to the lack of natural experiments that exogenously shift the tax burden among spouses. This paper takes a significant step towards addressing this gap.

In this paper, I exploit three specific features of the income tax in Germany to study newlywed spouses' labor supply responses to a move to individual taxation. First, married couples can freely choose between individual and joint income taxation for both their withholding income tax and final income tax. Second, the 2013 administrative reform to the German withholding income tax introduced the individual withholding tax schedule as the default for newlyweds. Third, I leverage newly available data for the universe of German taxpayers filing tax returns.

I show three main results: first, if a default is available, the likelihood that newlywed couples, for which the primary earner is male and the secondary earner is female, will opt for the default of individual withholding taxation, increases from 57% to 70%. Thus, the default matters. Second, the reform affects women's labor supply. Specifically, female labor income before taxes for secondary earners significantly increases by 3% in the initial years of marriage, while the labor income before taxes of male primary earners remains unchanged compared to the period before the default introduction. Consequently, the partner pay gap, representing the difference between the labor incomes of primary and secondary earner, is reduced by 2 percentage points. Third, female labor force participation increases by 1 percentage point among secondary earners. Therefore, gender equality in earnings improves and overall labor supply is raised.

[1] See Table 1.A.1 for an overview of joint taxation systems around the world.

In Germany, as in many other countries worldwide, income from employment is subject to automatic income tax withholding. This system involves deducting estimated income tax payments directly from an employee's monthly salary, which are then offset against their final income tax liability when they, or they and their spouse, submit their tax returns. Married couples have the option to select from two different tax withholding schemes: individual or joint.[2] The individual scheme taxes partners separately, while the joint scheme allows the primary earner, typically the husband, to shift part of the tax burden to the secondary earner, typically the wife. Thus, under the joint schedule, the average and marginal withholding tax rates for the secondary earner are higher relative to the individual schedule. The opposite is true for the primary earner. The difference in marginal tax rates between these two schemes can reach up to 25 percentage points for the secondary earner.

In the context of studying spouses' labor supply responses to joint or indiviudal taxation, the German withholding income tax schedules provide a unique setting compared to previous studies that analyze the introduction or abolition of joint income taxation. The key difference in my setting is that the choice between individual and joint withholding tax schemes does not affect a couple's overall annual tax liability. Instead, it only impacts the distribution of tax payments throughout the year and how much tax is withheld from each spouse's income individually. Consequently, the decision to marry, motivated by potential tax savings, is not influenced by the choice of withholding tax scheme and differential selection into marriage pre- and post-reform should not be an issue. As a result, this study can isolate the effects of tax rate changes on labor supply without having to account for varying marriage incentives due to tax benefits.

I exploit the 2013 reform as an exogenous shock to the choice of the withholding income tax schedule for newlyweds. Before the reform, newlyweds had to actively choose their withholding tax schedule – individual or joint schedule – right after being married. The reform introduced a default of the individual tax schedule: newlyweds are now automatically assigned to the individual schedule. Thus, the secondary earner, faces lower average and marginal tax rates and earns a higher share of the couple's intra-year net income if couples stick to the default. The opposite is true for the primary earner.

The responses to the reform can be interpreted as a move to individual taxation if the following conditions are fulfilled. First, defaults matter (see, e.g. Madrian and Shea 2001). Second, withholding taxes can affect spouses' labor supply in a similar manner as final taxes. I provide evidence for both conditions.

I am able to analyze novel and unique administrative microdata from German tax authorities. I have access to the universe of German tax units filing tax returns (Taxpayer panel). A tax unit is either an individual or a married couple filing jointly. It is available from 2001 to 2018. This allows me to link different-sex spouses and track both spouses once they are married.[3] The original dataset does not allow to observe women before being married. Herold and Wallossek (2023) and I link women pre and post

[2]This also applies to the final income tax. However, whereas almost all married couples choose the joint schedule for their final income tax, the choice for the withholding income tax has no clear majority. Thus, I from now on refer to the withholding tax when mentioning the choice between the individual and joint schedule.

[3]Same-sex couples were only allowed to file jointly in 2013 (Herold, Koch, and Neisser 2024). Thus, I focus on different-sex couples.

marriage ourselves by using a unique identifier that is only available since 2010. To our knowledge, we are the first to link German income tax data for spouses before the joint filing period. The tax return data provides precise information on individual earnings and taxes paid. I focus on a sample of dual earner newlyweds who both report positive labor income and no income from self-employment in the pre-marriage year, so that the withholding income tax is relevant.

First, I investigate the impact of the default introduction on the choice of withholding income tax schedules for newlyweds. Specifically, I explore whether married couples opt for the default of the individual schedule. Employing a stacked Difference-in-Differences approach, I find that the share of male primary earner newlyweds choosing the individual schedule increases from 57% to 70% after the default introduction. The effect, although less than half in size, remains highly significant for couples with a female primary earner and those with equal earners. In a heterogeneity analysis, I demonstrate that the impact of the default varies significantly based on specific characteristics, with a more pronounced effect observed for newlyweds living in West Germany and those with children. These findings provide evidence that defaults affect individuals' responses to taxation.

I then turn to my primary outcome of interest, examining the labor income responses of both spouses to the change in marginal withholding tax rates resulting from the move from joint to individual tax schedules. Using newlywed cohorts before the default introduction as the control group and those after as the treatment group, I employ marriage as the event. The findings indicate that secondary-earning newlywed women (constituting approximately 75% of newlywed women) experience a 3% increase in their labor income in the first year following the default introduction, with no discernible effect for newlywed men. This effect for women persists for two years post-marriage, gradually diminishing thereafter. Notably, the effects are more substantial and persistent when the sample is restricted to newlywed women with children. Additionally, I explore whether the default introduction influenced also the extensive margin of labor supply. I find an increase in the female labor force participation of 1 percentage point, while male labor force participation remains unchanged. Therefore, there is an overall increase in labor supply. Furthermore, I estimate a notable 2 percentage point reduction in the partner pay gap after the default introduction, demonstrating how changes in marginal withholding tax rates can enhance gender equality by equalizing spouses' labor income. Finally, I estimate an elasticity of labor income with respect to the net-of-withholding-tax rate of 1.7 for newlywed women following the standard approach to estimate the elasticity of taxable income. These results document a substantial response to the reform and provide evidence that not only final income taxes but also withholding income taxes do matter for labor supply.

Related literature This paper makes several contributions to existing research. The first strand of literature it contributes to is on labor supply responses of married women to family income tax reforms. The lack of (more recent) natural experiments limits the quasi-experimental literature in this field. While LaLumia (2008) and Kalíšková (2014) study the introduction of joint taxation in some US states in 1948 and the Czech Republic in 2005, respectively, Selin (2014) and Roantree (2023) evaluate the introduction of individual taxation for spouses in Sweden in the early 1970s and the

UK in 1990 on the extensive margin of labor supply. These previous studies have documented a significant decline (increase) in the labor force participation of married women in response to the introduction of joint (individual) taxation. Herold, Koch, and Neisser (2024) and Isaac (2023) analyze labor supply responses to the introduction of joint taxation for same-sex couples in Germany and the US in 2013, respectively. Herold, Koch, and Neisser (2024) examine the intensive margin of labor supply and find a substantial decrease in the secondary earner's income and a widening of the partner pay gap.[4]

I contribute to this literature in various dimensions. First, I leverage annual administrative panel data of the universe of German taxpayers. This rich and large dataset enables me to study the dynamics of the treatment effect and explore heterogeneity across numerous sub samples. Second, I exploit a recent reform that, while leaving the final tax burden for married couples unchanged, alters only the individual withholding tax burden for spouses. Third, I study the intensive margin of labor supply in addition to the extensive margin. The female labor force participation share has substantially increased in OECD countries since the earlier sample periods of the related studies. However, the majority of women still only work part-time, especially after childbirths. Hence, my setting allows me to uncover whether features of the tax code are relevant causes for these trends. Fourth, it is possible to link spouses with the dataset, allowing for an exploration of the impact on gender equality in earnings, specifically the partner pay gap. Overall, my findings align with previous literature, indicating that female labor force participation tends to increase after a decline in marginal tax rates. In contrast to the earlier literature, I additionally find effects on the intensive margin of female labor supply, proxied by labor income responses.

The second strand of literature my research relates to is the literature on labor supply responses to withholding taxes.[5] Becker, Fooken, and Steinhoff (2019) study this question using laboratory experiments. They find that individuals motivated by monetary incentives reduce their work effort in response to withholding taxes and increase their work effort after receiving tax refunds. Hence, withholding obscures tax incentives. This finding is supported by Bayer, Simon, and Wegmann (2023) who, through a survey, reveal that less than 20% of German married individuals understand that withholding income taxes serve as a prepayment for the final income tax. Using a reform that implied lower withholding taxes for some married women than for others, they estimate with a 5% sample of German taxpayers that women adjust their labor supply. By exploiting a reform with more substantial marginal withholding income tax varia-

[4]It is essential to note that beyond the quasi-experimental studies, a series of papers uses structural life cycle models to assess the effect of joint taxation on married women's labor supply (see Bick and Fuchs-Schündeln 2017, 2018; Borella, De Nardi, and Yang 2023; Guner, Kaygusuz, and Ventura 2012). For instance, Borella, De Nardi, and Yang (2023) simulate that eliminating joint income taxation and Social Security spousal and survivor benefits would significantly enhance married women's labor market participation in the US.

[5]Jones (2012) and Buettner, Erbe, and Grimm (2019) analyze behavioral channels of withholding taxes. Jones (2012) reveals that tax filers only partially adjust their withholding taxes following an external change in the level of an individual's withholdings relative to the tax liability. He attributes this overwithholding to inertia. Buettner, Erbe, and Grimm (2019) find that married couples with a female primary earner are less likely to choose the joint tax class compared to couples where the husband is the primary earner, resulting in higher marginal withholding taxes for female primary earners than male primary earners on average.

tion, I explore labor income responses for both primary and secondary earners within the universe of German taxpayers. In line with Becker, Fooken, and Steinhoff (2019) and Bayer, Simon, and Wegmann (2023), I find significant labor supply responses for newlywed women following a change in the withholding income tax. In addition, my analysis introduces a novel outcome: gender inequality in earnings.

The third strand of literature this paper speaks to is a large empirical literature that explores defaults. Defaults have been shown to matter for a variety of economic decisions, including retirement savings (Madrian and Shea 2001), charitable donations (Altmann et al. 2019), organ donation (Johnson and Goldstein 2003) and overdraft alerts (Adams et al. 2020). The existing literature illustrates that individuals can be nudged to certain choices, and once set, individuals tend to stick to defaults. I contribute to this literature by presenting evidence that defaults play a role in shaping individuals' responses to taxation. This is reflected in the increased likelihood of newlyweds choosing the individual tax schedule following the introduction of the default.

Finally, my research contributes to a strand of literature in behavioral public economics that explores the role of tax complexity and misperception. Existing studies establish that tax complexity changes individuals' responses to labor market incentives. Abeler and Jäger (2015) show that higher tax complexity results in a misunderstanding of incentives, leading to changes in experimental effort provision. Individuals may not know or may misunderstand the true incentive environment (Chetty, Friedman, and Saez 2013; Chetty and Saez 2013). Moreover, individuals tend to stick to salient features of a tax system (Chetty, Looney, and Kroft 2009; Finkelstein 2009). Notably, Blaufus et al. (2015) find that the perceived income tax rate significantly deviates from the objective tax rate for a majority of taxpayers in Germany, supported by survey evidence in Bayer, Simon, and Wegmann (2023). Hence, individuals may perceive their net-of-withholding tax rate as their final net-of-tax rate. I provide evidence that, despite that withholding income taxes should not impact labor supply responses under certain assumptions, spouses do adjust their labor income. Tax misperception, coupled with the potential rejection of the hypothesis of spouses' income pooling, offers a plausible explanation for this phenomenon.

The remainder of this paper proceeds as follows. Section 1.2 provides details on the institutional setting. Section 1.3 describes the employed data. Section 1.4 explains the empirical strategy to estimate the effect on the choice of the individual withholding tax schedule and documents its existence and magnitude. Section 1.5 outlines the empirical strategy to estimate the labor income responses and provides evidence on its impact. Finally, section 1.6 concludes.

1.2 Institutional Setting

Joint taxation The default for filing taxes for married couples in Germany is joint taxation. However, spouses have the option to choose individual taxation. This is the case in various other countries including the US and several European countries as presented in Table 1.A.1. Filing jointly in Germany implies benefits from income splitting. The taxable income of both spouses is first added and divided by two. The two halves are then taxed according to the tax schedule for each spouse. Due to the progressive

income tax schedule, this always implies a lower total tax burden if spouses' incomes differ. The splitting effect is greater the larger the difference of taxable incomes between spouses and the larger total household income (see Figure 1.A.1).[6] Joint taxation implies a lower marginal tax rate for the primary earner and a higher marginal tax rate for the secondary earner if spouses incomes differ, compared to individual taxation. The vast majority of married couples file jointly in Germany.

Withholding income taxation Germany has automatic income tax withholding for income from employment as numerous developed and developing countries. Expected income tax payments are withheld from the employee's salary by the employer and directly transferred to the tax authority on a monthly basis. The withholding income tax is credited against the final income tax once the individual/married couple files their tax return in the following year. Married couples have - similarly to the final income tax - also the option to be taxed individually or jointly for the withholding income tax.[7] The decision which withholding income tax schedule to choose is independent of the decision to file final income taxes jointly or individually. The withholding tax a taxpayer faces is determined by the so-called tax class she or he is assigned to. Spouses can choose either the individual tax schedule by assigning both spouses the so-called tax class IV or the joint tax schedule by assigning one spouse (typically the primary earner) tax class III and the other spouse (typically the secondary earner) tax class V.[8] Statistics show a large gender disparity in tax classes. The vast majority (79%) of taxpayers in tax class III are men, whereas almost 90% of taxpayers in tax class V are women (see Figure 1.A.2). The basic tax allowance from the spouse in tax class V is shifted to the spouse in tax class III (see Table 1.A.2 for more details). This leads to higher average and marginal tax rates for the spouse in tax class V (hence typically the wife) resulting in a higher individual tax burden relative to the individual tax schedule. The spouse in tax class III (hence typically the husband) faces lower average and marginal tax rates and thus a smaller monthly tax burden (see Figure 1.1a and 1.1b).[9] Therefore, the joint tax schedule can have significant consequences for the individual spouse and especially women. By reducing the net income of the secondary earner and increasing the net income of the primary earner the joint tax schedule reduces the secondary earner's income share and hence the secondary earner's bargaining power.

While the effective final income tax burden of the couple is not affected by the choice of the withholding tax schedule, the chosen withholding tax schedule affects the intra-year timing of the tax payments and the distribution of the individual tax burden within the

[6] In contrast to Germany, the US or Switzerland's joint taxation systems might imply a marriage penalty (higher taxes for married individuals compared to unmarried individuals with the same income).

[7] The monthly tax burden should be aligned with the final income tax burden so that spouses are able to already benefit from income splitting during the year (Bundestag 1974).

[8] Theoretically, a third option exists for married couples - the so-called factor method. The withholding tax is based on the individual tax schedule but labor income is multiplied with an additional factor that already takes income splitting into account. However, less than 1% of married couples use this option (Bundestag 2019). Therefore, I do not consider this option.

[9] Figure 1.1a and 1.1b illustrate the withholding tax rates for the pre-reform year of 2012. Figure 1.A.3a and Figure 1.A.3b display tax rates for the post-reform year of 2013 and demonstrate that tax rates changed only marginally.

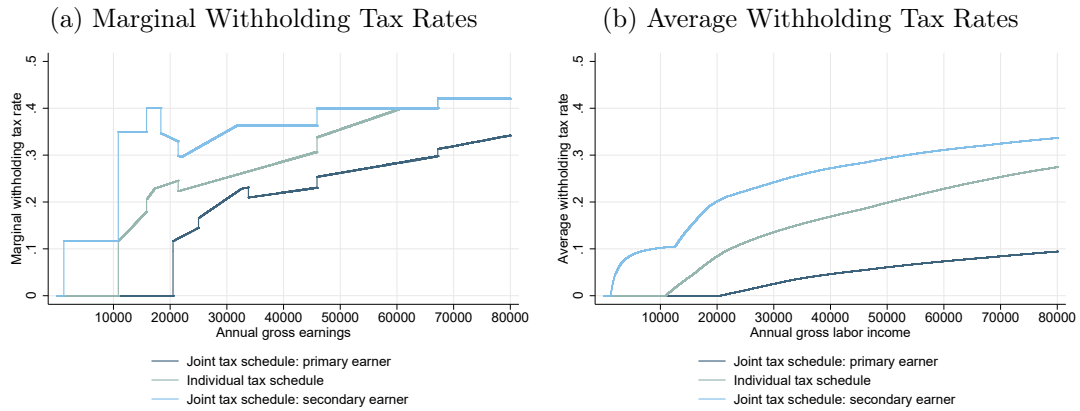


Figure 1.1: Withholding Income Tax Rates

Notes: The joint withholding tax schedule is based on the assumption that the primary earner contributes 60% to the couple's total labor income, while the secondary earner contributes 40%.
Source: Program flow chart by German Federal Ministry of Finance, 2012.

couple. In theory, withholding rates should not matter for the labor supply decision as long as households are not liquidity constrained in the course of the year, spouses pool their income completely and do not misperceive their final income tax as their withholding income tax.

However, studies illustrate that spouses do not pool their income completely (see e.g. Attanasio and Lechene 2014; Beblo and Beninger 2017; Blundell et al. 2007; Giommoni and Rubolino 2022; Lundberg, Pollak, and Wales 1997).[10] Furthermore, I use the German family survey panel *pairfam* to provide evidence that the majority of spouses only pool their income partly. Figure 1.A.4 shows the share of married couples according to a joint or separate bank account. More than 50% of married couples do not choose to have one joint bank account but at least one additional separate bank account. Thus, spouses' salary or parts of it is potentially transferred to the separate bank account.

Moreover, true effective net wages are very complicated to calculate. An individual filing jointly with her spouse will not be able to identify her individual final tax burden based on the tax assessment received from the tax authority. Instead this tax assessment only shows the couple's total final tax burden. Hence, individuals might take their monthly net-of-withholding-tax wage as a reference for labor supply decisions and not their final net labor income. Consequently, the choice of the withholding tax schedule may affect spouses' labor supply and thereby their partner pay gap.

Reform In 2013, an administrative reform of the withholding income tax was implemented. Before the reform, employees' relevant information for the withholding income tax (e.g. tax class, eligibility for child tax allowance) were stored on a paper card and

[10] Beblo and Beninger (2017) find that traditional couples with a clear primary earner are less likely to pool their incomes. Giommoni and Rubolino (2022) show that second earner men do not maximize family income by bunching at a cutoff so that the primary earner women is granted a tax credit - in contrast to second earner women who bunch.

kept with the employer. If circumstances changed, employees had to inform the local city office and let them manually adapt the information on the paper card. This implied that newlyweds had to actively choose their withholding tax schedule - individual or joint - right after being married. The administrative reform in 2013 digitized the process. The information relevant for the withholding income tax are now electronically stored by the tax authority. Although the legislation does not explicitly use the term default, this digitization effectively introduced a default of the individual tax schedule: Newlyweds are now automatically assigned to the individual schedule. If they intend to change their withholding tax schedule, they need to file a letter to the local tax authority.[11] Crucially, the reform did not affect the choice set available to newlyweds: Before and after, newlyweds could choose the individual and joint tax schedule.[12] As mentioned above, the choice of the withholding income tax schedule does not alter the couple's final tax burden and hence, the reform does not change potential tax benefits upon marriage.[13]

I use the reform as an exogenous shock to the choice of the withholding income tax schedule for newlyweds. Consistent with the literature on defaults, newlyweds presumably stick to the default and are hence more likely to choose the individual tax schedule after the reform (Madrian and Shea 2001). This implies that the secondary (primary) earner is more likely to face lower (higher) average and marginal tax rates and earn a higher (lower) share of the couple's intra-year net income if couples stick to the default. Figure 1.1 illustrates that the differences in average and marginal tax rates are substantial: an individual with an annual gross income of e.g. 30,000 € would face a decrease for the marginal tax rate from 35 to 25 percentage points when changing from the joint tax schedule to the individual tax schedule as the secondary earner. Not only are these sizable incentive changes but imply a considerable rise in the household bargaining power of the secondary earner. Figure 1.1b shows that even if individuals perceive their average tax rate as their marginal tax rate, so-called ironing (Liebman and Zeckhauser 2004), these individuals face significant changes in tax rates.[14] These effects are likely to be large enough that individuals had incentives to adjust their labor supply even though it may imply some non-trivial adjustment costs (Chetty 2012). Given that the reform was a purely administrative one, the effective default introduction was an unintended consequence and not designed by the government to affect spouses' withholding income tax. As such there was no public debate about the default introduction.[15]

1.3 Data

To determine newlyweds' choice of the tax schedule as well as labor income responses around the reform I exploit the *Taxpayer Panel* (TPP). It is provided and administered by the German Federal Statistical Office and the statistical offices of the states. This administrative micro data is obtained by combining income tax returns and tax assess-

[11] Already before the reform married couples could change their tax schedule once a year. The new tax schedule would be effective the next month onward. This was unchanged by the reform.

[12] See Table 1.A.3 for detailed information on the reform changes.

[13] Some couples may perceive differential marriage incentives post-reform as they e.g. cannot distinguish between the withholding and final income tax. Nevertheless, I do not find an increase in the marriage rate post-reform (see Figure 1.B.1).

[14] Rees-Jones and Taubinsky (2020) find that around half of the population irons.

[15] Google Trends do not reveal any striking changes around the reform introduction.

ments by the tax authority for the years 2001 to 2018. In addition, the dataset includes employer provided information for non-filers since 2012. The unit of observation is the income taxpayer, either a single individual or a married couple filing jointly. Once a married couple files jointly, one spouse (typically the wife) is appended to the other spouse's (typically the husband) spell. Hence, it is possible to observe both spouses of different-sex couples upon filing jointly.[16] My sample comprises the full population of German taxpayers. Data on the full population of German taxpayers is exclusively accessible upon specific research project inquiries and granted with a restricted set of variables. I have precise information on taxable income (divided in income sources), final and withholding income taxes and basic socio-demographic characteristics, such as gender, marital status, filing status, year of birth, state of residence, religion, year of birth of children, and number of children. The data do not provide information on hours worked, thus I use labor income as a proxy for labor supply. All variables are reported at the annual level.

The original data do not allow to observe single and joint filing spells of the same women. Individual tax IDs were only introduced in Germany in 2010. I exploit the individual tax ID to link spouses before filing jointly (marriage) if they got married in 2011 or later. To the best of my knowledge, together with Herold and Wallossek (2023), we are the first to link German income tax data for spouses before filing jointly. Combined with the full population, this gives me a rich and unique dataset to explore spouses around marriage.

Sample restrictions Since the default introduction of the individual withholding tax schedule in 2013 was only relevant for newlyweds, I restrict my sample to couples for whom I observe both spouses getting married. This is the case for couples marrying in 2011 or later. In addition for the labor income responses analysis, I keep a balanced panel: couples who I observe three years before and after getting married. Since 2018 is the last year in the dataset, I focus on newlyweds marrying between 2011 and 2015 for the main analysis. I include a smaller set of cohorts (2011-2014) and a larger set (2011-2017) for robustness checks.[17] I relax the balanced panel restriction for the "first stage" as I am primarily interested in the choice of the tax schedule in the marriage year. The withholding income tax applies to individuals with positive labor income from dependent employment. Therefore, I only keep newlyweds for whom both spouses earn an annual labor income above 5,400 € in the pre-marriage year to ensure at least some labor force attachment and who do not report income from self-employment in the pre-marriage year.[18] Hence, I measure spouses' intensive margin of labor supply. This is a relevant dimension to explore as female labor force participation has substantially increased over the last decades, however, the majority of women in Germany work part-time, especially after childbirth. I explore the extensive margin of labor supply in an additional analysis. I exclude non-filers since I can only observe them from 2012 onwards and this violates the balanced sample restriction. For the "first stage", I draw sub samples for different couple types: male primary earner couples, female primary

[16] Same-sex couples were only allowed to file jointly in 2013 (Herold, Koch, and Neisser 2024). Thus, I focus on different-sex couples.

[17] I cannot use the balanced panel when including newlyweds marrying between 2011 and 2017.

[18] Individuals could earn 5,400 € per year (450 € per month) without having to pay income taxes and social security contributions on this labor income during the sample period.

	Pre-reform	Post-reform
Female age	33.66	34.54
	(7.82)	(8.77)
Male age	36.65	37.68
	(8.15)	(9.22)
# children	0.55	0.55
	(0.86)	(0.98)
Catholic	0.44	0.43
	(0.50)	(0.50)
West Germany	0.85	0.85
	(0.36)	(0.35)
Female labor income	28,504 €	28,661 €
	(12,742)	(13,175)
Male labor income	45,295 €	45,391 €
	(36,246)	(36,935)
Female income share	0.39	0.39
	(0.10)	(0.11)
N couples	229,822	254,761

Table 1.1: Descriptive Statistics

Notes: The table shows summary statistics (mean values) for the year of marriage for the sample of newlyweds for the pre-reform (2011-2012) and post-reform (2013-2014) period. I restrict the sample to dual earner newlyweds for whom both spouses have positive labor income and no income from self-employment in the pre-marriage year. Income is adjusted by the consumer price index.

earner couples and equal earner couples. Husbands' (wives') income share is equal to or above 60% of the household income for male (female) primary earner couples. Spouses in equal earner couples earn within +/- 10% of their partner's income. I also include estimations in which I define male (female) primary earner couples with a husband's (wife's) income share of equal to or above 55%.

Descriptive statistics Table 1.1 shows the summary statistics for the sample for the pre and post reform years. The characteristics do not differ for the pre and post reform period. Women are 34 years and men 37 years old upon marriage. Couples have approx. 0.55 number of children in the year of marriage. Approx. 44% of newlyweds are catholic and the majority (approx. 85%) live in West German states. Labor income is adjusted by the consumer price index and in 2015 prices. Female newlyweds earn approx. 28,500 € and male newlyweds approx. 45,300 €.

1.4 "First Stage": the Choice of the Tax Schedule

In this section, I show that the introduction of the default had an effect on newlyweds' choice of the withholding income tax schedule: newlyweds became significantly more likely to choose the individual tax schedule after the reform. Thus, the reform represents an exogenous shift in the choice of the withholding income tax schedule which I will

leverage in the subsequent analysis.

1.4.1 Empirical Strategy

I use a stacked Difference-in-Differences approach to determine the causal effect of the default's introduction on newlyweds' choice of the tax schedule (Cengiz et al. 2019; Deshpande and Li 2019). For this I create datasets for each newlywed cohort from 2011-2014. In each dataset, I label couples marrying that year as treated, while couples who have married before 2011 and are married throughout 2011-2018 are the control group. Thus, treated couples never act as controls in other datasets, whereas controls can be controls in several datasets. Finally, I append all four datasets. The resulting dataset has 484,583 newlywed couples. Table 1.B.1 compares characteristics of couples in the treatment and control group pre- and post-reform. Treated and control couples' are similar in the pre- and post-reform period along all dimensions, showing the absence of differential selection of couples into marriage pre- and post-reform or diverging trends for already married couples after the reform.[19]

I regress the following baseline specification:

$$D_{ct}^{individual} = \alpha_c + \gamma_t + \delta Treated_c \cdot Post_t + X_{ct} + \epsilon_{ct} \tag{1.1}$$

where $D_{ct}^{individual}$ denotes an indicator variable equal to 1 if newlywed couple c chooses the individual withholding tax schedule in the year of their marriage t. $Treated_c$ is an indicator equal to 1 if couple c is newlywed. $Post_t$ is an indicator variable equal to 1 after the reform (2013 and later). α_c and γ_t are couple and year-fixed effects, respectively. X_{ct} includes controls that capture newlyweds' characteristics such as wife's age, spouses' age difference, number of children, state of residence, an indicator equal to 1 if one spouse is catholic, an indicator equal to 1 if one spouse has other income (such as rental income, but no income from self-employment). I also control for the share of income substitutes (sum of unemployment benefits, parental leave benefits, sickness benefits, etc. of the secondary earner relative to the individual gross income) since the withholding tax schedule can also affect the amount of future income substitutes. In addition, I also include an indicator variable for additional taxes (assessed final income taxes are higher than withholding income taxes) which implies additional tax payments once the couple has filed their tax return and might therefore also affect the choice of the withholding tax schedule. δ is the coefficient of interest and estimates the effect of the reform, i.e. the introduction of the default, on the choice of the individual withholding tax schedule for newlyweds compared to already married couples. I cluster standard errors on the couple level since this is the level of variation.

The identifying assumption of the Difference-in-Differences model is that, in the absence of the default introduction, the share of newlyweds choosing the individual tax schedule would have evolved similarly relative to the share of already married couples. I validate this in the next section (section 1.4.2). The empirical strategy may be confounding if

[19] Treated couples do not grow older in the post-reform period as is the case for control couples since newlyweds are only in the treatment group in the year of marriage. The number of children decreases for control couples in the post-reform period as only children up to 18 years old are relevant for the tax return if they finished their education. The age of the first child for the control group in the post-reform period shows that first children are on average 20.

selection into marriage changed as a result of the reform, in a way that would have led the choice of the withholding tax schedule after the reform to differ irrespective of the reform. Figure 1.B.1 illustrates that this concern is not valid. The number of marriages is stable across periods. Table 1.1 provides summary statistics on newlyweds' observable characteristics before and after the reform, showing the absence of differential selection of couples into marriage after the reform. Furthermore, it has recently been shown that the two-way fixed effect design with variation in treatment timing may produce biased treatment effects by comparing early treated and late treated cohorts (Callaway and Sant'Anna 2021; De Chaisemartin and d'Haultfoeuille 2020; Sun and Abraham 2021). This is of less relevance for my setting since the pre-reform newlywed cohorts act as never-treated groups. Nevertheless, the stacked Difference-in-Differences approach is one method to overcome heterogeneous treatment effects (Baker, Larcker, and Wang 2022).

1.4.2 Graphical Evidence

Figure 1.2a shows graphical evidence for the effect of the default on the choice of the individual withholding income schedule for newlyweds. The graph shows the unconditional mean of the share of newlywed couples who choose the individual tax schedule in their marriage year (blue dots) and already married couples who choose the individual tax schedule in any year (red triangles). Already married are defined as having married before 2007 and staying married throughout 2017. Hence, newlyweds do not change to the group of already married couples as explained above.

The share of already married couples and the share of newlyweds choosing the individual withholding tax schedule is relatively stable before the default introduction in 2013. Approx. 43% of already married couples choose the individual tax schedule, whereas this is true for approx. 57% of newlyweds. The higher share for newlyweds is arguably driven by smaller income differences among spouses right at the start of marriage. The share of already married couples displays a small positive trend starting pre-reform in 2012, but no substantial change around the reform year, indicated by the dashed line, is visible. In contrast, there is a clear discontinuity at the reform date for newlyweds. The share jumps to 67% in 2013, the reform year, and increases further to 73% over the next four years. This positive trend for newlyweds post-reform could be driven by spillovers - newlyweds communicating with their social networks and updating their knowledge about the process of choosing a withholding income tax schedule. Figure 1.2a also shows that approx. 30% of newlyweds directly opt out of the default in their marriage year and choose the joint withholding income tax schedule.

While Figure 1.2a just presented a discontinuous change in the share of newlyweds choosing the individual tax schedule pre- and post-reform, it could be possible that the discontinuity exists only at marriage but then vanishes in the subsequent years. Nevertheless, Figure 1.2b demonstrates a sharp change in the share of newlyweds choosing the individual tax schedule post-reform (newlywed cohort 2013 and 2014) that persists in the years after marriage. For both pre- and post-reform cohorts the share of couples choosing the individual tax schedule declines over time with roughly the same rate. This is arguably driven by childbirth and couples switching to the joint tax schedule. Yet, three years after marriage the difference in the pre- and post-reform shares is still substantial.

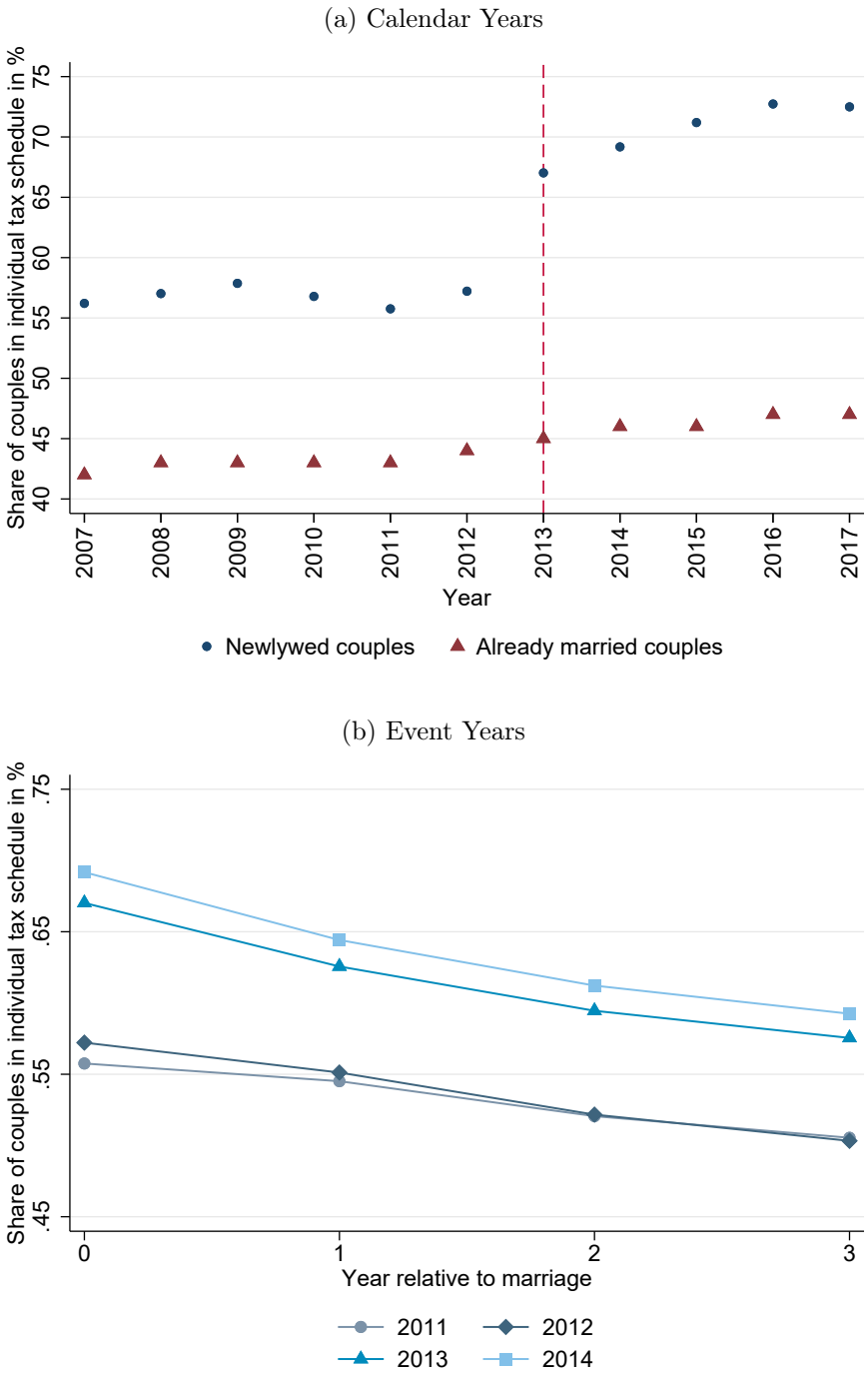


Figure 1.2: Share of Newlyweds in Individual Withholding Tax Schedule over Calendar and Event Years

Notes: Panel (a) shows the unconditional means in the share of the individual withholding tax schedule for newlyweds in the respective marriage year and already married couples who married before 2007 and stay married, complying with sample restrictions. Panel (b) shows the unconditional means in the share of the individual withholding tax schedule for newlywed cohorts 2011 to 2014 in the years after marriage. *t=0* presents the marriage years.

1.4.3 Results

Table 1.2 reports the corresponding results from Equation 1.1 and confirms the graphical evidence from Figure 1.2a: The introduction of the default significantly increases the share of newlyweds choosing the individual withholding income tax schedule by approx. 10 percentage points in the marriage year. Column 1 of Table 1.2 refers to the specification defined in Equation 1.1. Column 2, 3, and 4 are sub samples for specific couple types: newlyweds with a male primary earner, female primary earner or equal earner spouses, respectively. Newlyweds with a male primary earner present approx. 75% of all newlyweds and are thus the vast majority. This sub sample reacts more strongly to the introduction of the default relative to the general sample: the share of male primary earner couples choosing the individual tax schedule increases by 13.6 percentage points after the reform. In contrast, the default has a weaker effect on female primary and equal earner newlyweds. The differences in effect sizes might arise due to the following factors. Female primary and equal earner newlyweds are more gender equal in earnings than male primary earner newlyweds. Consequently, they already choose the individual tax schedule more often before the introduction of the default. In addition, Buettner, Erbe, and Grimm (2019) find that female primary earner couples are less likely to choose the joint tax schedule than male primary earner couples. This is supported by Giommoni and Rubolino (2022) who show that male secondary earners do not maximize family income. The pre-period mean of male primary earner newlyweds' share of the individual tax schedule is with 40.5% much lower than the pre-period mean of female primary and equal earner newlyweds, 62.3% and 83.1%, respectively.

	Entire sample (1)	Male primary earner (2)	Female primary earner (3)	Equal earner (4)
$Treated_c \cdot Post_t$	0.101***	0.136***	0.055***	0.033***
	(0.004)	(0.006)	(0.005)	(0.002)
Controls	x	x	x	x
Pre-period mean	0.585	0.405	0.623	0.831
Observations	5,453,260	3,244,176	409,244	1,799,840

Table 1.2: Effect on Choice of Individual Tax Schedule

Notes: Effect of the default introduction on the choice of the individual withholding taxation schedule in the year of marriage for newlyweds complying with sample restrictions for the newlywed cohorts 2011-2014. Coefficients from the regression specified in Equation 1.1. *,**,*** denote statistical significance at the 5, 1 and 0.1% level.

Newlyweds could actively or passively comply with the default. Jones (2012) shows that tax filers only partially adjust their withholding tax after an external change in the level of an individual's withholdings relative to the tax liability. He finds that individuals offset less than one third of the change in their refund level after one year and less than two thirds after three years. Thus, at least part of the effect I estimate could be due to passive compliers. However, since I am ultimately interested in spouses' labor income responses to their withholding income tax, it is not so relevant whether newlyweds stick

to the default actively or passively.

1.4.4 Heterogeneous Effects of the Default

I explore which observable characteristics relate to a stronger response with respect to the default introduction - which characteristics relate to a higher compliance rate to the default introduction? Understanding the interplay between the default and different demographic characteristics is of great relevance for policy makers, since it allows for understanding the impact of the default for different groups of the population. To account for potentially heterogeneous effects, I interact the newlywed indicator variable $Treated_c$ and the reform indicator variable $Post_t$ with individual characteristics X_c. This gives the following estimation equation:

$$D_{ct}^{ind} = \alpha_c + \gamma_t + \delta Treated_c \cdot Post_t + \mu Treated_c \cdot Post_t \cdot X_{ct} + X_{ct} + \epsilon_{ct} \tag{1.2}$$

For each individual characteristic, the interaction term μ captures heterogeneity in the effect of the default for different values of the control variables. A positive value implies a larger increase in the share of the individual withholding income tax schedule after the default introduction.

Figure 1.3 shows the results for the total effect from Equation 1.2. The results document that the overall effect of the default is persistent when allowing for heterogeneous effects. However, the impact of the default significantly differs for certain characteristics. Catholic newlyweds are significantly less likely to choose the individual withholding income tax schedule, but the effect of the default is significantly larger for them. This is similar for newlyweds living in West German states and newlyweds having children. Couples in West Germany are still more traditional and are more inclined to the male breadwinner norm than couples in East Germany who are shaped by the former socialist GDR. The default has no differential effect on newlyweds whose age difference is larger than three years and the husband is older than on the general sample.[20] I also study different household labor income percentiles. Newlywed couples whose household labor income is below the 50th percentile are less likely to choose the individual tax schedule than the general sample after the reform. In contrast, newlywed couples whose household labor income is above the 90th percentile respond to the default introduction very similarly as the general sample.

There are two potential explanations for the heterogeneous effects of the default. First, newlyweds may face different (dis-)incentives for choosing the individual tax schedule. For example, couples with a household labor income below the 50th percentile may be liquidity constrained and thus choose the joint tax schedule to minimize their tax payments during the year, whereas liquidity is not a concern for high income households. Second, the heterogeneous effects could stem from different levels of salience and newlyweds' understanding of the default. If the default and the option to deviate from the default are less salient for certain newlyweds, then these newlyweds may be more likely to stick to the default. The administrative tax data unfortunately does not provide a direct measure of tax literacy to explore whether different levels of understanding of the

[20] Previous literature has used this characteristic as a proxy for conservatism. The husband potentially has a longer employment history and is thus on a higher wage trajectory than the wife.

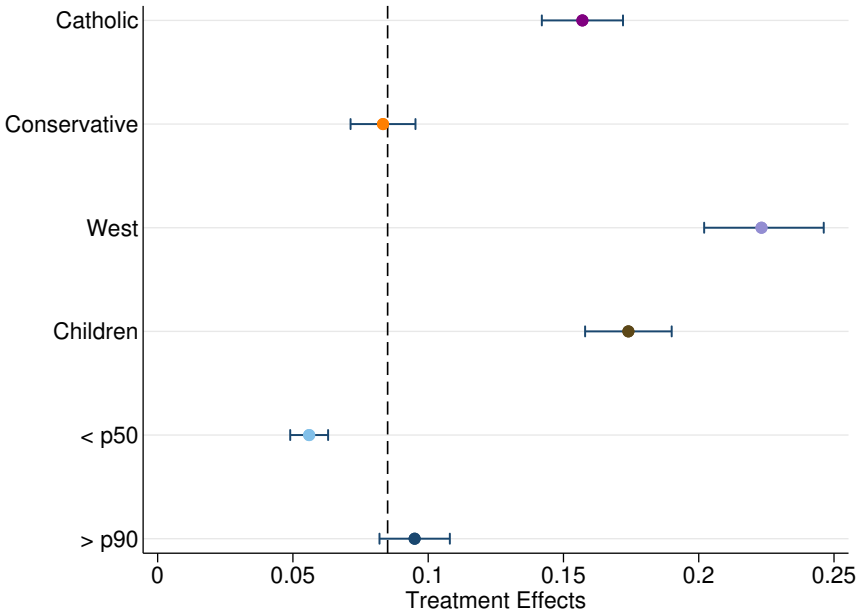


Figure 1.3: Heterogeneous Effects of Default

Notes: The figure presents the point estimates for the total effect when interacting the reform indicator $Post_t$ and $Treated_c$ with individual characteristics X_{ct}. The dashed vertical line indicates the baseline effect when including the interactions with the individual characteristics. Couples are defined as conservative if spouses' age difference is greater than three and the husband is older than the wife. Horizontal segments indicate the 95% confidence interval.

tax system affect the responses. Avenues for future research could entail constructing a measure of tax "sophistication" such as e.g. the usage of tax deductions as a proxy for tax literacy.[21]

1.5 "Reduced Form": Labor Income Responses

In this section I document that the exogenous shift towards the individual tax schedule implied a significant increase in labor income for women, whereas no change can be observed for male labor income.

[21] The German income tax code allows for a vast number of potential deductions once an individual files her tax return. Kirchhof (2011) counts at least 534 potential deduction possibilities. Deductions can be classified into income-related deductions that are necessary expenses to earn income and other deductions that are special expenses. Whereas income-related deductions are third-party reported and automatically deducted from gross income, other deductions are self-reported and imply selection by the individual. These other deductions include e.g. child care costs, own or children's education, donations to charity or political parties and church tax payments. Doerrenberg, Peichl, and Siegloch (2017) find that other deductions make up 87% of total deductions on average and that total deductions account for approx. 20% of gross income. Some individuals might not know about these deduction possibilities and do not exploit them.

1.5.1 Empirical Strategy

Having established that the reform changed newlyweds' choice of the tax schedule, I analyze the effect on spouses' labor income responses in an event-study setting. For this, I focus on male primary earner couples. The reason is twofold. First, couples with a male primary earner and a female secondary earner represent the vast majority of married couples in Germany (approx. 75%). Second, section 1.4 demonstrated that the individual tax schedule already is the preferred option for female primary earner and equal earner couples before the reform and that the effect of the default is much weaker for these couple types. Hence, spouses of these couples experienced much smaller changes in labor supply incentives due to smaller changes in marginal withholding tax rates.[22]

The effective treatment date varies across couples depending on their year of marriage. For example, a newlywed couple who married in 2011 would not be treated, whereas a couple marrying in 2013 would be treated. The empirical design, therefore, compares earlier and later newlywed cohorts. The specification is the following:

$$ln(y_{it}) = \alpha_i + \beta_t + \sum \gamma_k \cdot D_k + \sum \delta_k \cdot D_k \cdot Post_t + X_{it} + \epsilon_{it} \tag{1.3}$$

where y_{it} is labor income or another outcome variable measured at the individual level. D_k are event time indicator variables. X_{it} are time-varying controls as in Equation 1.1. In addition, I include indicator variables for having a child in a certain age group (0-2 years, 3-5 years, 6-9 years, 10-14 years, 15-18 years). α_i are individual-by-marital-status fixed effects and β_t are calendar-year-fixed effects. Standard errors are clustered at the couple level. The key coefficient of interest is δ_k, measuring the change in labor income of newlywed individual i in event year k post-reform relative to a newlywed individual's labor income in event year k marrying in the pre-reform period. I use $t = -1$ as the baseline year.

Since I estimate the reduced form effect, I obtain an intention to treat (ITT) estimate. All newlyweds marrying in 2013 and after are affected by the default introduction. However, not all stay with the default of the individual tax schedule (see Figure 1.2a). Thus, although these couples who immediately opt out after having married are not treated, they are included in the reduced form analysis. The ITT estimate directly evaluates the actual effect of the reform. Therefore, the treatment on the treated (TOT) estimates would be larger.

The underlying assumption of my identification strategy is that absent the reform untreated and treated newlyweds' labor income would have followed the same trend. The following threats to identification could arise. First, the effects I am measuring could just be capturing a time trend in labor income. To check this, I estimate a separate event study for each cohort (see section 1.5.4). Second, selection into marriage could have changed around the reform. Since the purpose of the reform was only to digitize the withholding income tax system, it is unlikely that the reform had an effect on marriage rates and treatment timing is plausibly exogenous (see section 1.4.1). Figure 1.B.1 shows that the marriage rate has been stable around the reform in 2013. Hence,

[22] Descriptive statistics for the male primary earner couple sample is shown in Table 1.B.2.

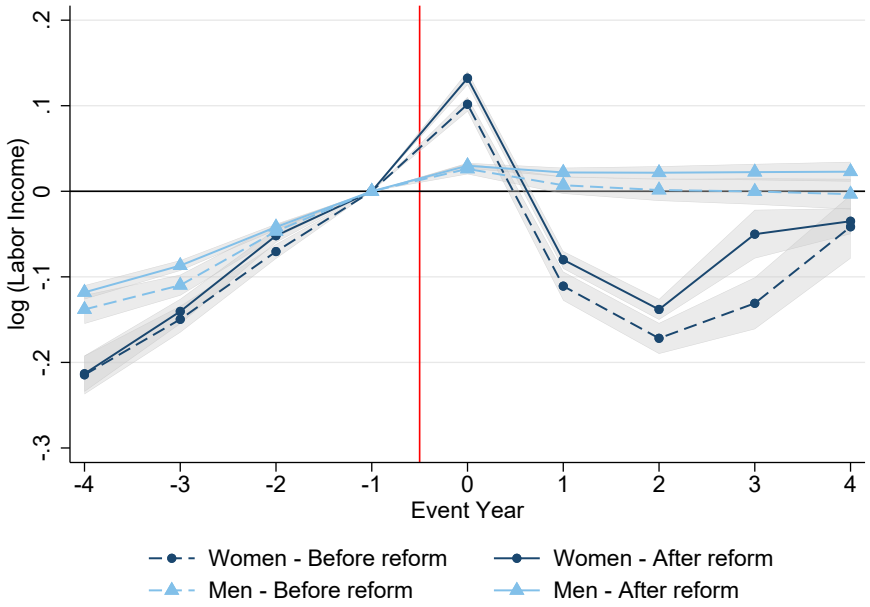


Figure 1.4: Labor Income Relative to Year of Marriage

Notes: This figure shows female and male labor income relative to the year of marriage for male primary earner couples. Plotted estimates are based on a balanced sample. t=0 is the year of marriage. Gray areas are 95% confidence intervals.

selection into marriage does not seem to change around the reform. Thus, early and late cohorts of newlyweds can be meaningfully compared.[23]

1.5.2 Results

Labor income relative to year of marriage Figure 1.4 displays the evolution of female and male labor income relative to the year of marriage for the sample of male primary earner couples. Labor income for newlywed women drops substantially in the first year after marriage and increases again for the subsequent years. This pattern emerges irrespective of the reform. However, the drop in female labor income is significantly smaller in the post reform period. Thus, the introduction of the default seems to affect female labor income. The substantial drop in the female labor income in the year after marriage is consistent with the literature on "child penalties" (see e.g. Kleven, Landais, and Søgaard 2019). In addition, Herold and Wallossek (2023) document a marriage earnings gap even after accounting for the child penalty.[24] Male labor income significantly increases up to the year of marriage and is relatively stable afterwards. In contrast to women, this pattern evolves similarly for the pre- and post-reform period. Therefore, the reform does not seem to affect male labor income.

[23] Since my setting involves a single treatment date, I do not need to assume homogeneous treatment effects as highlighted in the new Difference-in-Differences literature (e.g. Goodman-Bacon 2021).

[24] In contrast to this literature, I do not evaluate an earnings gap by estimating the difference between the individual observed and counterfactual earnings, but use the reform as an exogenous variation.

Net labor income Before analyzing the effect on gross labor income, I evaluate the treatment effect $\hat{\delta}_k$ on female and male net labor income. Given that more newlyweds choose the individual tax schedule in the post-reform period, newlywed women face a lower while newlywed men face a higher withholding tax liability. This should mechanically lead to a higher female net labor income and a lower male net labor income. Figure 1.5a illustrates exactly this. For $t < 0$, the effects are around zero and insignificant. Therefore, newlyweds marrying before and newlyweds marrying after the reform do not follow a differential trend. In $t = 0$, newlywed women significantly increase their net labor income by 3.9% relative to newlywed women before the reform. The treatment effect is persistent for $t = 1$ to $t = 3$, but decreases in magnitude and becomes insignificant for $t = 4$. For men, net labor income significantly decreases by 1% after being married in the post-reform period relative to the pre-reform period.

Gross labor income Figure 1.5b shows the treatment effect $\hat{\delta}_k$ for gross labor income for men and women. The evolution of the effect on labor income closely mirrors that of the net labor income, with no discernible differential trends in the pre-reform period for both women and men. The point estimates for women are slightly smaller compared to those for net labor income. In $t = 0$, newlywed women significantly increase their labor income by 3.6% relative to their counterparts before the reform. This treatment effect remains relatively stable in the early years of marriage and diminishes in magnitude thereafter. It is crucial to note that these effects are likely lower bounds, given that the withholding income tax is used as a proxy for the final income tax.

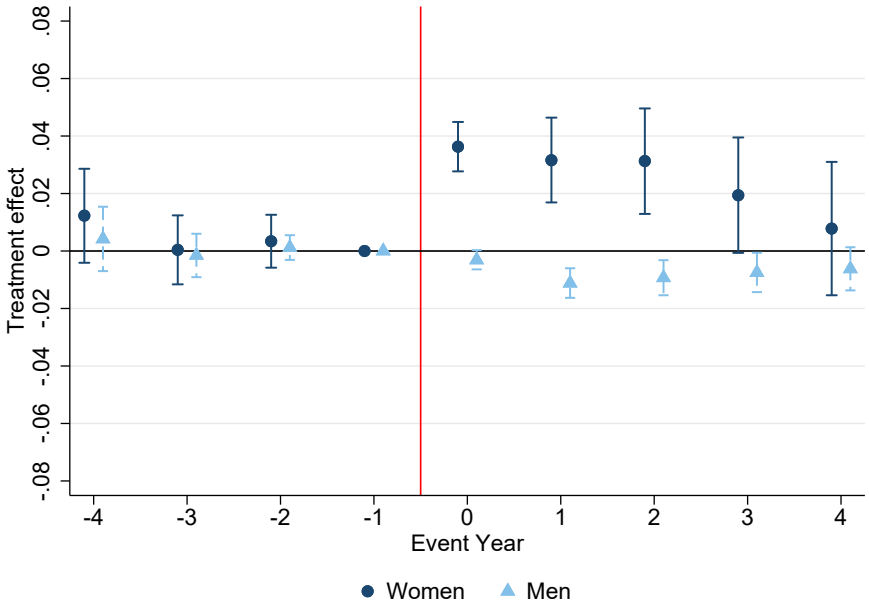
These substantial responses provide evidence that not only final income taxes but also withholding income taxes matter for spouses' labor supply. The large estimated effects may be driven by the facts that spouses do not (entirely) pool their income and that (some) spouses cannot distinguish between withholding and final income taxes. These findings for the intensive margin of female labor supply contrast with previous literature that did not find significant intensive margin responses following the introduction of joint taxation (Kalíšková 2014; LaLumia 2008). Kalíšková (2014) suggests that the small effect on the intensive margin of female labor supply can be attributed to the low availability of jobs offering flexible working hours in the Czech Republic during the early 2000s. However, this constraint is likely to be less pronounced for Western European countries, including Germany.

For men, contrary to the small negative effect for net labor income, the point estimates and confidence intervals are around zero. Hence, the introduction of the default and the resulting change in the marginal withholding tax rate predominantly affected female gross labor income, with no impact on male gross labor income. These findings are in line with prior literature documenting that women's labor supply is more elastic to income taxes than that of men (Blundell and MaCurdy 1999; Keane 2011).[25]

Next, I explore whether newlyweds with children reacted differently. For this analysis, I restrict the sample to newlyweds who have a child below the age of 18 within $t = -4$

[25] I conduct additional robustness checks by including only the 2011 to 2014 newlywed cohorts and employing an alternative definition of male primary earner couples (male primary earner's income share is at least 55% instead of 60%). The results from both robustness checks are very similar and available upon request.

(a) Net Income



(b) Gross Income

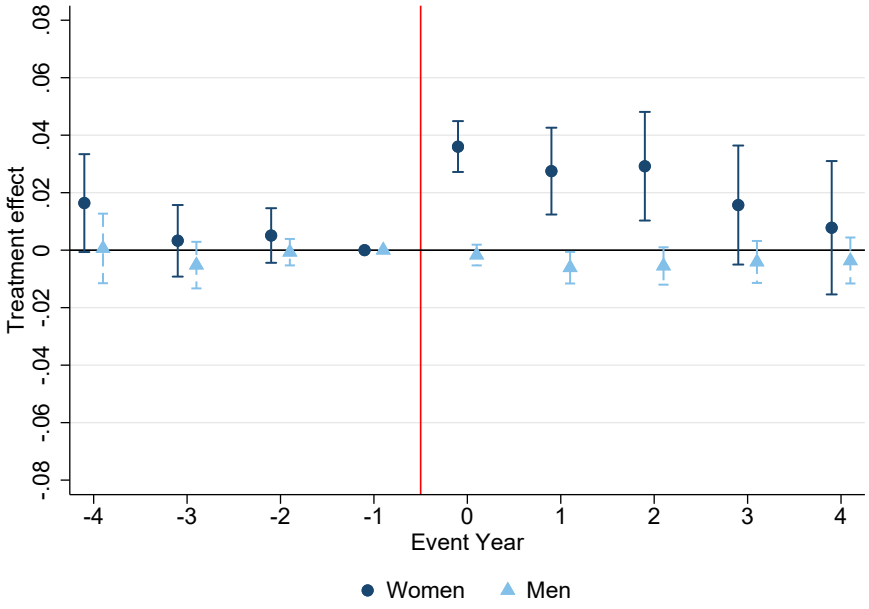


Figure 1.5: Treatment Effect for Labor Income

Notes: Panel (a) shows the treatment effect for female and male net labor income and panel (b) for gross labor income for male primary earner couples. Plotted estimates are based on balanced sample. t=0 is the year of marriage. Vertical segments are 95% confidence intervals.

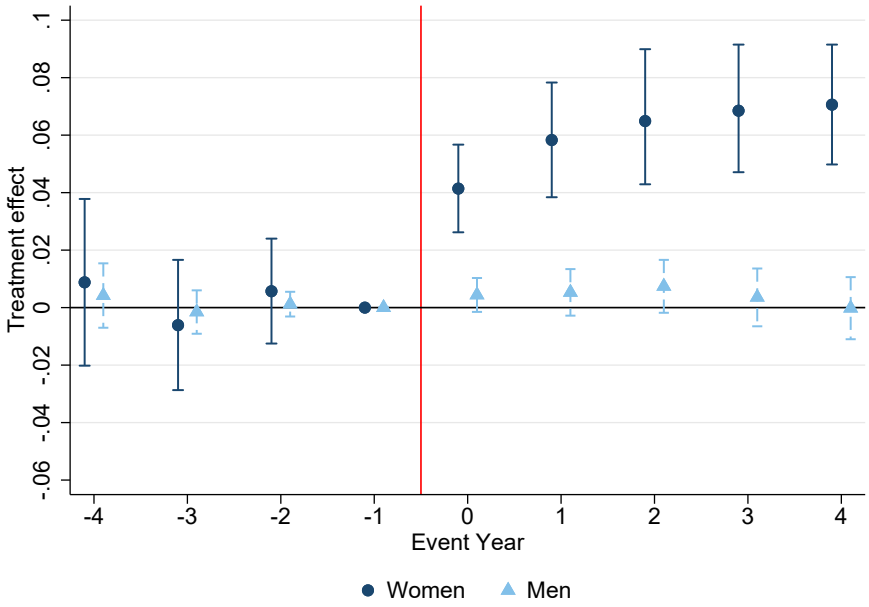


Figure 1.6: Treatment Effect for Labor Income for Couples with Children

Notes: This figure shows the treatment effect for female and male labor income for male primary earner couples with children below the age of 18. Plotted estimates are based on balanced sample. t=0 is the year of marriage. Vertical segments are 95% confidence intervals.

and $t = 4$. In addition, I include the squared age of the child as a control variable.[26] For women, the point estimate in $t = 0$ is similar to the main sample (see Figure 1.6). However, in contrast to the main sample, the effect on female labor income significantly increases and persists in the years after marriage. Female labor income increases by approx. 7% four years after marriage for couples marrying post-reform relative to pre-reform newlyweds. Given that the default introduction had a stronger effect on newlyweds with children (see section 1.4.4), the larger responses for labor income of women with children are not surprising. Furthermore, these results are consistent with the previous literature finding that women with children are more responsive to income taxes (see e.g. Guner, Kaygusuz, and Ventura 2012; LaLumia 2008; Selin 2014). When I restrict the sample to couples without having children within $t = -4$ and $t = 4$, the point estimates are smaller but the pattern is similar to the main sample. Labor income significantly increases for the first years of marriage for post-reform newlyweds relative to pre-reform newlyweds. The change then becomes insignificant in subsequent years.[27] For men with children, the point estimates are positive, but not statistically significant.[28]

[26] For this specification, I cannot include individual fixed effects as these are collinear with the child's age.

[27] Results are available upon request.

[28] This is similar for men without children.

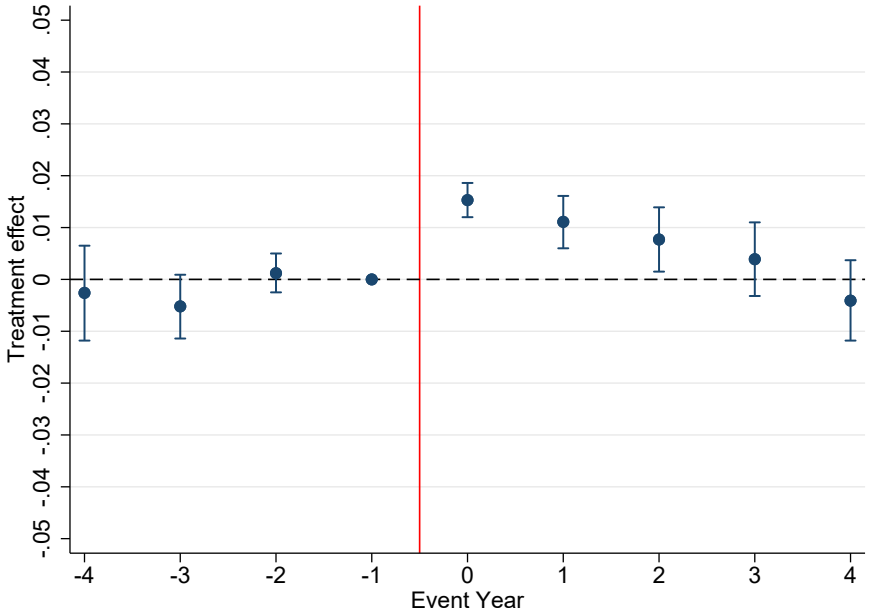


Figure 1.7: Treatment Effect for Household Labor Income

Notes: This figure shows the treatment effect for household labor income for male primary earner couples. Plotted estimates are based on balanced sample. t=0 is the year of marriage. Vertical segments are 95% confidence intervals.

Household labor income Besides husband's and wife's individual labor income, I explore the impact of the default introduction and the implied marginal withholding tax rate change on household labor income. I define household labor income as the sum of spouses' labor income. Figure 1.7 illustrates that household labor income increases significantly by 1.5% in the year of marriage for post-reform newlyweds relative to pre-reform newlyweds. Hence, the increase in female labor income is not offset by a decrease in male labor income. However, similar to the effect on female labor income, the point estimates for household income decrease in size and become insignificant in $t = 3$.

Partner pay gap It is important to evaluate whether the reform improved gender equality in earnings by equalizing spouses' labor income. Previous studies, with the exception of Herold, Koch, and Neisser (2024), have not analyzed this yet. To determine the effect on gender equality in earnings, I study the partner pay gap. I define the partner pay gap as the difference between the husband's and wife's labor income before taxes relative to the total income before taxes of the couple. Figure 1.8 plots the treatment effects for the partner pay gap for couples with a male primary and a female secondary earner. The partner pay gap is reduced by up to 1.9 percentage points post-reform in the first three years of marriage. The effect vanishes for $t = 4$. This finding demonstrates that taxation of couples matters for gender equality in earnings. In this setting, gender equality is improved by the default introduction and the implied marginal withholding tax rate change in the first years of marriage. Complementary,

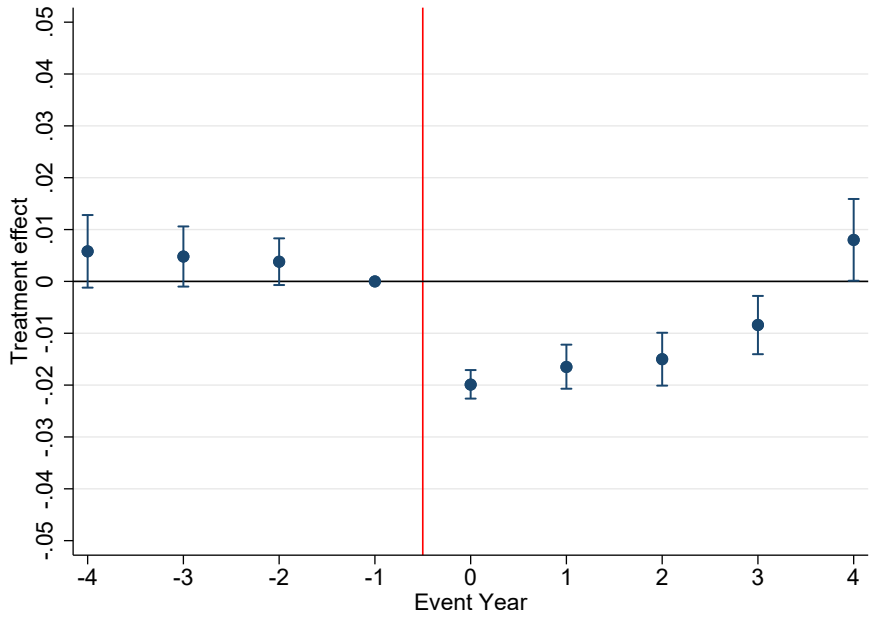


Figure 1.8: Treatment Effect for Partner Pay Gap

Notes: This figure shows the treatment effect for the partner pay gap for male primary earner couples. Partner pay gap is defined as the difference between primary earner's and secondary earner's labor income relative to the total income of the couple. Plotted estimates are based on balanced sample. t=0 is the year of marriage. Vertical segments are 95% confidence intervals.

Herold, Koch, and Neisser (2024) find a widening of the partner pay gap after the introduction of joint taxation for the final income tax for same-sex couples.

Labor force participation So far, I exclusively explored the intensive margin of labor supply of spouses, using changes in labor income as a proxy. Yet, the reform's impact may extend to the extensive margin - the labor force participation of spouses. I use an indicator variable, taking the value of 1 if an individual earns positive labor income in the respective year, to proxy labor force participation. Figure 1.9 illustrates the treatment effects on female and male labor force participation. Newlywed women earning positive labor income in the year of marriage show a significant increase of 1 percentage point in the post-reform period compared to the pre-reform period. This treatment effect further rises to approximately 2 percentage points in the subsequent years. Notably, the effect on the extensive margin is more persistent than the effect on the intensive margin, seen in Figure 1.5b. One plausible explanation for this upward trend in the treatment effect on labor force participation could be that newlywed women are less likely to exit the labor force post-reform compared to the pre-reform period. This aligns with findings in the literature on labor supply responses to joint taxation, which has predominantly focused on the extensive margin (see e.g. Kalíšková 2014; LaLumia 2008; Selin 2014). LaLumia (2008) estimates a 2 percentage points decrease in the employment rate of married women, whereas Kalíšková (2014) finds a slightly larger (3 percentage points) decline for married women with children following the introduction

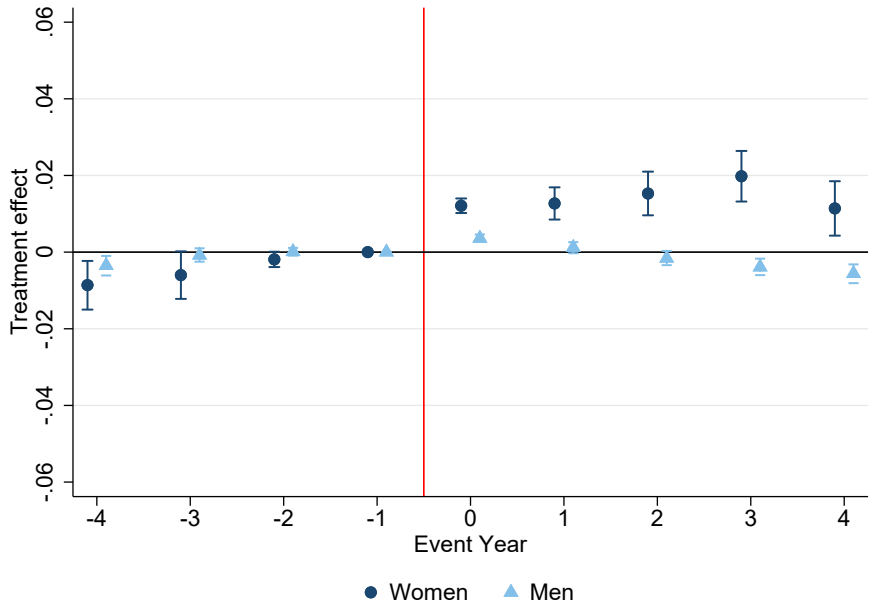


Figure 1.9: Treatment Effect for Labor Force Participation

Notes: This figure shows the treatment effect for female and male labor force participation for male primary earner couples. Labor force participation is proxied with an indicator variable equal to 1 if the individual earns positive labor income. Plotted estimates are based on balanced sample. t=0 is the year of marriage. Vertical segments are 95% confidence intervals.

of joint taxation in the US and the Czech Republic, respectively.

Some women might have left the labor force a few years after marriage in the pre-reform period due to the diminished labor supply incentives implied by the joint tax schedule. The higher take-up of the individual tax schedule post-reform, along with the resulting higher labor supply incentives, might have induced fewer women from leaving the labor force. The treatment effect for male labor force participation is similar to the intensive margin around zero. In $t = 4$, there is a small significant negative effect - the share of newlywed men earning positive labor income decreases by 0.5 percentage points. This small negative effect for men could be driven by the positive treatment effect observed for women. With more women remaining or entering the labor force, some husbands may exit the labor force, no longer serving as the sole breadwinner.

1.5.3 Elasticity

In addition to the event study from Section 1.5.1 that shows how labor supply responses evolve over time, I estimate the elasticity of labor income with respect to the net-of-withholding-tax rate to scale the effects by the change in the net-of-withholding-tax rate. For this, I follow the standard approach to estimate the elasticity of taxable income (ETI). An ETI measures the responsiveness of taxable income to changes in the net-of-tax rate - defined as one minus the MTR. In a progressive tax system, income and the marginal tax rate are jointly determined. To overcome this endogeneity problem, I employ a common approach and construct mechanical tax rate changes. For this, I

apply post-reform tax schedules on pre-reform labor income and instrument for the net-of-withholding-tax rate, as proposed by Gruber and Saez (2002). I follow the standard regression specification and estimate it separately for women and men (s = w,m):

$$ln\left(\frac{y^s_{it}}{y^s_{it-1}}\right) = \varepsilon^s \cdot ln\left(\frac{1-\tau_{it}}{1-\tau_{it-1}}\right) + \delta^s_{inc} + \lambda^s_{share} + \theta^s X_{it-1} + \eta^s_{age} + \mu^s_t + \epsilon^s_{it} \quad (1.4)$$

With the following first stage:

$$ln\left(\frac{1-\tau_{it}}{1-\tau_{it-1}}\right) = \left(\frac{1-\tau^{hyp}_{it-1}}{1-\tau_{it-1}}\right) + \delta^s_{inc} + \lambda^s_{share} + \theta^s X_{it-1} + \eta^s_{age} + \mu^s_t + \epsilon^s_{it} \quad (1.5)$$

y^s_{it} is the gross labor income of individual i in marriage year t. $t-1$ denotes the base-year, i.e. the pre-marriage year. $1-\tau_{it}$ is the net-of-withholding-tax rate for i for the marriage year t, and $1-\tau_{it-1}$ for the pre-marriage year. ε is the parameter of interest. $1-\tau^{hyp}_{it-1}$ presents the hypothetical net-of-withholding-tax rate i would have faced under the withholding tax schedule from $t-1$. δ^s_{inc} denotes income-bin-fixed effects for individual gross labor income bins (1000 €-bins) in $t-1$ to capture non-tax related income trends (e.g. reversion to the mean and heterogeneous income growth across the earnings distribution). I follow the most standard approach and control for the pre-reform income (Auten and Carroll 1999). λ^s_{share} are income-share-fixed effects for individual i's income share of the total couple income (2%-bins). X_{it-1} is a vector of control variables that are the same as in Equation 1.3. η^s_{age} and μ^s_t are a set of age and year-fixed effects.

The 2 stage least squared elasticity (2SLS) estimates from Equation 1.4 for women are presented in Table 1.3. Column (1) does not include pre-marriage income or income-share-fixed effects. Column (2) and (3) include pre-marriage income and income-share-fixed effects, respectively and column (4) includes both fixed effects. The estimates are relatively similar and significant for all four specifications, and range between 1.6 and 1.7. These estimates are comparatively large but in line with the literature documenting elasticities of taxable income with respect to the net-of-tax rate for married women (see e.g. Hermle and Peichl 2018; Neisser 2021). Given that I study newlywed women who are arguably more responsive compared to older women, the larger estimate is reasonable. Additionally, I study more recent income responses than those explored in the previous literature, reflecting the increased flexibility of the labor market. This might also explain the larger estimates in my setting. Notably, Herold and Wallossek (2023) estimate elasticities with respect to the net-of-(final)-tax rate that are roughly similar in size for women in Germany in the same time period I am studying. The elasticity estimates for men are presented in Table 1.C.1. The 2SLS results are statistically significant and negative but economically very small. Therefore, changes in the marginal withholding tax affect the income of newlywed women, while the impact on men's income is minimal, consistent with the literature.

1.5.4 Robustness Checks

I employ the following robustness checks to assess whether the estimated labor income responses are driven by other factors than the default introduction and the implied change in the marginal withholding tax rate.

	Baseline specification	Control for		
		income	income share	both
	(1)	(2)	(3)	(4)
Second stage	1.623***	1.676***	1.640***	1.685***
	(0.051)	(0.053)	(0.052)	(0.053)
First stage	0.555***	0.582***	0.582***	0.581***
	(0.005)	(0.005)	(0.005)	(0.005)
Initial income FE		x		x
Initial income share FE			x	x
Age FE	x	x	x	x
Year FE	x	x	x	x
Controls	x	x	x	x
Observations	404,188	404,188	404,188	404,188

Table 1.3: Elasticity Estimates for Women

Notes: Estimates for the elasticity of gross income with respect to the withholding net-of-tax rate from the 2SLS estimation of Equation 1.4 and 1.5 for women. Column (1) does not include pre-marriage income and income-share-fixed effects (FE). Column (2) includes pre-marriage income-fixed effects and column (3) includes pre-marriage income-share-fixed effects. Column (4) includes both pre-marriage income and income-share-fixed effects. All columns include age and year fixed effects and other control variables. *,**,*** denote statistical significance at the 5, 1 and 0.1% level.

Cohort event study While the results just presented in section 1.5.2 show that the labor income responses for women are systemically different between newlyweds before and after the 2013 reform, they could just be capturing a time trend. I estimate a separate event study for each female newlywed cohort to check whether there is a discontinuous change around the reform. The specification is the following:

$$log(y_{it}) = \alpha_i + \beta_t + \sum \sum \gamma_k \cdot D_k \cdot D_{cohort} + X_{it} + \epsilon_{it} \tag{1.6}$$

All variables are defined as in Equation 1.3. D_{cohort} are cohort indicator variables for each newlywed cohort. Standard errors are again clustered at the couple level. Figure 1.D.1 in the appendix illustrates female labor income around marriage for newlywed cohorts 2011 to 2014. Female labor income is higher for post-reform newlywed cohorts than for pre-reform newlywed cohorts. However, two factors complicate the identification of a clear discontinuous change around the reform. Firstly, the share of newlyweds choosing the individual tax schedule continued to rise after the reform (as depicted in Figure 1.2a). Consequently, we would expect higher labor income responses for the 2014 newlywed cohort. Secondly, the majority of couples marry in the latter half of the year, as illustrated for the reform year 2013 in Figure 1.D.2. Hence, the majority of couples marrying in 2013 had limited time to react during that year and possibly respond in the subsequent year.

Large roll out of daycare centers Several policy reforms impacting the availability of all-day daycare slots for children under the age of three have been implemented in Germany since the mid-2000s (Spieß 2011). These reforms resulted in a significant increase in publicly subsidized daycare slots in both West and East Germany. However, the expansions in East and West Germany varied in terms of magnitude and initial levels. In West Germany, the coverage for children under three years old rose from approximately 12% in 2008 to 29% in 2018, while in East Germany, it increased from about 43% to 55% during the same period (Destatis 2019). Starting from the mid-2000s, West Germany witnessed a sharp rise that leveled off from 2014 onward. Moreover, the expansion led to substantial regional disparities in the speed of growth, which differed significantly across counties (Barschkett 2022). As the expansion coincides with my sample period, the female labor income responses could be driven by the roll out of daycare centers. The dataset does not allow to identify individuals' region, but only provides the state of residence. Therefore, I control for the expansion of daycare centers on the state level. For this, I obtain data for the change in the daycare coverage for each state for the sample period of 2011-2015 from the German Federal Statistical Office and include them as control variables (see Table 1.D.1 in the Appendix). Daycare coverage is defined as the share of children being in daycare, entailing daycare centers and childminders. Table 1.D.1 shows that there already was substantial variation among states, and not only among East and West Germany. East German states had a higher initial level and thus a smaller rise. The labor income and labor force participation responses are very similar when controlling for the daycare expansion (see Figure 1.D.3 and 1.D.4). This confirms that the responses are not driven by the daycare expansion.

Parental leave benefits Since 2007, Germany offers paid parental leave to parents during the first year following their child's birth. The benefits are proportional to the pre-birth net income, with a cap on benefits for those earning above a certain threshold. This benefit design creates incentives for individuals below the cap to increase the pre-birth net income in order to maximize parental benefits post-birth. The net income can be increased by changing the withholding tax schedule. A reform in 2015 allowed for part-time work while receiving parental leave benefits. Therefore, the estimated labor income responses could be partly driven by newlyweds planning to maximize their benefits and not by the responses to the change in marginal withholding income tax rates. To address this, I employ two robustness checks. Firstly, restricting the analysis to post-reform years 2013 and 2014 (excluding the parental benefit reform year) yields results consistent with the main findings.[29] Secondly, I focus on newlyweds above the benefit cap (which relates to an annual labor income of approx. 60,000 €), who do not have an incentive to maximize their net income. Some challenges arise due to the small number of women in this income bracket and the uniform marginal withholding income tax rates for joint and individual schedules above 60,000 € (see Figure 1.1a). Hence, individuals earning above the cap do not face differential labor supply incentives pre- and post-reform and should not react to the introduction of the default in 2013. Estimates for female newlyweds from this sub sample show very large confidence intervals due to the small sample size and indicate no significant labor income responses (see Figure 1.D.5). Glogowsky et al. (2024) support this by finding

[29] Results are available upon request.

minimal changes in withholding tax schedules by expecting parents seeking benefit maximization. Collectively, this suggests that observed female labor income responses are not influenced by the incentives stemming from the parental leave benefits.

1.6 Conclusion

In this paper, I have studied the role of couple's taxation on the labor supply of newlywed spouses. By leveraging the introduction of a default of individual withholding taxation in a regime of joint withholding taxation in Germany, I establish three main facts. First, the share of newlyweds choosing the individual tax schedule increases from 57% to 70% percent for primary earner couples after the default introduction and thus female secondary earner face lower average and marginal tax rates. Second, female labor income increases by 3% on average in the first years of marriage post reform for female secondary earners, while male labor income remains unchanged. The effects are larger and more persistent when restricting the sample to only newlyweds having children within the sample period. Consequently, the partner pay gap narrows by 2 percentage points. Finally, female labor force participation persistently increases by 1 percentage point. Therefore, the reform raises overall labor supply. These results not only show that institutional settings (such as the withholding rate or defaults) of a tax system matter for labor supply responses, but additionally imply that joint taxation of spouses can harm gender equality in earnings.

Avenues for future research The narrowing partner pay gap and the implied rise in the female income share likely strengthened women's household bargaining power. Subsequent research should delve into this aspect further.[30] Moreover, I focus on short-term labor income responses. However, income taxes also affect the incentives on human capital accumulation, fertility and savings decisions. Roantree (2023) develops a dynamic life cycle model. Simulations suggest that abolishing joint taxation of couples can increase women's employment and earnings in the long-run and lead to women opting to have children at a later stage in life. Hence, an exploration of long-term labor market outcomes is essential.

Policy implications Since labor supply of women is more elastic than labor supply of men, tax rates should be lower for women than for men (see e.g. Alesina, Ichino, and Karabarbounis 2011, who study gender-based taxation). However, joint taxation of spouses precisely counters this finding. Instead of lower taxes, married women being typically the secondary earner are facing a higher tax burden compared to men. While the elimination of joint income taxation might pose constitutional challenges, abolishing the joint withholding income tax schedule, as currently discussed by the German government, could present a readily implementable and cost-effective reform. Based on

[30] Studies have illustrated that job characteristics such as commuting affect the sizable inequality in labor market outcomes of men and women and that there persists a gender commuting gap (Bertrand 2020; Goldin 2021). Women choose jobs with a shorter commute and move to lower-paying jobs closer to home, especially after having children (Oreffice and Sansone 2023; Petrongolo and Ronchi 2020). Hence, commuting distance could be used as a proxy for household bargaining power to evaluate whether the reform changed the household bargaining composition of spouses.

the findings of this study, the abolition of the joint withholding income tax would yield a significant boost to female and overall labor supply and is likely to be only a lower bound compared to a move to individual taxation for the final income tax. This holds particular relevance in the current era of skill shortages and demographic changes that we are currently facing.

Appendix to Chapter 1

1.A Additional Institutional Details

	Tax filing		
Country	Joint	Optional	Family
Brazil		x	
Chile		x	
Estonia		x	
France			x
Germany		x	
Ireland		x	
Israel		x	
Luxembourg	x		
Malta		x	
Norway		x	
Panama		x	
Poland		x	
Portugal			x
Spain		x	
Switzerland		x	
United States		x	

Table 1.A.1: Personal Income Tax Systems

Notes: This table presents all countries whose personal income tax systems are characterized either by joint filing of spouses, an option to file jointly or individually as spouses or income taxation based on the family level.
Source: OECD 2016.

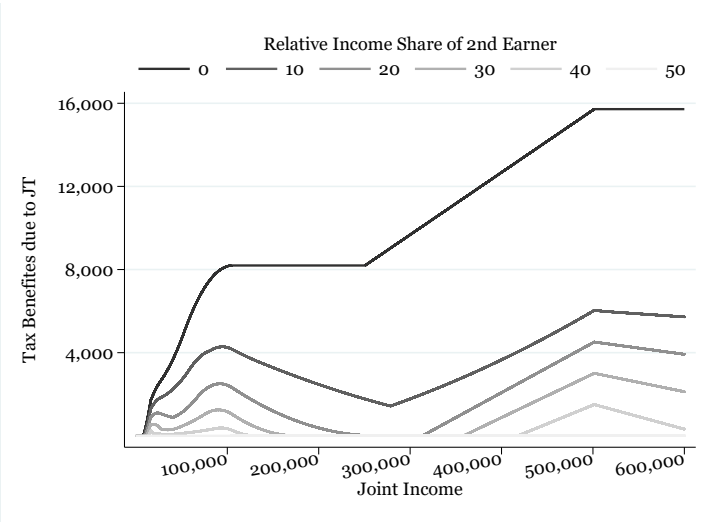


Figure 1.A.1: Tax Benefits under Joint Taxation

Notes: This figure shows the absolute benefits of joint taxation compared to single taxation for specific relative income distributions among partners. The results are based on the tax schedule in 2013.
Source: Own calculation.

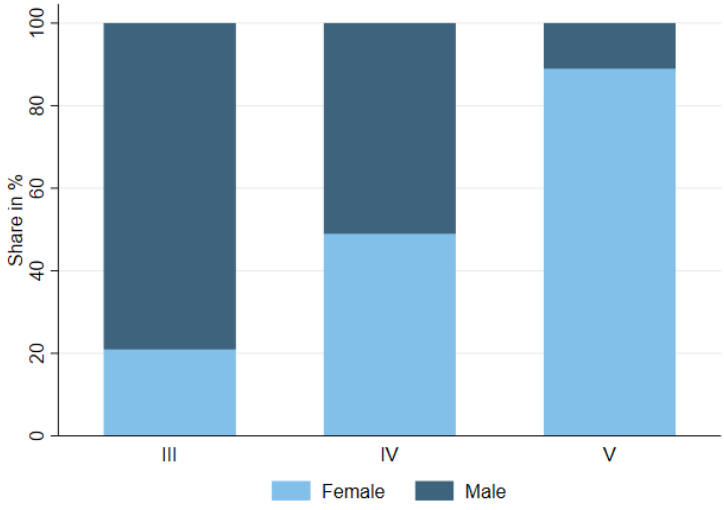


Figure 1.A.2: Gender Composition of Tax Schedules

Notes: Figure illustrates the gender composition of tax schedules. III relates to the primary earner of the joint tax schedule, IV relates to the individual tax schedule and V relates to the secondary earner of the joint tax schedule.
Source: Income tax returns 2015, German Federal Statistical Office.

Withholding tax schedule	Tax class	Tax allowances	Amount
Joint	III	Basic tax allowance of both spouses	16,010 €
		Employee's tax allowance	1,000 €
		Special expense allowance	36 €
		Provisional expense allowance	yes
		Child allowance	7008 €
Individual	IV	Basic tax allowance	8,005 €
		Employee's tax allowance	1,000 €
		Special expense allowance	36 €
		Provisional expense allowance	yes
		Child allowance	3504 €
Joint (taxed as second income)	V	Basic tax allowance	-
		Employee's tax allowance	1,000 €
		Special expense allowance	36 €
		Provisional expense allowance	yes
		Child allowance	-

Table 1.A.2: Withholding Tax Schedules

Notes: Table presents the withholding tax schedules, corresponding tax classes and the relevant tax allowances.
Source: Program flow chart by German Federal Ministry of Finance, 2013.

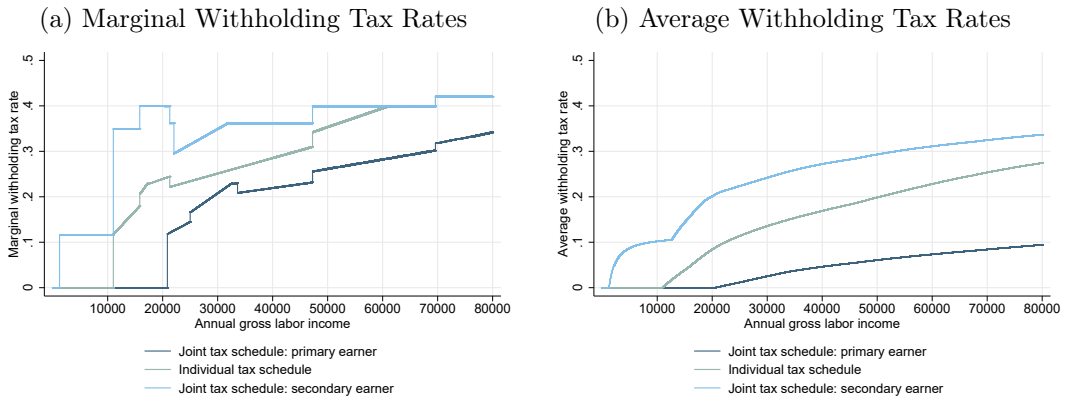


Figure 1.A.3: Witholding Income Tax Rates, 2013

Source: Program flow chart by German Federal Ministry of Finance, 2013.

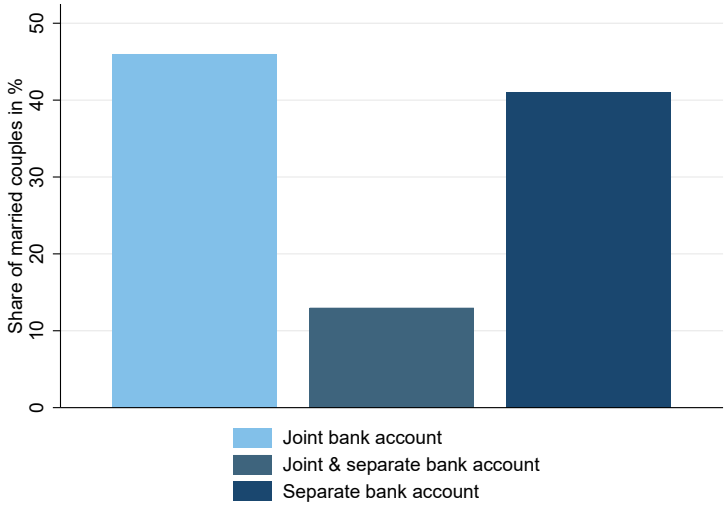


Figure 1.A.4: Joint or Separate Bank Account

Notes: Figure illustrates the share of married couples having a joint, separate or joint and separate bank account. The sample is restricted to married couples aged 20-45 years to align it to the sample of newlyweds with the TPP.
Source: Pairfam, 2012/13 wave.

	Pre-reform (up to 2012)	Post-reform (after 2012)
Form	Paper card	ELSTAM (electronically stored)
Responsible authority	Local city office	Local financial authority
Tax class		
After marriage	Active choice (III/V or IV/IV)	Default: IV/IV, letter for III/V
One working spouse	III/-	IV/IV

Table 1.A.3: Features of the Withholding Income Tax System Pre & Post Reform

1.B Additional Descriptive Statistics

	Treatment group		Control group	
	pre-reform	post-reform	pre-reform	post-reform
Female age	33	33	46	48
Male age	36	36	49	51
# children	0.55	0.55	0.99	0.88
Age first child	3	3	18	20
Catholic	0.44	0.43	0.39	0.38
West Germany	0.85	0.85	0.77	0.77
Female labor income	28,504 €	28,661 €	28,912 €	30,057 €
Male labor income	45,295 €	45,191 €	50,135 €	51,150 €
Female income share	0.39	0.39	0.39	0.38
N couples	229,822	254,761	2,052,730	2,072,131

Table 1.B.1: Descriptive Statistics Treatment vs. Control Group

Notes: The table shows summary statistics (mean values) for the treatment and control group for the pre-reform (2011-2012) and post-reform (2013-2014) period. The treatment group are newlyweds in the year of marriage. The control group are married couples who have married before 2011 and stay married throughout 2011-2018. I restrict the sample to dual earner couples for whom both spouses have positive labor income and no income from self-employment. Income is adjusted by the consumer price index.

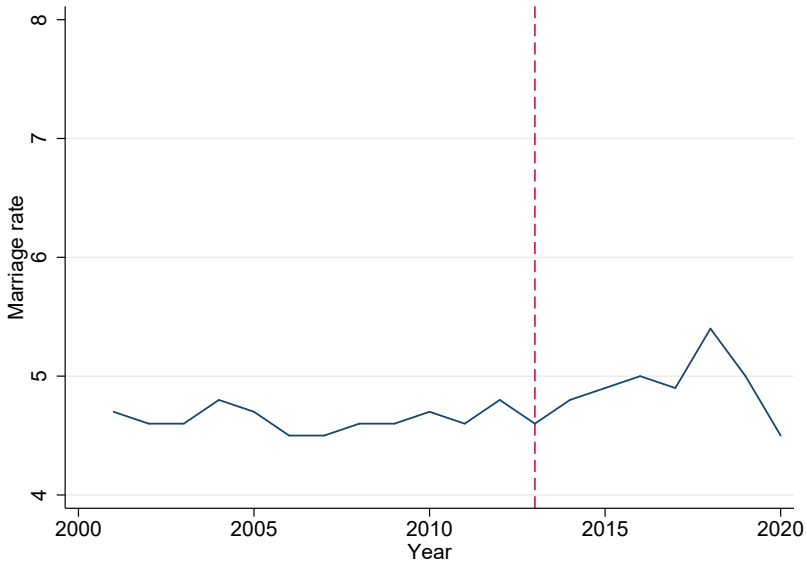


Figure 1.B.1: Marriages per 1,000 People in Germany

Source: OECD.

Female age at marriage	34
Male age at marriage	37
# children	0.45
Catholic	0.45
West Germany	0.85
Female labor income	23,858 €
Male labor income	56,145 €
Female income share	0.30
N couples	159,957

Table 1.B.2: Descriptive Statistics for Male Primary Earner Couples

Notes: The table shows summary statistics (mean values) for the pre-marriage year (except for age) for the male primary earner couples who get married between 2011-2015 and who are observed +/-3 years of marriage. I restrict the sample to dual earner couples for whom both spouses have positive labor income and no income from self-employment in the year before marriage. Income is shown in 2015 Euros.

1.C Additional Results

	Baseline specification	Control for		
		income	income share	both
	(1)	(2)	(3)	(4)
Second stage	-0.064***	-.083***	-0.66***	-0.084***
	(0.011)	(0.011)	(0.011)	(0.011)
First stage	0.956***	0.975***	0.956***	0.975***
	(0.003)	(0.003)	(0.003)	(0.003)
Initial income FE		x		x
Initial income share FE			x	x
Age FE	x	x	x	x
Year FE	x	x	x	x
Controls	x	x	x	x
Observations	404,188	404,188	404,188	404,188

Table 1.C.1: Elasticity Estimates for Men

Notes: Estimates for the elasticity of gross income with respect to the withholding net-of-tax rate from the 2SLS estimation of Equation 1.4 and 1.5 for men. Column (1) does not include pre-marriage income and income-share-fixed effects (FE). Column (2) includes pre-marriage income-fixed effects and column (3) includes pre-marriage income-share-fixed effects. Column (4) includes both pre-marriage income and income-share-fixed effects. All columns include age and year fixed effects and other control variables. *,**,*** denote statistical significance at the 5, 1 and 0.1% level.

1.D Additional Robustness Checks

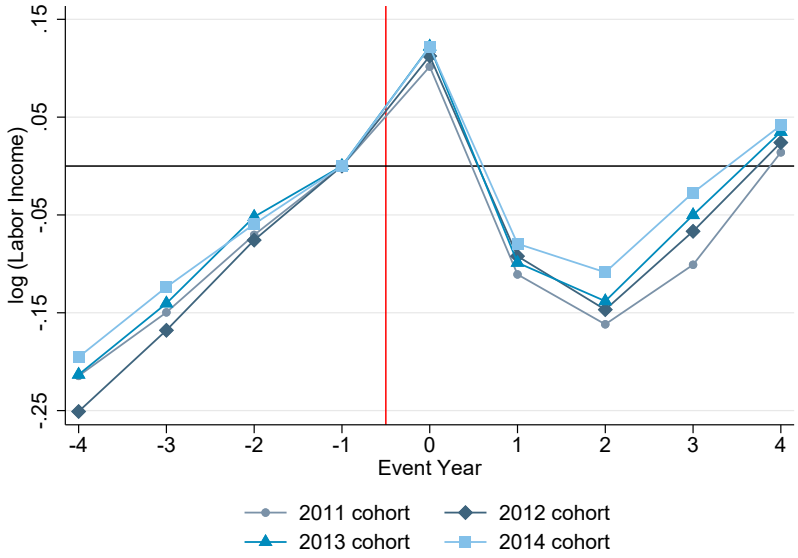


Figure 1.D.1: Cohort Event Study

Notes: This figure shows female labor income relative to the year of marriage for male primary earner couples and different newlywed cohorts. Plotted estimates are based on balanced sample. t=0 is the year of marriage.

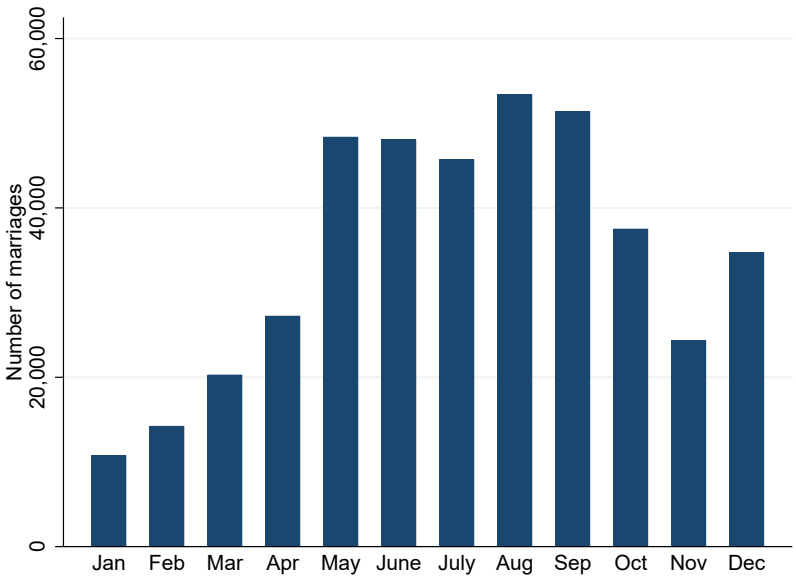


Figure 1.D.2: Distribution of Marriages Across Months in 2013

Notes: This figure shows the number of marriages per month in Germany in 2013.
Source: Destatis.

State	Percentage change
Schleswig-Holstein	0.538
North-Rhine-Westfalia	0.517
Lower-Saxony	0.443
Hessia	0.371
Saarland	0.341
Hamburg	0.332
Bremen	0.301
Bavaria	0.299
Baden-Wuerttemberg	0.267
Rhineland-Palatinate	0.185
Saxony	0.091
Mecklenburg-Western Pomerania	0.087
Brandenburg	0.084
Thuringia	0.074
Berlin	0.060
Saxony-Anhalt	-0.037

Table 1.D.1: Change in the Share of Day Care Centers

Notes: The table shows the change in percent in the day care center share in the respective state between 2011-2015.
Source: Destatis.

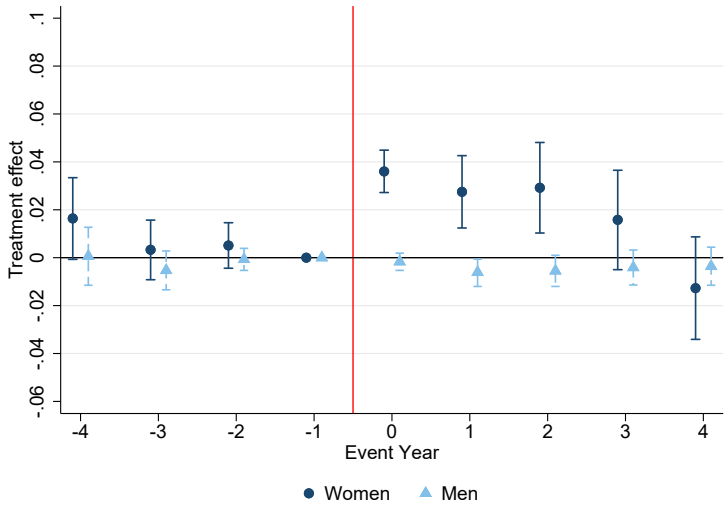


Figure 1.D.3: Treatment Effect for Labor Income Controlling for Roll Out of Daycare Centers

Notes: This figure shows the treatment effect for female and male labor income for male primary earner couples, controlling for the roll out of day care centers. Plotted estimates are based on balanced sample. t=0 is the year of marriage. Vertical segments are 95% confidence intervals.

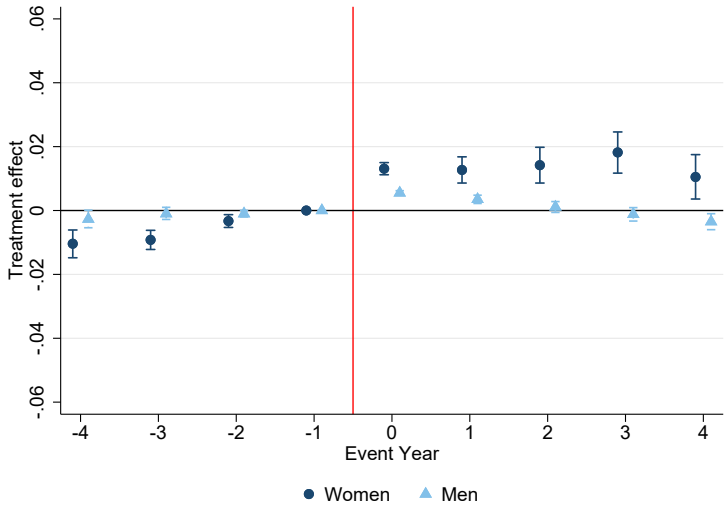


Figure 1.D.4: Treatment Effect for Labor Force Participation Controlling for Roll Out of Daycare Centers

Notes: This figure shows the treatment effect for female and male labor force participation for male primary earner couples, controlling for the roll out of day care centers. Plotted estimates are based on balanced sample. t=0 is the year of marriage. Vertical segments are 95% confidence intervals.

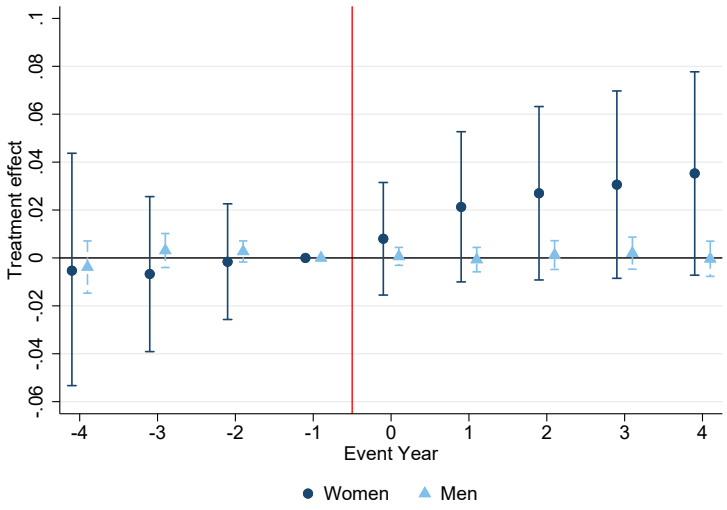


Figure 1.D.5: Treatment Effect for Labor Income for Newlyweds above the Parental Benefit Cap

Notes: This figure shows the treatment effect for female and male labor income for male primary earner couples whose individual labor income is above the income threshold for the parental benefit cap. Plotted estimates are based on balanced sample. t=0 is the year of marriage. Vertical segments are 95% confidence intervals.

Chapter 2

Disentangling Gender Norms and Tax Incentives - Analyzing the Introduction of Joint Income Taxation for Same-Sex Couples*

2.1 Introduction

The persistent disparity in long-run earnings between married men and women, particularly mothers, has been a concern in the recent economic literature (see for instance, Blau and Kahn 2017; Kleven, Landais, and Søgaard 2019). Joint taxation is considered to be one of the contributing factors behind the gender earnings gap by reducing incentives to work of the secondary earner, typically the wife, and thereby limiting the labor market participation of married women (Borella, De Nardi, and Yang 2023). In a progressive income tax system, joint taxation leads to an increase in one spouse's marginal tax rate not only in their own income but also in the spousal income. The secondary earner faces a higher marginal tax rate compared to individual taxation, while the reverse is true for the primary earner. However, analyzing the effect of joint income taxation is challenging due to the following reasons. First, finding exogenous variation that changes the tax burden between spouses as natural experiments is difficult due to the lack of the introduction or abolition of joint taxation. Second, transitions between tax regimes are generally accompanied by changes in marriage incentives, introducing an additional margin of distortion. Third, traditional gender norms that promote the disproportionate burden of care work and division of labor often relegate women to the role of secondary earners and can consequently have parallel negative effects on female earnings as it is the case for joint taxation. This paper aims to overcome these challenges.

We exploit the introduction of joint taxation for same-sex civil partnerships in Germany in 2013 to isolate the incentive part of joint income taxation on partners' labor supply. Since 2001, same-sex couples have been able to enter registered civil partnerships in Germany. However, it was not until 2013 that they could benefit from joint taxation, compared to different-sex couples who already had access. This reform left all other aspects of civil partnerships unchanged. It offers a quasi-experimental setting that exogenously changes partners' tax schedules. Due to kinks in the German tax code, changes in marginal tax rates are heterogeneous depending on a household's overall and relative income. Additionally, marriage (or civil partnership) incentives do not change for couples who already are in a civil partnership as the reform only changed partners' tax burden. To disentangle gender norms and the role of the tax system in place, same-sex couples are a valuable comparison group as they are arguably less affected by traditional within-couple gender norms (Goldberg 2013).

*This chapter is based on co-authored work with Elena Herold and Carina Neisser.

We employ novel and unique administrative tax return data for the years 2008 to 2018. More specifically, we are able to observe the universe of all civil-partnered same-sex couples who file their income tax return jointly at some point in time and link these observations to both individuals' income tax data before the legalization of joint filing. Administrative tax records do not suffer from misreporting or small sample sizes and offer precise and comprehensive information on individual and household income as well as the tax liability of couples.

To empirically quantify the impact of joint taxation, we exploit the introduction of joint taxation as a quasi-experiment and use both the individual (secondary and primary earner) and the couple as our observational unit. We apply an event study and (dynamic) Difference-in-Differences approach using same-sex couples as our treatment and different-sex couples as our control group. This specification allows us to estimate effects across (absolute) income levels but also across relative incomes within households.

Our dynamic Difference-in-Differences approach shows that a same-sex secondary earner experiences a significant decline in his income of up to 17% after filing taxes jointly relative to a different-sex secondary earner and prior to the introduction of joint taxation for same-sex couples. The partner pay gap increases by up to 14 percentage points three years after the introduction of joint filing for same-sex couples relative to the partner pay gap of different-sex couples. Therefore, joint taxation has substantial effects on partners' income and intra-household inequality. We plan to check the robustness of these results with several approaches.

We additionally explore whether same-sex couples engage in tax planning in the form of shifting the withholding tax burden among partners. We find that same-sex civil partners are, irrespective of the size of their partner pay gap, less likely to choose a withholding income tax schedule which results in a higher marginal tax burden for the secondary earner compared to different-sex spouses with a male primary earner.

Related literature We contribute to the following strands of literature. First, we contribute to the literature studying spouses' labor supply responses to family income tax reforms. There is only a small quasi-experimental literature due to the lack of natural experiments. The findings consistently document a negative (positive) impact of joint (individual) taxation on women's labor supply, who are typically the secondary earner, whereas no effect for men (Kalíšková 2014; Koch 2024; LaLumia 2008; Selin 2014).[1] Closest related to our paper, Isaac (2023) analyzes the impact of the introduction of joint filing for same-sex marriages on the labor supply of same-sex spouses in the US using data from the American Community Survey (ACS). He exploits the switch from individual to joint taxation created by federal recognition of existing state-level same-sex marriages in the US in 2013. Employing a generalized Difference-in-Differences approach with different-sex couples as the control group, Isaac (2023) estimates the effect of the predicted change in the marginal net-of-tax rates on annual hours worked and labor force participation of same-sex couples. He finds no significant intensive margin responses (hours worked) but relatively large extensive margin responses (labor force participation) among secondary earners which is in line with Selin (2014).

[1] Besides quasi-experimental studies, a series of papers use structural life cycle models to evaluate the effect of joint taxation on married women's labor supply (see Bick and Fuchs-Schündeln 2017, 2018; Borella, De Nardi, and Yang 2023; Guner, Kaygusuz, and Ventura 2012).

Our project adds to this literature in various ways. First, by using high-quality administrative data, we avoid problems of income misreporting and identification errors of treated individuals that are common in survey data.[2] The tax return data offers us precise data on earnings and taxes over a longer time window, allowing us to follow individuals/couples and observe how effects evolve over time. Second, we have access to the universe of German same-sex couples filing jointly, which gives us a large sample size. Third, the income tax system in Germany offers clear financial benefits for the couple's total tax burden when switching from single to joint income taxation, whereas the joint taxation system in the US imply a marriage benefit for some newlyweds and a marriage penalty for others. We estimate an average tax benefit of 829 € for same-sex couples, whereas about 40% of marriages in the US face a "marriage penalty" and Isaac (2023) reports an average marriage penalty of $449 for same-sex couples. Therefore, in Germany, both the income and substitution effect imply disincentives for the secondary earner to work, whereas this is not consistently so for the US. Moreover, the German withholding income tax system allows married couples or civil partnerships to engage in tax planning and shift the withholding income tax burden among partners to minimize the couple's total intra-year tax burden. We can exploit this feature to explore the effect of tax planning on intra-household inequality of same-sex couples. Overall, our results are consistent with the previous literature, indicating that the secondary earner's labor supply declines after the introduction of joint taxation. In contrast to the earlier literature, we find substantial responses on the intensive margin of labor supply.

Our study also relates to a growing literature on the labor supply and earnings of same-sex individuals and couples as well as LGBTQ individuals in general (see Badgett, Carpenter, and Sansone 2021, for an excellent review). Economic research is limited as there are only few available sources of data. Studies focusing on labor market outcomes show that same-sex couples distribute their household and labor allocation more equally (e.g. Van der Vleuten, Jaspers, and Lippe 2021). Male same-sex couples spend less and female same-sex couples more time in the labor force than their same-gender counterpart in different-sex couples, however the gap among female secondary earner between different- and same-sex couples is significantly reduced when controlling for children (Antecol and Steinberger 2013; Leppel 2009; Tebaldi and Elmslie 2006). In addition, Carpenter (2007) and Drydakis (2022) show that there exists a wage gap between hetero- and homosexual individuals, which is positive for women and negative for men. We explore how joint taxation affects earnings of same-sex couples filling a notable research gap highlighted by Badgett, Carpenter, and Sansone (2021) regarding public policies affecting LGBTQ individuals. To the best of our knowledge, our study is the first that examines such a large sample of same-sex couples with high-quality administrative data.

This paper is set up as follows. Section 2.2 describes the institutional setting and section 2.3 provides an overview on the literature on same-sex couples' labor supply. Section 2.4 presents the employed data. Section 2.5 explains the identification strategy. Section 2.6 shows the results on same-sex couples' earning responses. Section 2.7 presents the descriptive results on tax planning of same-sex couples. Section 2.8 concludes.

[2]Same-sex couples are clearly identified since this information has to be declared upon tax filing.

2.2 Institutional Setting

2.2.1 Taxation of Couples in Germany

Final income tax As in many other countries, including the US and various European countries (see Table 1.A.1 for an overview), the German tax system offers married couples and civil partnerships the option to file their taxes jointly. The default for filing taxes for married couples in Germany is joint taxation and the vast majority of married couples opt for filing taxes jointly. However, spouses have the option to choose individual taxation. To formalize the difference between individual and joint taxation, we compare the total tax liabilities for both systems. Under individual taxation, the total tax liability, $T_{individual}$, is simply the summation of tax liabilities based on income from spouse 1 (denoted as y_1) and spouse 2 (y_2, respectively):

$$T_{individual} = T(y_1) + T(y_2) \tag{2.1}$$

Thus, the marginal and average tax rates of spouse 1 and 2 are independent of the partners' incomes. Conversely, under joint taxation, the couple's income tax liability, T_{joint}, is based on half of their combined income, which is then doubled:

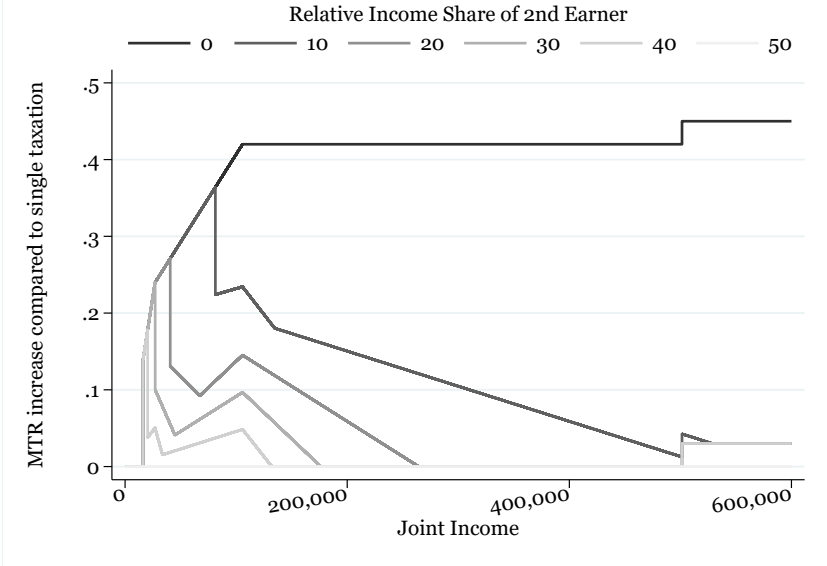
$$T_{joint} = 2 \cdot T((y_1 + y_2)/2) \tag{2.2}$$

In the context of a progressive tax system, joint income taxation has two effects. First, the marginal tax rate (MTR) of the primary earner is lower than under individual taxation, while the secondary earner is subject to a higher MTR. Figure 2.1a illustrates the variation in the MTR for the secondary earner based on both the secondary earner's and the overall household income for joint taxation relative to individual taxation. Generally, the larger the income difference between spouse 1 and 2, the larger is the increase in the MTR for the secondary earner. Various kinks in the German tax code contribute to variations in the MTR for the secondary earner in relation to the overall household income. Second, joint taxation results in a reduced joint tax burden, translating to marriage tax benefits.[3] Figure 2.1b shows these benefits for specific household incomes and relative income shares. The magnitude of benefits increase in the size of spouses' relative income difference, notably favoring the single-earner model. The impact of joint household income on benefits exhibits heterogeneity owing to kinks in the German income tax schedule.

Although tax systems may not explicitly discriminate between genders, they can feature implicit gender biases that arise from the interaction between the tax system, economic behavior and social expectations from male and female taxpayers. For Germany, Hermle, Herold, and Hildebrand (2024) show that the average female income share is around 36% for all married women. Differences in relative household income, in combination with a progressive and joint income tax system, introduce lower labor supply incentives for women as the secondary earner. This stems from the fact that under joint taxation, women face higher MTR relative to individual taxation. Consequently, this higher tax burden may deter women from fully participating in the labor force - at

[3] In contrast to Germany, the US or Switzerland's joint taxation systems might imply a marriage penalty (higher taxes for married individuals compared to unmarried individuals with the same income).

(a) Change in Marginal Tax Rates for the Secondary Earner under Joint Taxation



(b) Tax Benefits under Joint Taxation

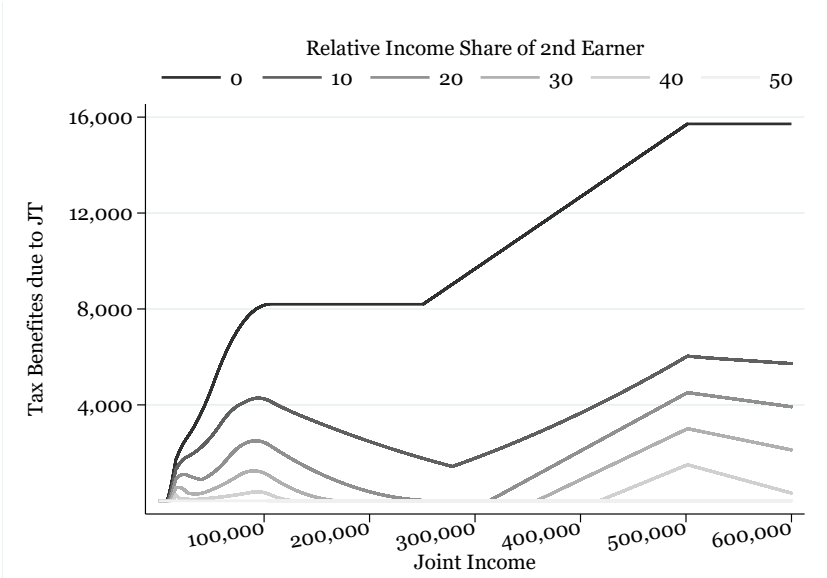


Figure 2.1: Joint Taxation vs. Individual Taxation

Notes: Figure (a) shows the difference in marginal tax rates for the secondary earner when switching from single to joint taxation for different levels of the income of the partner. Figure (b) shows the absolute benefits of joint taxation compared to single taxation for specific relative income distributions among partners.
Source: Own calculation based on the tax schedule in 2013.

the extensive and intensive margin, potentially limiting their earnings relative to their male counterparts.

Withholding income tax In the German setting, the adverse labor supply effects of joint taxation on the secondary earner may further be amplified by spouses' choice of the withholding income tax schedule. Germany has automatic income tax withholding for income from employment as numerous developed and developing countries. Expected income tax payments are withheld from the employee's salary by the employer and directly transferred to the tax authority on a monthly basis. The withholding income tax is credited against the final income tax once the individual/married couple files their tax return in the following year. In Germany, the withholding tax is determined by the so-called tax class. By default, both spouses are assigned to tax class IV, translating to the individual tax schedule.[4] Due to tax benefits from tax splitting under joint income taxation, this often results in overestimated withholding income taxes for couples with unequal earnings. To align monthly withholding income taxes with final annual income taxes, spouses can also opt for assigning one spouse (typically the primary earner) tax class III and the other spouse (typically the secondary earner) tax class V, translating to joint taxation. The basic tax allowance from the spouse in tax class V is shifted to the spouse in tax class III. This leads to higher average and marginal tax rates for the secondary earner resulting in a higher individual tax burden and lower net earnings relative to the individual tax schedule. The primary earner faces lower average and marginal tax rates and thus a smaller monthly tax burden and higher net earnings. Notably, the effective final income tax burden of the couple is not affected by the choice of the withholding tax schedule, but the intra-year timing of the tax payments and the distribution of the individual tax burden within the couple. In theory, withholding rates should not matter for the labor supply decision as long as households are not liquidity constrained in the course of the year, spouses pool their income completely and do not misperceive their final income tax as their withholding income tax. However, studies illustrate that spouses do not pool their income completely (see e.g. Attanasio and Lechene 2014; Beblo and Beninger 2017; Blundell et al. 2007; Giommoni and Rubolino 2022; Lundberg, Pollak, and Wales 1997) and that individuals misperceive complex tax schedules (see e.g. Abeler and Jäger 2015). Consequently, the German income tax not only incentivizes unequal earnings among spouses but also reinforces an unequal allocation of net wage income, exacerbating the negative impact of joint taxation on the secondary earner.

2.2.2 Same-Sex Civil Partnerships

To establish a legal framework for same-sex couples in Germany, the concept of civil partnership (*Eingetragene Lebenspartnerschaft*) was introduced in 2001, marking a pivotal shift from a historical backdrop tainted by anti-homosexual regulations.[5] [6] While

[4] The default of IV/IV was introduced in 2013. Before 2013, spouses had to actively choose their tax classes at their local city office (Koch 2024).

[5] Most notably, until 1994, the *law against homosexuality act* criminalized relationships between two males, albeit softened in 1969 to solely prohibit homosexual prostitution and illicit relations before the age of 21.

[6] Note that a civil-partnership was only an option for same-sex partners in Germany, distinguishing it from the legal framework in countries like France, where different-sex civil partnerships exist.

civil partnerships in Germany granted homosexual couples access to many rights akin to different-sex marriages, certain distinctions persisted due to political dynamics within the reigning coalition of the Social Democrats and the Green Party. These distinctions were strategically maintained to secure a majority in the Federal Council (Sanders 2016).

In 2013, joint income taxation was extended to civil partnerships, rectifying a prior inequality. The different treatment of same-sex and different-sex couples under income tax laws was ruled incompatible with the constitution on May 7, 2013 (BVerfGE, May 2013). The reform was introduced in July 2013. Moreover, joint taxation was not only allowed for the years 2013 onwards but civil partnerships could file their taxes jointly retroactively for earlier years of their civil partnership if they had not already filed their taxes for the respective years or they had filed objections against their local tax authority for not having granted them joint filing. As a consequence, these tax assessments were not legally binding yet. Same-sex couples received the refunds for the years they filed retroactively as a one-off payment including interest. This observation aligns with expectations, considering the substantial benefits outlined in Figure 2.1b. The implementation of joint filing for same-sex couple did not immediately entail the automatic assignment of tax classes IV/IV, as is the case for married different-sex spouses; rather, this change was only instituted in 2015. Same-sex couples were required to file a letter to their local tax authority requesting a change in their tax classes to IV/IV or III/V in the years of 2013 and 2014.

In October 2017, marriage for same-sex couples was finally introduced.[7] Existing civil partnerships could decide whether to maintain their status or transform into marriage. This legislative shift was coupled with the abolition of the option to establish new civil partnerships.

Figure 2.2 illustrates the absolute number of registered civil partnerships spanning the years 2007 to 2020. Total numbers increase each year leading up to 2017 when same-sex marriage was introduced, and no new civil partnerships were established, with existing ones having the choice to transition to marriage. The absolute number of male same-sex partnerships exceeds the number of female partnerships every year. Importantly, there is no discernible spike in the years 2013 or 2014, indicating no discontinuous pattern in the dynamics of civil partnerships after the introduction of joint filing for same-sex civil partnerships.

2.3 Labor Supply Responses of Same-Sex Couples

We aim to examine the impact of introducing joint taxation on partners' labor market outcomes and intra-household income distribution. In heterosexual relationships, conventional gender norms often discourage women, typically the secondary earner, from participating in the labor force, aligning with the potential effects of joint taxation. Conversely, gender roles may exert less influence in same-sex partnerships due to the absence of established social expectations and anticipated sanctions, thereby allowing for a less gendered division of labor. Previous research by Andresen and Nix (2022) indicates a reduced child penalty for women in same-sex couples, attributing this phenomenon to differences in gender norms and preferences. Thus, same-sex couples offer

[7] Same-sex marriage was supported by a majority of the population (Küpper, Klocke, and Hoffmann 2017).

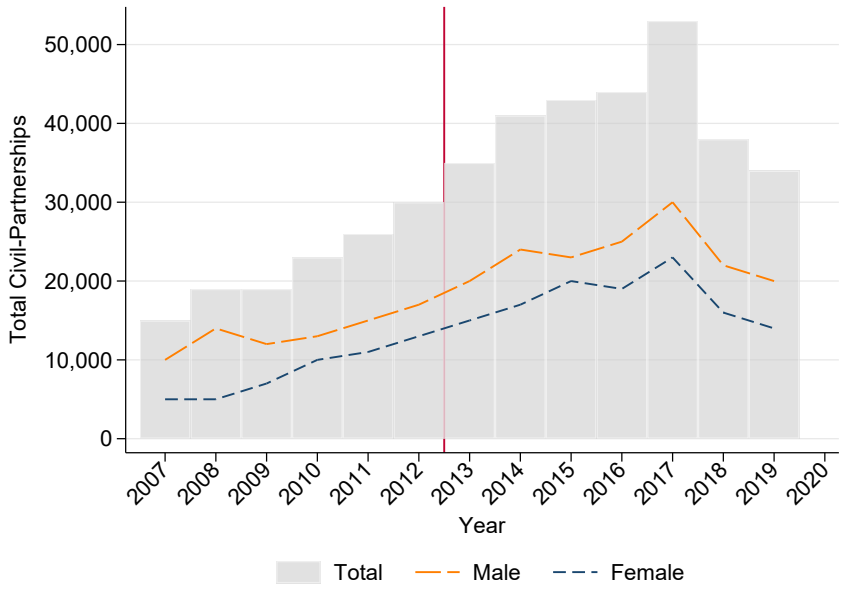


Figure 2.2: Absolute Number of Same-Sex Civil Partnerships

Notes: This figure shows the annual absolute number of registered same-sex civil partnerships in Germany for male, female, and total same-sex partnerships. The red line indicates the introduction of joint taxation.
Source: German Federal Statistical Office, based on Micro Census.

an advantageous setting for isolating tax incentives from gender-related norms. However, same-sex and different-sex couples may diverge in several attributes that could contribute to divergent labor supply responses and intra-household specialization as well (Hansen, Martell, and Roncolato 2020).

First, drawing from the theory of Becker (1981) on comparative advantage and family specialization, partners may derive mutual benefits through specialization in market and household work. Unlike different-sex couples, where distinct gender roles often dictate specialization, same-sex partners may not experience such clear relative efficiencies due to similar socialization. Consequently, same-sex couples are anticipated to specialize less, reflecting their reduced potential gains from comparative advantage (Becker 1981). This is generally confirmed by the literature (Bauer 2016; Dilmaghani and Dean 2023; Hofmarcher and Plug 2022). Furthermore, this is consistent with findings showing a positive wage gap for women and a negative wage gap for men in same-sex couples (Badgett, Carpenter, and Sansone 2021; Jepsen and Jepsen 2022). However, Giddings et al. (2014) document a narrowing specialization gap between same-sex and different-sex couples over time in the US, with Hofmarcher and Plug (2022) attributing these disparities primarily to the most traditional different-sex couples.[8]

Second, according to bargaining models, partners with distinct preferences negotiate

[8] Additionally, the model of Becker (1981) implies that married spouses, given their higher commitment level and stability, have a greater incentive to specialize compared to civil partners. Notably, same-sex marriage was only legalized in various countries in the last decade, including Germany. Therefore, same-sex partners may have specialized less.

to reach a mutually acceptable division of labor (Manser and Brown 1980). In these models, higher earnings typically confer greater bargaining power, leading the higher-earning partner to specialize in market work. Importantly, bargaining models do not make gender-specific predictions, making them applicable to both same- and different-sex couples. Despite differences in suggested mechanisms between family economics and bargaining approaches, the predictions from both theories are notably similar. Specifically, given the narrower earnings gap on average in same-sex couples, bargaining models, akin to comparative advantage models, suggest that same-sex couples will specialize less than different-sex couples. However, Oreffice (2011) finds that same-sex couples exhibit similar intra-household bargaining dynamics in their labor supply decisions compared to heterosexual couples.

Third, gender-conforming norms and intra-household specialization are often reinforced by parenthood. Biological differences, such as breastfeeding, and caregiving skills become particularly relevant in households with children. Consequently, the division of housework in same-sex couples may appear more balanced simply because gay and lesbian couples are less likely to live together with children (Black, Sanders, and Taylor 2007). Controlling for the presence of children in the household is expected to mitigate differences between same- and different-sex couples. Indeed, this is corroborated by Antecol and Steinberger (2013).

Finally, even in the absence of socially defined gender roles, couples might tend to specialize over time. According to family economic theory, there is an anticipated progression towards increased differentiation and reduced sharing of household responsibilities over the course of a couple's relationship, driven by specific investments made within the partnership. Given the higher likelihood of dissolution among same-sex couples compared to different-sex couples, it is plausible that same-sex couples exhibit greater intra-household equality (Lau 2012; Valfort 2017).

To summarize, gender-conforming social norms are less relevant within same-sex couples. Additionally, same-sex couples exhibit relatively comparable patterns to different-sex couples in labor market decisions, especially when accounting for factors such as children and relationship stability. Consequently, we can exploit the introduction of joint taxation for same-sex civil partners to disentangle the influence of gender norms from tax incentives.

2.4 Data

To assess the effect of joint taxation on same-sex couples' earnings, we leverage the *Taxpayer Panel* (TPP) for the years 2008-2018. It is an annual panel data set on the full population of Geman taxpayers, provided and administered by the German Federal Statistical Office and the statistical offices of the states. This administrative micro data is obtained by combining income tax returns and tax assessments from the tax authority. The unit of observation is the income taxpayer, encompassing single individuals, married couples, or civil partners filing jointly.

We have access to data on the complete population of German taxpayers which requires specific research project inquiries and is granted with a restricted set of variables.[9]

[9]The TPP is usually provided as a stratified 5% random sample.

The data include precise information on taxable income (divided by income sources), final and withholding income taxes and basic socio-demographic characteristics, such as gender, marital status, filing status, year of birth, state of residence, religion, year of birth of children, and the number of children. The data lacks information on hours worked, thus we use labor income as a proxy for labor supply.

In 2013, a same-sex marker was introduced in the data when joint filing was granted to civil partners. This marker facilitates accurate identification of same-sex couples and constitutes a valuable advantage compared to survey data from other countries. We utilize this marker to create a unique sample of the universe of all same-sex civil partners that file their taxes jointly at some point in time.[10]

Upon joint filing, one partner is appended to the other partner's spell.[11] Hence, it is possible to observe both partners whenever filing jointly. However, the original data do not allow to observe single filing spells of both partners before and after joint filing. Leveraging individual tax IDs introduced in Germany in 2010, we manage to link partners before they start to file jointly and thus across all years. To the best of our knowledge, we are the first to link German income tax data for same-sex couples before filing jointly. Coupled with our unique access to the entire population of jointly filing same-sex civil partnerships, this data set enables us to explore the dynamics of same-sex couples around civil partnerships.

We cannot identify same-sex partners who never file taxes jointly. Given the substantial benefits associated with joint filing, we expect the number to be proportionally small as for different-sex couples.[12] Furthermore, we cannot identify whether same-sex couples have already been in a civil partnership before 2013 or established one in that year for partners that first appear in the data in 2013, given the introduction of the same-sex marker variable in 2013. Therefore, our estimates may be subject to potential confounding factors arising from same-sex couples forming new civil partnerships in 2013 to capitalize on the tax advantages offered by joint filing.

2.4.1 Sample Restrictions

We focus on the years from 2008 to 2018, centered around the introduction of joint filing for same-sex civil partnerships. Additionally, our analysis includes only tax units who file their taxes jointly in 2013 - the year of the reform - and are already observed in 2012 and 2013, encompassing the periods before and after the introduction of joint income taxation. Moreover, we impose an age criterion, limiting our sample to individuals within the prime working age range, specifically those aged between 20 and 50 in 2013. To ensure a minimum level of labor force attachment, we further restrict our sample to tax units in which both individuals exhibit positive income. Hence, we measure partners' intensive margin of labor supply. We plan to explore the extensive margin of labor supply in additional analyses.

One potential concern arises regarding couples who registered their partnership in 2013

[10] We observe inconsistencies in the same-sex marker for certain individuals across different years. Nevertheless, we can rectify these inconsistencies by using the individual and their partner's IDs.

[11] For different-sex couples the wife is typically appended to the husband's spell.

[12] We estimate an average tax benefit of 829 € for same-sex couples.

following the reform. These registered partnerships may be spurred by the introduction of joint filing and are therefore endogenous to the reform. However, Figure 2.2 reveals that there was not a significant jump in registrations in the year of the reform. Consequently, it is plausible to assert that these partnerships constitute only a marginal part within the sample.

As mentioned in section 2.2, same-sex civil partners could file taxes jointly not only from 2013 onwards but also had the option to retroactively file their taxes jointly for preceding civil partnership years, provided their tax assessments for those years were still pending. Within our dataset, we can identify the retrospective joint filing and utilize this as a marker for civil partners predating 2013. However, we do not use retroactive filing same-sex civil partners as our main sample as these couples are a specific group due to two reasons. First, the act of filing taxes retroactively likely suggests a higher level of tax literacy and/or significant financial advantages associated with retroactive filing. Secondly, the intra-household income distribution of this subgroup exhibits compositional biases. We observe heterogeneity concerning the number of preceding years for which couples file retroactively. Specifically, we observe the largest share of couples opting for retroactive filing for only one year, with a minority electing to do so for multiple years, including a select few who commence retroactive filing several years prior to 2013. This can be driven by the fact that it is more difficult for us to link couples or that individuals are not included in the dataset for these earlier retroactive filing years yet. Resultingly, the sample is biased, because the changes in the average relative income share do not come from changes over time but from the composition of the sample. Thus, we currently exclude retroactively filing couples from our empirical analysis until we devise a robust methodology to address the compositional shifts within the sample. For now, we focus on same-sex civil partners who first filed their taxes jointly in 2013 and did not file retroactively.

2.4.2 Descriptive Statistics

The descriptive statistics for the entire sample, encompassing partners who retroactively file and those who do not, alongside the control group consisting of married different-sex spouses for the pre-reform year of 2012, are presented in Table 2.1. Notably, we define different-sex spouses in primary and secondary earner, irrespective of the gender, as we do for same-sex couples. Approx. 40% of our same-sex couples sample are female couples. This is consistent with the data shown in Figure 2.2 which is based on the German Micro Census. The average age of the primary earner is 40 years and of the secondary earner 38 years. This is similar across the three same-sex samples and the age difference among partners is approx. 5 years. Different-sex spouses are slightly older and have a smaller age difference. Furthermore, the number of children varies among the different samples, with different-sex couples having on average 1.72 children compared to a notably lower number of 0.32 children for same-sex couples. Column 3 highlights that among same-sex partners, those who do not engage in retroactive filing exhibit the lowest likelihood of having children. Same-sex couples are slightly less likely to live in Western Germany. Additionally, they are more likely to be self-employed, with retro-filers having the highest share of self-employed. This presents suggestive evidence that retroactively filing same-sex partners might be more tax knowledgeable as they are likely to have a tax advisor.

	Same-sex couples			
	(1) Main	(2) Retro	(3) No Retro	(4) DSC
Female couple	0.39	0.38	0.42	-
	(0.11)	(0.10)	(0.12)	
Age 1st earner	39.60	39.89	38.84	42.16
	(6.34)	(6.27)	(6.45)	(5.11)
Age 2nd earner	38.16	38.22	38.00	43.05
	(6.81)	(6.83)	(6.75)	(4.87)
Age difference	4.62	4.71	4.37	2.95
	(4.21)	(4.34)	(3.83)	(2.67)
Nr. children	0.32	0.38	0.19	1.72
	(0.71)	(0.78)	(0.51)	(0.98)
West Germany	0.81	0.80	0.82	0.86
	(0.40)	(0.40)	(0.39)	(0.35)
Self-employed	0.57	0.60	0.53	0.44
	(0.49)	(0.49)	(0.50)	(0.50)
Income 1st earner	59,490	61,266	54,803	59,567
	(71,748)	(78,685)	(48,612)	(151,704)
Income 2nd earner	20,664	17,880	28,017	12,329
	(21,040)	(21,100)	(19,015)	(17,646)
Partner pay gap	0.51	0.55	0.43	0.65
	(0.33)	(0.33)	(0.26)	(0.33)
MTR change 1st earner	-	-	-0.03	-
			(0.04)	
MTR change 2nd earner	-	-	0.07	-
			(0.09)	
Couple's tax benefit	-	-	829	-
			(1,509)	
N	12,221	8,365	3,856	2,769,383

Table 2.1: Descriptive Statistics

Notes: Table illustrates the descriptive statistics of the analyzed samples of same-sex civil partnerships and the control group of different-sex couples (*DSC*) in the baseline year before the reform (2012). The sample *Main* includes all individuals that first appear in 2013. Samples *Retro* and *No Retro* are subsamples of the *Main* sample and include only couples that made use of the option to file retroactively before in 2013 and those that did not, respectively. Income is the sum of income from business, self-employment and employment. The partner pay gap is defined as the difference between partners' income relative to the total household income. The MTR change is the hypothetical change in MTR if the same-sex couple was already filing jointly in the baseline year of 2012. The tax benefit is the same-sex couple's tax saving if joint taxation was already introduced in 2012. Raw data (unweighted), mean values.

The sum of income from business, self-employment and employment for the primary earner is very similar across the four samples, with no retro-filer primary earners having a slightly smaller income. In contrast, the secondary earner's income is substantially lower for different-sex couples and the largest for no retro-filer same-sex couples. This is consistent with the partner pay gap, which is defined as the difference between partners' income relative to the total household income. The partner pay gap is highest for different-sex couples and lowest for no retro-filer same-sex couples. Specifically, retroactively filing same-sex couples display a 12 percentage points lower gap compared to their

counterparts who do not file retroactively. This discrepancy suggests that, on average, these partners recognized the greater benefit of joint filing in cases of larger income disparities and opted to exploit the option for retroactive filing accordingly. We estimate the hypothetical change in MTR for the primary and secondary earner if the same-sex couple was already filing jointly in the baseline year of 2012. We do not do this for the retro-actively filing couples, as we do not observe their individual taxable income in 2012. The MTR would have decreased by 3 percentage points for primary earners and increased by 7 percentage points for secondary earners. Furthermore, we estimate the tax benefit if joint taxation was already introduced in 2012. Same-sex couples, who do not file retroactively, face 829 € of tax *savings*. This is substantially larger compared to the average marriage *penalty* of $449 reported by Isaac (2023).

2.5 Empirical Strategy

To estimate the causal effect of introducing joint taxation, we employ a dynamic Difference-in-Differences design with the treatment being the introduction of joint filing for same-sex civil partners in 2013. Hence, same-sex couples are our treatment group since they enjoy potential tax benefits resulting from the introduction of joint income taxation. Our control group are married different-sex couples because they could already benefit from joint filing and the reform in 2013 does not affect them. The dynamic Difference-in-Differences approach allows us to observe how the effect of joint taxation on partners' income evolves over time.

Difference-in-Differences For our Difference-in-Differences specification, we estimate the following dynamic regression:

$$log(y_{it}) = \delta_{h(i)} + \gamma_t + \mu T_i + \sum_{k \neq 2012} \beta^k \cdot (D_t^k \cdot T_i) + X_{it} + \epsilon_{it} \tag{2.3}$$

where y_{it} refers to the respective outcome variable of unit i at time t. Unit i either refers to the secondary earner, the primary earner or the couple. The outcome variables are secondary earner's income, primary earner's income, and the partner pay gap. Income is the sum of income from business, self-employment and employment. We adjust income variables by the consumer price index. T_i is a binary treatment indicator variable equal to 1 when i is in a same-sex civil partnership. D_t is a binary time indicator variable that takes the value of 1 in each post-treatment year. $(D_t^k \cdot T_i)$ is the interaction term, and hence, β_k the coefficient of interest, which captures the difference in treatment effects before and after the reform. γ_t and $\delta_{h(i)}$ are year and age-fixed effects, respectively. Year-fixed effects will capture national-level time-varying shocks that may affect labor supply among same- and different-sex couples. X_{it} includes control variables on unit i such as age difference, combined age, children, and region.[13]

The identifying assumption we make in this Difference-in-Differences framework is that incomes of same- and different-sex couples would have evolved parallel to each other in the absence of the introduction of joint filing for same-sex civil partners in 2013. We explore this in Section 2.6.

[13]Since our setting involves a single treatment date, we do not need to assume homogeneous treatment effects as highlighted in the new Difference-in-Differences literature (e.g. Goodman-Bacon 2021).

Elasticity of taxable income To estimate behavioral responses to taxation, we plan to rely on a well-known concept in public finance - the elasticity of taxable income (ETI). The ETI measures the responsiveness of taxable income to changes in the net-of-tax rate (NTR) - defined as one minus the marginal tax rate. It summarizes all different types of behavioral responses to income taxation. Compared to labor supply elasticities, it also captures tax avoidance and tax evasion behavior. We plan to leverage variation in each partner's marginal tax rate as a result of the switch from individual to joint taxation due to the reform, illustrated in Figure 2.1a.

In a progressive tax system, the marginal tax rate and income are jointly determined. To overcome endogeneity problems, we will take a common approach and construct mechanical tax rate changes. We will hereby apply post-reform tax rules on pre-reform taxable income and instrument for the NTR. Since pre-reform taxable income is not available by definition, we will construct hypothetical pre-reform taxable income for each couple by relying on a tax simulation model. The resulting change in the net-of-tax rates is free of any behavioral responses. We will apply instruments proposed by Gruber and Saez (2002).

An interesting aspect of such an approach is that when using different outcome variables, e.g. taxable income vs. overall earnings, we limit the range of responses. For example, it allows us to see if individuals change their real behavior (e.g., labor supply adjustments) or if we see any changes in reporting behavior (e.g., a stronger response with taxable income).

2.6 Results

In this section, we present the estimation results of the effect of the introduction of joint taxation on the primary and secondary earners' income and the partner pay gap, employing a Difference-in-Differences approach.

Income We first run an event study for incomes of the primary and secondary earner for same- and different-sex couples to check how incomes evolve around the reform. Figure 2.3a illustrates the income trend for primary earners and Figure 2.3b for secondary earners. Note that we define primary and secondary earner based on pre-reform years $t = -3$ to $t = -1$, so that partners do not switch between being the primary or secondary earner. The income of primary earners in same-sex partnerships remains relatively stable prior to the reform, experiences a decrease during the reform year, and then exhibits an upward trend starting from $t = 1$. In contrast, primary earners in different-sex couples display a consistent upward trend both before and after the reform. Secondary earners' income is relatively stable around the reform and starts an upward trend in $t = 2$ for same-sex couples, whereas different-sex secondary earners' income shows a stable upward trend for pre- and post-reform years.[14] Consequently, it appears that primary and secondary earners from same- and different-sex partnerships follow slightly divergent trends in $t = -3$ leading up to the reform.

[14]Koch (2024) studies a reform to the withholding income tax in the same year of 2013 that decreased the marginal withholding income tax for female newlywed secondary earners. She finds that this group of women increase their income in response to the reform. We do not include newlyweds in our control group of different-sex couples. Consequently, the withholding income tax reform in 2013 should not affect the incomes of our control group.

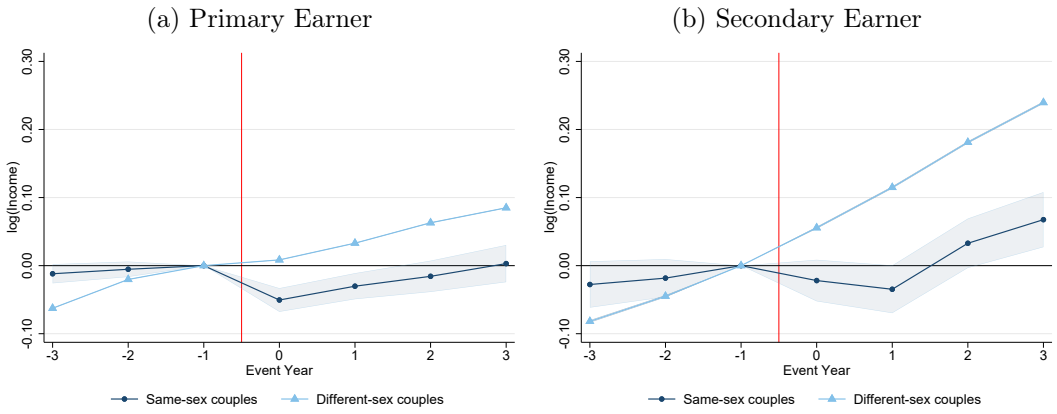


Figure 2.3: Income Relative to Reform Year

Notes: This figure shows event study results from estimating Equation 2.3 for the secondary and primary earner without the interaction term $D_t \cdot T_i$. The sample includes all same-sex couples that first filed jointly in 2013 and did not retroactively file taxes jointly. Panel (a) shows the result for primary earners. Panel (b) shows the result for secondary earners. Primary and secondary earners are defined on the basis of the pre-reform years. $t = 0$ is the reform year of 2013. Income is the sum of income from business, self-employment and employment. Shaded areas are 95% confidence intervals. Confidence intervals of different-sex couples are very narrow due to the large sample size.

To explore the divergent trends, we employ the following robustness checks. First, we use an alternative definition for primary and secondary earner based only on the baseline year $t = -1$, illustrated in Figure 2.A.1a and 2.A.1b. Pre-trends of same-sex and different-sex primary earners seem to be more similar for this definition. This might be driven by the hump-shaped trend for same-sex primary earners that is due to the restriction on the baseline year $t = -1$. For secondary earners this pattern is somehow reversed. Income falls until $t = 0$, remains relatively stable and increases from $t = 2$. Therefore, we prefer the definition on all pre-reform years, for which partners consistently remain the primary or secondary earner before the reform. Second, we employ a donut-hole approach and exclude couples whose individual incomes differ by only +/- 5%. Figures 2.A.1c, 2.A.1d, 2.A.1e, and 2.A.1f show that this robustness check does not change the income trends of primary and secondary earners of both the treatment and control group. We plan to exclude couples whose incomes differ by a larger percent to explore the income trends of couples for who joint taxation is in general more beneficial. Moreover, we plan to investigate the role of children for income trends and focus on only couples with/without children.[15]

To eliminate general time trends, we next turn to our Difference-in-Differences approach, for which we employ different-sex couples as our control group. One implication of the identifying parallel trend assumption is that we should not observe an economically

[15]In Germany, a large roll out of day care centers for children below three years took place between 2004-2014. Mothers in different-sex couples, who are typically the secondary earner and do the majority of child care, could more easily return to work. This might explain the strong income growth for different-sex secondary earners.

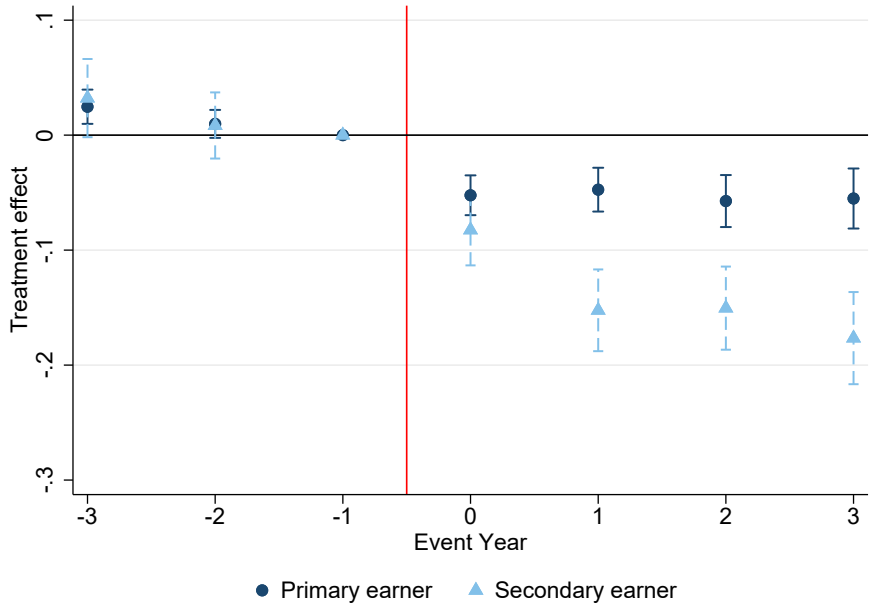


Figure 2.4: Treatment Effect for Income

Notes: This figure shows the results from estimating Equation 2.3 for the sample, which includes all same-sex couples that first filed jointly in 2013 and did not retroactively file taxes jointly. The graph represents the estimated coefficients from the interaction term (β_k) for the primary earners' and secondary earners' income. Primary and secondary earner are defined on the basis of the pre-reform years. $t = 0$ is the reform year of 2013. Income is the sum of income from business, self-employment and employment. Vertical segments are 95% confidence intervals.

significant effect in the pre-reform period. In fact, Figure 2.4 shows that the pre-reform estimates are not statistically significantly different from zero for secondary earners. Income significantly decreases by up to 17% for a same-sex secondary earner after filing taxes jointly post-reform relative to a different-sex counterpart. As already observed in Figure 2.3a, same- and different-sex primary earners seem to be on slightly different trends in $t = -3$, indicated by the small positive coefficient in $t = -3$ that is statistically significantly different from zero in Figure 2.4. Similar to same-sex secondary earners, same-sex primary earners experience a reduction in income, but smaller in magnitude. Same-sex primary earners' income decreases by 5% after filing jointly post-reform relative to different-sex primary earners.[16] Given that marginal tax rates fall for primary earners after the introduction of joint taxation for same-sex couples, the negative estimate could indicate that the income effect is larger than the substitution effect for the primary earner as these two effects go in the opposite direction for primary earners.

We include initial income levels in further analyses to explore whether our findings are

[16]Figure 2.A.2 illustrates the results from our dynamic Difference-in-Differences approach employing the alternative definition for the primary and secondary earner based on the baseline year. Results are relatively similar except for a small negative pre-trend in $t = -3$ for primary earners and slightly larger coefficients for secondary earners post-reform.

robust. Results remain relatively similar.[17] Moreover, as already explained above, we plan to explore whether our results are robust to expanding our donut-hole approach and focusing on the role of children.

The previous literature has primarily focused on the extensive margin of labor supply. Isaac (2023) and Kalíšková (2014) study intensive margin responses and both estimate insignificant small responses. In contrast, we find substantial responses from secondary earners. Kalíšková (2014) argues that the small effect on the intensive margin of female labor supply can be attributed to the low availability of jobs offering flexible working hours in the Czech Republic during the early 2000s. This should be less of a concern for Western European countries, including Germany. Furthermore, Isaac (2023) suggests that the use of cross-sectional data and measurement errors in self-reported hours worked could be reasons why it might be difficult to precisely estimate hours responses in his setting. We employ administrative micro tax data in a panel format for which a majority of variables is third-party reported. Hence, we are able to precisely estimate the responses. Additionally, the joint taxation system in Germany offers clear tax benefits. Hence, both the income and substitution effect imply disincentives for the secondary earner to work. This is not as clear for the US and could partly explain our relatively larger estimates compared to Isaac (2023).

Partner pay gap It is important to evaluate the impact of the introduction of joint taxation also on the couple level to study the intra-household income distribution. We define the partner pay gap as the absolute difference between partners' incomes before taxes relative to the total income before taxes of the couple. Notably, the advantage of this concept is that we do not need to define primary and secondary earner. Figure 2.5a shows the partner pay gap for different- and same-sex couples around the reform year. The pre-trends are stable and very similar for both groups. In the reform year $t = 0$, the partner pay gap trends diverge for the two groups. Same-sex couples' partner pay gap significantly widens, whereas different-sex couples' partner pay gap follows the slight upward trend from the pre-reform period. This can also be seen in the dynamic Difference-in-Differences results, illustrated in Figure 2.5b. The coefficients are around zero for the pre-reform years. The partner pay gap then increases by up to 14 percentage points three years after the introduction of joint filing for same-sex couples relative to the partner pay gap of different-sex couples. The stable pre-trends and the substantial widening of the partner pay gap in the reform year indicate that couples for which primary and secondary earner are not consistent over the years should not be problematic for our identification of the treatment effect. The previous literature has only studied individual responses and has not looked at the couple level so far. An exception is Koch (2024) who studies a move to the individual withholding income tax schedule for newlywed different-sex spouses in Germany and finds a narrowing of the partner pay gap. Our findings can give additional insights and demonstrate that joint taxation can have substantial effects on intra-household inequality.

Discussion Our results from the dynamic Difference-in-Differences approach offer new insights on the effect of joint taxation. This holds true regardless of two factors:

[17] Results are available upon request.

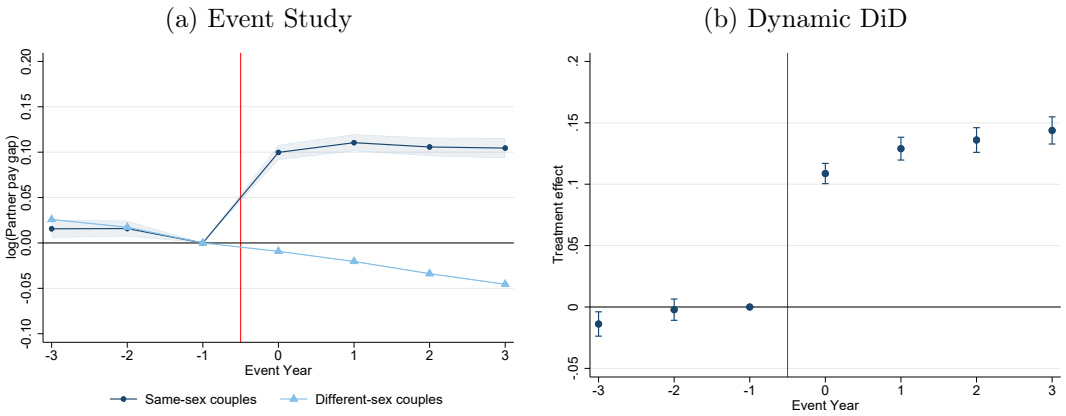


Figure 2.5: Partner Pay Gap Relative to Reform Year and Treatment Effect

Notes: This figure shows the partner pay gap relative to the reform year and the treatment effect. Panel (a) shows event study results from estimating Equation 2.3 for the partner pay gap without the interaction term $D_t \cdot T_i$. Panel (b) shows the results from estimating Equation 2.3 for the partner pay gap. The graphs represent the estimated coefficients from the interaction term (β_k). The sample includes all same-sex couples that first filed jointly in 2013 and did not retroactively file taxes jointly. $t = 0$ is the reform year of 2013. Income is the sum of income from business, self-employment and employment. Shaded areas and vertical segments are 95% confidence intervals. Confidence intervals of different-sex couples in panel (a) are very narrow due to the large sample size.

First, our analysis does not differentiate between income and substitution effects, and second, we (so far) do not scale the effects by the change in MTRs (i.e. estimate an ETI). As explained in Section 2.3, same-sex couples are much less affected by traditional gender norms than different-sex couples. Our findings indicate that the tax incentives implied by the introduction of joint taxation for same-sex couples cause substantial responses. Same-sex secondary earners decrease their income significantly and persistently. Moreover, the partner pay gap widens considerably. Consequently, our results suggest that gender norms among different-sex couples may not be the only factor for the low labor supply of women, who are typically the secondary earner, in countries with joint taxation, but that joint taxation plays a major role.[18]

Next steps We plan to explore heterogeneous effects of the introduction of joint taxation for same-sex civil partners. Does the reform have a different impact on female and male same-sex civil partners? Additionally, how do same-sex civil partners with children respond? This is likely to be linked, given that female same-sex civil partners are more likely to have children.

Furthermore, we would like to investigate the reform's effect not only on the intensive margin but also on the extensive margin of labor supply. We plan to proxy the extensive margin with an indicator equal to 1 if same-sex couples have positive income. The previous literature has primarily focused on the extensive margin of labor supply and find significant responses. Therefore, it would be beneficial to be able to put estimates from our setting in context to the other studies.

[18] Tax incentives of joint taxation and gender norms could interact for different-sex couples, making it challenging to identify both channels individually for this group.

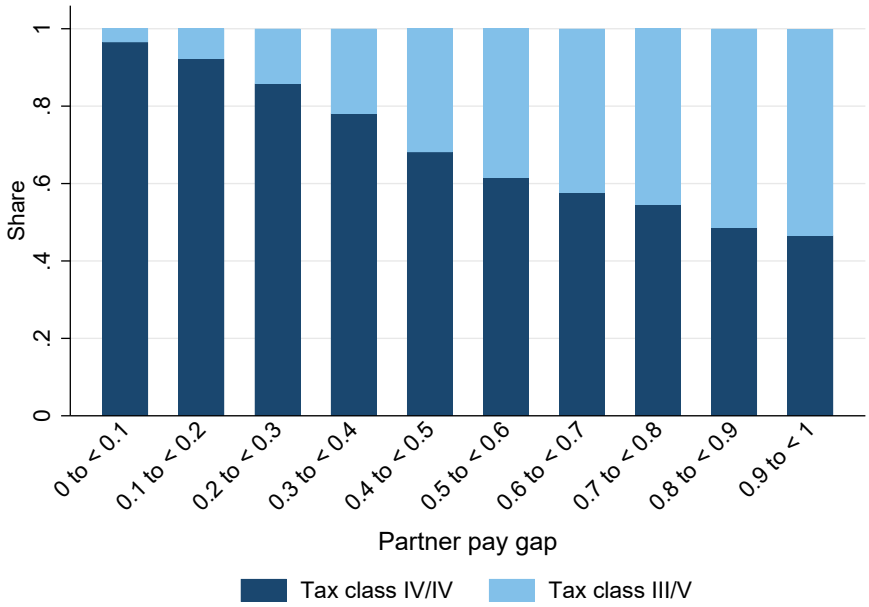


Figure 2.6: Choice of Tax Class Combination of Same-Sex Civil Partners

Notes: This figure shows the chosen tax class combination of same-sex civil partners along the partner pay gap distribution for 2017. The partner pay gap is defined as the absolute difference between partners' incomes before taxes relative to the total income before taxes of the couple. Tax class combinations III/V and V/III are combined and represented as III/V.

2.7 Tax Planning of Same-Sex Couples

As explained in section 2.2, the German withholding income tax systems allows spouses to shift individual withholding tax burdens among partners to already benefit from income splitting during the year. This may reinforce the unequal allocation of net labor income and consequently affect the household bargaining power of partners as well as the intra-household inequality. Koch (2024) provides evidence for this and shows that withholding income tax schedules do matter for labor supply. Specifically, she finds that a default introduction of the individual withholding tax schedule for different-sex newlyweds leads to significantly higher female labor income and a narrowing of the partner pay gap. The joint withholding income tax schedule option was not simultaneously introduced for same-sex civil partners as joint taxation in 2013, but only in 2015. Thus, our treatment effects in Section 2.6 are not affected by responses to the withholding income tax. For now, we explore whether same-sex civil partners use this tax planning option from 2015 onwards to shift withholding tax burdens.

Figure 2.6 plots the shares of the chosen tax class combinations by same-sex civil partners for the year 2017 against the relative partner pay gap. We only look at same-sex couples for which both partners have positive labor income as only for these couples tax classes are relevant for both partners. Additionally, we do not distinguish between tax class combinations III/V and V/III as the difference is only relevant for different-sex

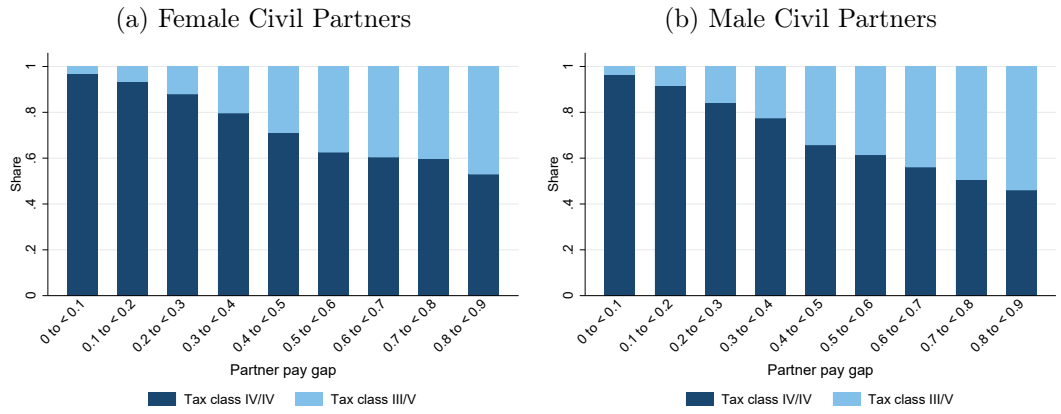


Figure 2.7: Choice of Tax Class Combination of Same-Sex Civil Partners by Gender

Notes: This figure shows the chosen tax class combination by partner pay gap for (a) same-sex female civil partners and (b) same-sex male civil partners for the year 2017. The partner pay gap is defined as the absolute difference between partners' income before taxes relative to the total income before taxes of the couple. The shares for partner pay gaps "0.9 to $<$ 1" were not released by the Federal Statistical Office due to the small sample size. Tax class combinations III/V and V/III are combined and represented as III/V.

couples.[19] Figure 2.6 illustrates that a same-sex primary earner is more likely to have the favorable tax treatment of combination III/V when the partner pay gap is larger.[20] This is very similar for female and male civil partners as shown in Figures 2.7a and 2.7b. Female same-sex couples opt for tax class III/V slightly less frequently than male same-sex couples. This is in line with the literature finding that female same-sex couples specialize the least and might be less willing to increase the intra-household inequality (see e.g. Van der Vleuten, Jaspers, and Lippe 2021).

Buettner, Erbe, and Grimm (2019) study tax planning of different-sex spouses using cross-sectional tax return data for Germany for the year 2004. They document that there is a significant number of households that do not minimize taxes. Couples tend to abstain from tax planning if this has a stronger negative impact on the after-tax income of the secondary earner. Although distributional concerns matter irrespective of which of the spouses is the primary earner, they show that couples where the husband is the primary earner are more likely to choose the joint withholding tax schedule, implying a lower (higher) net income for the wife (husband). We follow their approach and use cross-sectional tax return data for the year 2017. The pattern observed for different-sex couples by Buettner, Erbe, and Grimm (2019) remains very similar over the years and is illustrated in Figure 2.8.

Relative to different-sex couples with a clear male primary earner, same-sex couples are more likely to choose the tax class combination IV/IV, which results in a lower marginal tax burden for the secondary earner, irrespective of the size of the partner pay gap. In contrast, the choice of tax class combination IV/IV is similarly correlated

[19]The husband's tax class is placed first. Hence, e.g. for the tax class combination III/V, the husband has tax class III and the wife V.

[20]The number of observations can be found in Table 2.A.1.

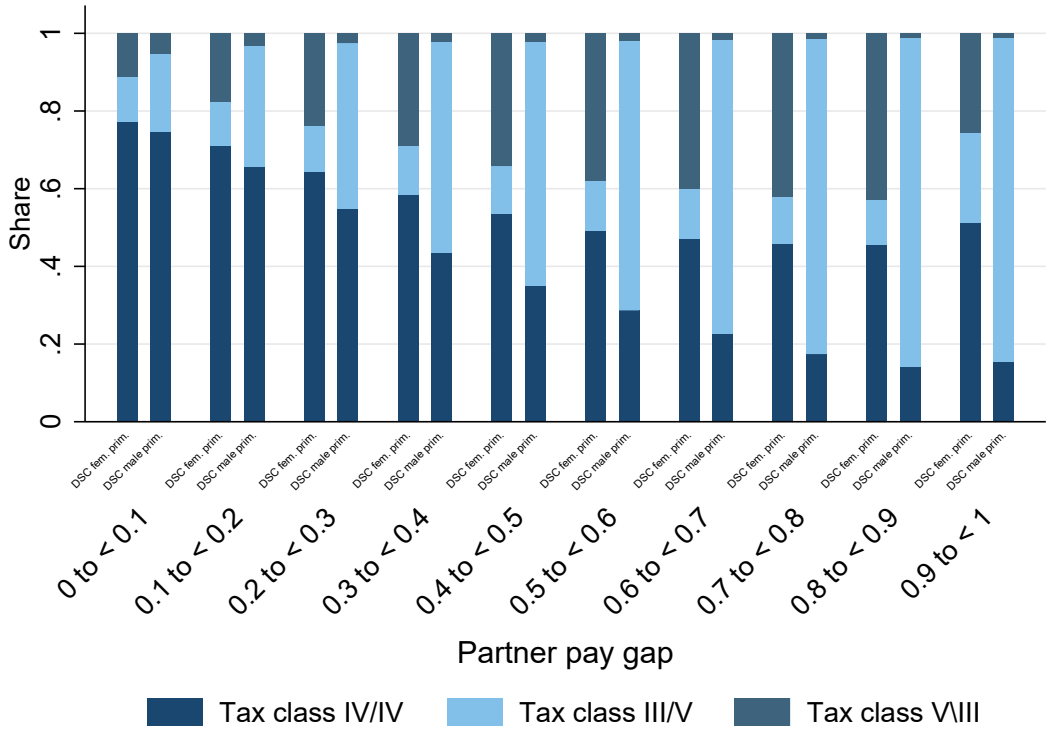


Figure 2.8: Choice of Tax Class Combination of Different-Sex Couples

Notes: This figure shows the chosen tax class combination of different-sex married couples with a female primary earner (DSC fem. prim.) and with a male primary earner (DSC male prim.) along the partner pay gap distribution for 2017. The partner pay gap is defined as the absolute difference between partners' income before taxes relative to the total income before taxes of the couple.
Source: Cross-section of income tax return statistic (FAST 2017).

with the partner pay gap for same-sex civil partners as for different-sex spouses with a female primary earner. Hence, same-sex couples intra-household bargaining power and potential labor income responses may be less affected by the withholding income tax schedule as they choose the joint schedule less often compared to different-sex couples (Koch 2024).[21] Moreover, this descriptive finding could suggest that different-sex couples for which the husband is the primary earner might be affected by gender norms when choosing the withholding income tax class combination.

2.8 Conclusion

Joint taxation of spouses is often debated as a determinant behind the gender earnings gap since it implies negative income incentives for the secondary earner, typically the wife. However, quantifying the impact of joint taxation on income has been challenging due to the lack of exogenous variation and the interaction with traditional gender norms.

[21] The majority of different-sex couples are male primary earner couples who choose the joint withholding tax schedule more often.

We overcome these challenges by exploiting the introduction of joint taxation for same-sex couples in Germany in 2013. We use novel, uniquely linked, administrative income tax return data on the universe of all same-sex couples who file jointly and employ a Difference-in-Differences approach for identification. We find that a same-sex secondary earner experiences a significant decline in their income after filing taxes jointly relative to a different-sex secondary earner and relative to prior to the introduction of joint taxation for same-sex couples. Additionally, the partner pay gap significantly widens three years after the reform relative to different-sex couples. The results show that joint taxation indeed negatively impacts the income of the secondary earner. We shed light on the interaction of tax incentives and gender norms as same-sex couples are arguably less affected by gender norms. Our findings suggest that gender norms are not the only reason for low labor supply of women, who are typically the secondary earner, but that joint taxation plays a major role. Furthermore, we explore whether same-sex couples engage in tax planning in the form of shifting the withholding tax burden among partners. We find that same-sex civil partners are, irrespective of the size of the partner pay gap, less likely to choose a withholding income tax schedule, which results in a higher marginal tax burden for the secondary earner, compared to different-sex spouses with a male primary earner.

Policy implications Governments worldwide are facing the challenge of addressing the shortage of skilled workers. Abolishing joint taxation of spouses is currently debated in several countries such as Germany and Switzerland to boost labor supply, particularly among secondary earners, who are often women. This study not only offers the opportunity of further understanding the particular economic dynamics and challenges experienced by same-sex couples, but can also inform policy-makers about the economic costs of joint taxation and its detrimental impact on gender equality within the labor market.

Appendix to Chapter 2

2.A Additional Results

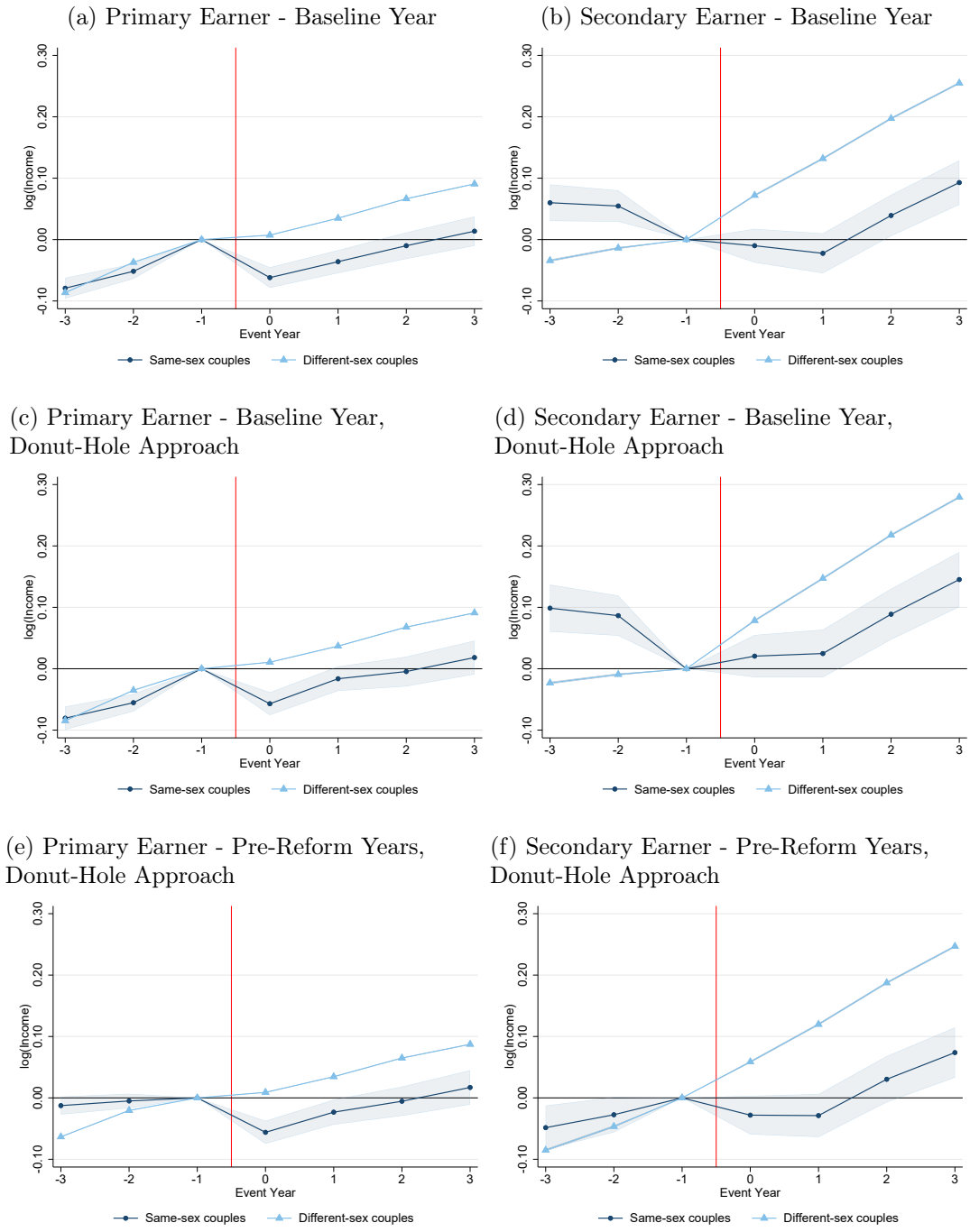


Figure 2.A.1: Income Relative to Reform Year for Different Specifications

Notes: This figure shows event study results from estimating Equation 2.3 for the secondary and primary earner without the interaction term $D_t \cdot T_i$. The sample includes all same-sex couples that first filed jointly in 2013 and did not retroactively file taxes jointly. Panel (a) and (b) show the results for primary and secondary earners, respectively, defining the primary and secondary earner based on the baseline year $t = -1$. Panel (c) and (d) employ the same definition as Panel (a) and (b) and exclude civil partners whose incomes differ by less than +/- 5%. Panel (e) and (f) show the results for primary and secondary earners, respectively, defined on the basis of all pre-reform years as in Figure 2.3 and exclude civil partners whose income differs by less than +/- 5%. $t = 0$ is the reform year of 2013. Income is the sum of income from business, self-employment and employment. Shaded areas are 95% confidence intervals. Confidence intervals of different-sex couples are very narrow due to the large sample size.

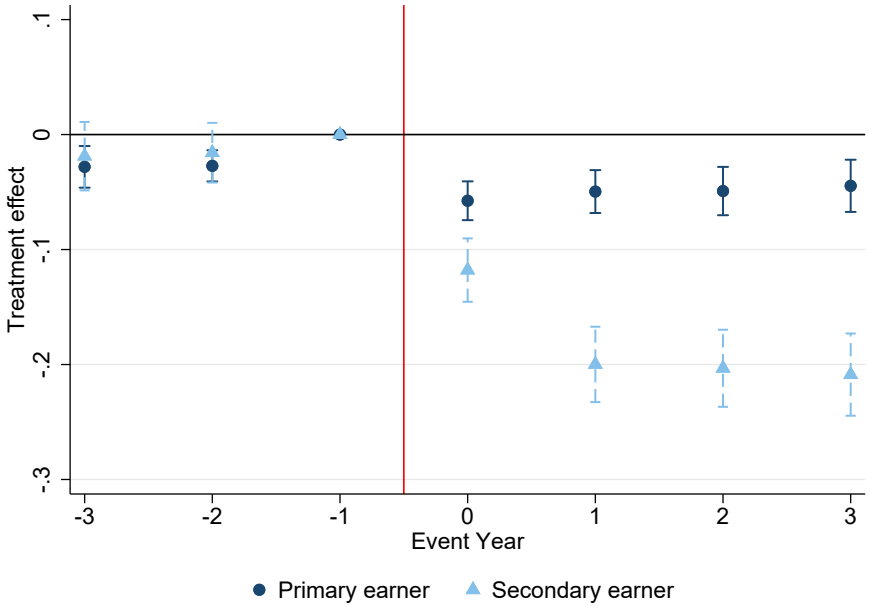


Figure 2.A.2: Treatment Effect for Income - Baseline Year

Notes: This figure shows the results from estimating Equation 2.3 for the sample, which includes all same-sex couples that first filed jointly in 2013 and did not retroactively file taxes jointly. The graphs represent the estimated coefficients from the interaction term (β_k) for the primary earners' and secondary earners' income. $t = 0$ is the reform year of 2013. Primary and secondary earner are defined in the baseline year. Income is the sum of income from business, self-employment and employment. Vertical segments are 95% confidence intervals.

Tax class combination	IV/IV	III/V
Partner pay gap	Observations	
0 to $<$ 0.1	4,741	198
0.1 to $<$ 0.2	3,921	378
0.2 to $<$ 0.3	3,049	573
0.3 to $<$ 0.4	2,225	712
0.4 to $<$ 0.5	1,551	842
0.5 to $<$ 0.6	1,048	782
0.6 to $<$ 0.7	716	630
0.7 to $<$ 0.8	591	584
0.8 to $<$ 0.9	434	567
0.9 to $<$ 1	395	564

Table 2.A.1: Choice of Tax Class Combination - Number of Observations

Notes: This table shows the chosen tax class combination by the partner pay gap for the year 2017, illustrated in Figure 2.6.

Chapter 3

The Impact of Parental Leave Benefits on Pre-Birth Earnings*

3.1 Introduction

Parental leave benefit policies exist around the world (OECD 2022). These policies can have important economic impacts on both parents and children. They, for example, can affect post-birth female labor supply and earnings in the long run as examined by various studies (e.g. Asai 2015; Canaan 2022; Dahl et al. 2016; Lalive and Zweimüller 2009; Olivetti and Petrongolo 2017; Schönberg and Ludsteck 2014). Parental leave policies also affect the incentive to work prior to birth as the amount of the benefit is linked to pre-birth earnings.[1] Therefore, parental leave policies may affect pre-birth earnings of expecting mothers. Understanding these effects is crucial as they can hold substantial implications for government expenditure in the form of higher benefit payouts. These effects remain largely unexplored in the previous literature. This paper aims to fill this gap.

To do so, we study how expecting parents adapt their labor income in the pre-birth year in response to parental leave benefits. We exploit quasi-experimental variation from the introduction of the paid parental leave benefit in Germany in 2007. Studying the German reform is interesting for several reasons. First, the parental leave benefit imply incentives for expecting parents to increase earnings as they are linked to earnings prior to birth. The implied incentives are substantial, which provides us with large identifying variation, enabling us to apply a Difference-in-Income Trend approach for identification. Second, the benefit can be increased via two channels: by increasing labor supply or by leveraging tax planning. Hence, we can observe both labor supply and tax planning responses simultaneously. Third, we leverage unique and novel data encompassing the universe of German taxpayers. Importantly, the relevance of our findings extends to a range of transfer policies. A multitude of other countries' parental benefits and wage replacement transfers, such as unemployment benefits, share a similar design, relying on pre-event income as a determining factor. This highlights the relevance of our measured anticipation effects for understanding and designing effective social and labor market interventions.

We show three main results: We first document novel stylized facts on the labor market behavior of women in the pre-birth year. Expecting mothers reduce their earnings in the pre-birth year, especially at the bottom of the income distribution. Second, the introduction of the parental leave benefit leads women at the bottom to decrease their

*This chapter is based on co-authored work with Ulrich Glogowsky, Amelie Grosenick, Andreas Peichl and Dominik Sachs.

[1]Table 3.A.1 presents an overview of OECD countries' parental leave benefit systems.

pre-birth earnings less and this effect is largely driven by fewer women exiting the labor force in the year preceding childbirth. Third, when descriptively looking at the alternative channel to maximize the benefit (i.e. tax planning), we find that a minority of expecting parents exploit this option. The benefit's complex incentive structure is likely to play a role here.

The German setting allows us to derive these findings since it has a couple of helpful features. In 2007, the German government introduced a new parental leave benefit, replacing the previous means-tested child-rearing payment which had provided a flat benefit independent of pre-birth earnings. The new German parental leave benefit is tied to the average net earnings of the 12 months preceding childbirth up to a specific benefit cap. One additional Euro net income in the pre-birth year implies up to one additional Euro of the parental leave benefit. Hence, this policy significantly reduces the effective marginal tax rate on pre-birth income for expecting parents planning to take parental leave. The resulting changes in effective marginal tax rates during the pre-birth year are very large and amount up to 50 percentage points for many expecting parents. This surpasses the variation used in previous studies estimating income responses to effective tax rates. Importantly, we can exploit substantial variation in the size of the implied pre-birth incentives as they vary significantly with pre-birth income.

Expecting parents can increase their pre-birth net income through two channels. Firstly, they can adjust their labor supply in the 12 months before birth, thereby increasing their gross and net income and consequently their benefit. Secondly, even with gross incomes held constant, the parental benefit's design allows married couples to exploit tax planning to increase their benefit due to a peculiarity of the German income tax. Notably, the parental leave benefit depends on the monthly income net of the monthly withholding income tax. Couples in Germany have the option to choose between two different withholding income tax schedules. While this tax class choice does not alter the ultimate annual income tax burden for the couples, it enables them to redistribute the monthly withholding income tax burden between both spouses and lowers the overall monthly withholding tax burden of the couple. Consequently, the parent taking parental leave can increase the benefit by shifting parts of the withholding income tax burden to the other spouse. Hence, the introduction of the parental benefit implied that tax class choices suddenly have real economic consequences.[2]

We leverage administrative micro data from German tax authorities, specifically the *Taxpayer Panel*. Notably, we have access to the universe of German taxpayers filing tax returns. This data set spans the years 2001 to 2018 and provides precise information on individual earnings, taxes paid and withholding income tax schedules. This rich data set allows us to measure labor income responses and tax planning and enables us to estimate the effects on pre-birth earnings along the income distribution. We focus on women, given that almost all mothers take parental leave (Destatis 2014).

To evaluate the impact of the introduction of the paid parental leave benefit on pre-birth earnings, we adopt the approach of Jakobsen and Søgaard (2022). This approach compares income trends of treated and untreated parts of the income distribution before and after a reform, similar to a Difference-in-Differences approach. We estimate

[2]Due to the fact that the tax class choice only matters for the allocation of the withholding income tax across the year and between spouses, it does not have real economic consequences for couples that are not credit constrained and are unitary decision makers.

income growth across the income distribution for the pre-reform period, characterized by stable marginal tax rates, and the post-reform period, marked by a reduction in marginal tax rates for expecting parents due to the introduction of the parental leave benefit. We define a validation region, where marginal tax rates remain stable in both periods, and an identification region, where significant changes in marginal tax rates occur post-reform. Women receiving the benefit minimum or earning an income above the income threshold are included in the validation region, while those earning income in between are part of the identification region. We employ a *Difference-in-Income Trends* estimator by comparing the trend differential in the outcome during the post-reform period within the identification region to that in the validation region. This comparison is juxtaposed with the corresponding differential in the pre-reform period.

Our empirical findings can be summarized as follows. In a first step, prior to the causal analysis, we document that expecting mothers reduce their earnings in the pre-birth year before the reform. This reduction is particularly large for women with incomes below the median income. Turning to the estimation of the causal impact of the introduction of the parental leave benefit, we find that the reform heavily impacted pre-birth earnings. Expecting mothers decrease their labor income less in the pre-birth year following the introduction of the paid parental leave benefit. The magnitude of this effect is substantial. Due to the reform, pre-birth income of expecting mothers increases by up to 60%. This response is predominantly driven by lower-income mothers (with earnings below the median income). Importantly, this is not just a time trend, but persists for alternative pre- and post-reform years. We reveal that a key driver behind this phenomenon is a significant reduction in the number of women exiting the labor force in the year preceding childbirth.

We explore the impact of the introduction of the parental leave benefit on the second channel, i.e. on expecting parents' propensity to engage in tax planning, descriptively. Even holding gross income constant, couples can increase their net income to increase their parental benefit. They can do so by shifting parts of the withholding income tax burden from expecting mothers to fathers. We observe that few expecting parents exploit tax planning in the form of changing their tax schedule in the year prior to childbirth to maximize their parental benefit. We plan to employ a standard Difference-in-Differences approach to determine the causal impact. For this, we will use expecting mothers who do not have an incentive to increase their pre-birth earnings as they face stable effective marginal tax rates as the control group.

Related literature Our paper contributes to different strands of literature. We add to the literature on the impact of parental leave policies on female labor market outcomes (Asai 2015; Canaan 2022; Dahl et al. 2016; Girsberger et al. 2023; Lalive and Zweimüller 2009; Olivetti and Petrongolo 2017; Schönberg and Ludsteck 2014). Notably, previous studies predominantly focus on post-birth female earnings and employment outcomes. The results depend on the duration of the leave period, ranging from positive or insignificant impacts for leaves shorter than one year to negative effects for longer leaves. Analyzing the introduction of the paid parental leave benefit in Germany, we contribute

to this literature in the following dimensions.[3]

First, our study extends the literature by examining pre-birth outcomes. To date, there is only one paper that (as a side result) examines the impact of a parental leave benefit reform on pre-birth earnings. Girsberger et al. (2023) study the introduction of a short mandated maternity leave in Switzerland. Their overall focus lies on post-birth female labor supply and fertility and they additionally study anticipatory labor supply effects. We provide a richer and more systematic analysis. Our approach and data do not only allow us to identify overall effects, but also enables us to examine changes and response margins (extensive and intensive) along the income distribution. Moreover, we discuss how tax incentives can interact with parental leave policies. In addition, we study a longer leave program, the length of which is closer to the OECD average (OECD 2022). While Girsberger et al. (2023) find small and temporary intensive margin responses of labor income after the introduction of the mandate, we estimate sizeable effects on the intensive and extensive margins for expecting mothers at the bottom. Second, we extend the literature by studying an alternative channel for benefit maximization - tax planning. We explore whether expecting parents exploit the option to shift the withholding income tax burden between them.

The second strand of literature this paper speaks to is the growing literature on the optimal design of parental leave policies (Lalive et al. 2013; Zurla 2022). Optimal policy design has predominantly been shaped by considerations of post-birth outcomes. Our analysis introduces a novel perspective by demonstrating that linking benefits to pre-birth earnings impacts labor supply prior to take-up and consequently influences the size of the benefit strongly. In the context of wage replacement benefits, past earnings are often relevant in determining benefit size and duration. Consequently, our findings carry implications for a broader spectrum of wage replacement policies, as they underscore the presence of incentives for (unintended) responses before benefit take-up.

The structure of this paper is as follows. Section 3.2 describes the institutional background. Section 3.3 describes the employed data. Section 3.4 provides descriptive evidence on female pre-birth labor supply. Section 3.5 explains the empirical strategy to estimate the labor income responses and provides evidence on its impact. Section 3.6 presents descriptives on tax planning of expecting parents and outlines our future plans. Finally, section 3.7 concludes.

3.2 Institutional Setting

In 2007, the German government transitioned from a flat child-rearing payment (called *Erziehungsgeld*) to an earnings-related transfer (called *Elterngeld*).

3.2.1 Pre-Reform Regime

The old scheme targeted low-income families. A flat benefit of 300 € per month was offered to eligible recipients for 24 months. An alternative option allowed individuals

[3]Bergemann and Riphahn (2022), Geyer, Haan, and Wrohlich (2015), Kluve and Tamm (2013), and Kluve and Schmitz (2018) find limited effects on long-run female labor force participation, with faster returns to the labor market by mothers eligible to the new parental leave benefit. Additionally, Raute (2019) observes a significant increase in fertility of tertiary-educated women and Cygan-Rehm (2016) documents changes in the timing of higher-order births.

to receive 450 € per month, but for a limited duration of 12 months. Eligibility for the two options required a net household income below 30,000 € (or 22,086 € for the budget option) for the first six months, and below 16,470 € for both options from the seventh month onward. Generally, the relevant income was the sum of all positive incomes (including income from employment, self-employment, business, agriculture, capital) earned by the father in the calendar year prior to birth if the mother was not working during the benefit receipt period - an arrangement applicable to almost all families. Consequently, the benefit did not depend on mothers' pre-birth earnings as illustrated in Figure 3.1a.

3.2.2 Post-Reform Regime

By contrast, the new benefit scheme provides more generous transfers which are linked to mothers' earnings in the 12 months prior to birth and offers universal coverage.[4] The maximum duration of the parental leave benefit is 12 months for one parent, with an additional two months reserved specifically for the other partner, resulting in a total of 14 months for both parents or a single parent. If both parents take at least two months, flexible sharing of the 14-month period is permitted. However, if only one parent takes leave, they are entitled to a maximum of 12 months of the benefit. The reform did not alter the 36 months of job-protected leave and the mandatory maternal leave (*Mutterschutz*) period.[5] Leave payments are determined as a function of an individual's average net income over the 12 months preceding the birth and the replacement rate, which determines the proportion of the net income received as benefit. The net income considered for benefit determination encompasses positive labor income, with deductions made for monthly withholding income tax, monthly solidarity surcharge[6], monthly church tax (if applicable), and monthly social security contributions (see Table 3.A.2).

Figure 3.1a illustrates the parental leave payment and Figure 3.1b the replacement rate as a function of net earnings from the year before birth. The minimum monthly payment is set at 300 € and a maximum cap is defined as 1,800 € per month. For net earnings above the minimum threshold, there is a gradually decreasing replacement rate from 100% to 67% until 1,200 € of monthly net income. For earnings exceeding 1,200 €, the replacement rate is 65% of net income up to a net earnings level of 2,769 €. This amount corresponds to the earnings level linked to the benefit cap of 1,800 €.

The take-up of the parental benefit by mothers is exceptionally high: 98% of mothers claim some leave and 90% opt for the maximum duration of 12 months (14 months for a single parent), resulting in an average leave length of 11.7 months (see Figure 3.A.1a and 3.A.1b). The average monthly benefit for mothers amounts to 700 € (see Figure 3.A.1c).

To qualify for the benefit, recipients must not exceed 30 hours of work per week during

[4] An income cap of 500,000 € was implemented in 2011, beyond which the benefit were no longer disbursed. Subsequently, in 2021, this cap was further reduced to 300,000 €.

[5] Mandatory maternal leave requires six weeks before and eight weeks after birth to be taken off at full pay.

[6] The solidarity surcharge (*Solidaritätszuschlag* in German) is a supplement to the income tax (amounting to 5.5% of the income tax liability during the period we study), reintroduced in 1993 to finance the costs associated with the reunification of East and West Germany.

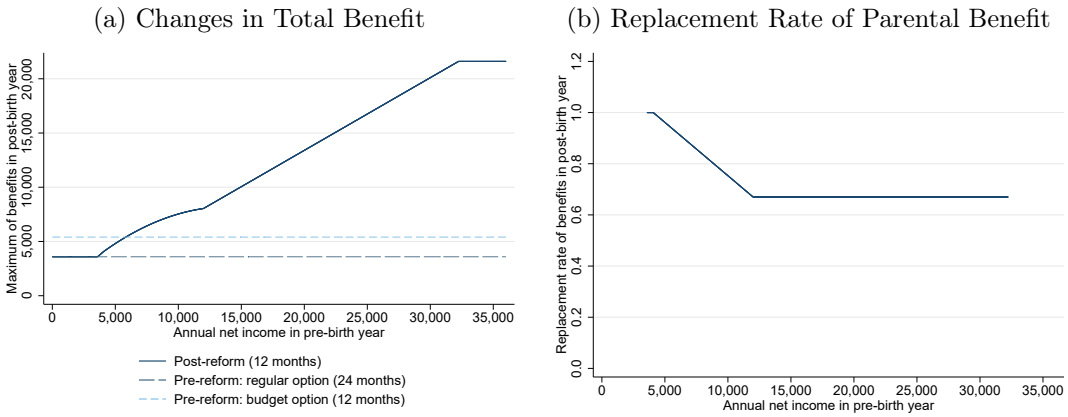


Figure 3.1: The Parental Leave Benefit in Germany

Notes: Panel (a) shows the annual benefit pre- and post-reform given the annual pre-birth net income. Panel (b) illustrates the replacement rate of the post-reform parental benefit. Below the net income of 3,600 € the minimum amount of the benefit is received, i.e. 300 € per month. Above the net income of 32,400 € the maximum amount of the benefit is received, i.e. 21,600 €.
Source: Parental leave benefit law (BEEG), 2007, and child rearing money law (BErzGG), 2004.

parental leave. There is also an option to work part-time and receive a reduced benefit. However, only 0.1% of mother make use of this option (Destatis 2014).[7]

3.2.3 Labor Supply Incentives

The explicit connection between pre-birth net income and the parental leave benefit creates significant incentives for expecting parents to adjust their income. Figure 3.2 demonstrates the incentives to adapt income for individuals eligible for the parental benefit and those who are not. Furthermore, it provides insights into how these incentives vary across the income distribution for expecting parents.

Typically, individuals face the statutory marginal tax rate as defined by the tax code if they are not expecting a child in the next year (illustrated by the dark blue line). This rate starts at zero for a monthly income up to the tax allowance and gradually increases, reaching a maximum of 45% for high incomes. The effective marginal tax rate (EMTR, depicted by the light blue line) follows a distinct trajectory when the parental benefit is factored into the calculation. We make the following thought experiment to derive the EMTR: Consider a mother with income y_{-1} in the pre-birth year. She pays a tax of $T(y_{-1})$ in the same year and receives a benefit of $B(y_{-1})$ in the year of childbirth. The benefit is equal to the net income multiplied with the replacement rate: $(y_{-1} - T(y_{-1})) \cdot RR(y_{-1})$. We assume a replacement rate as shown in Figure 3.1b. The "2-year effective marginal tax rate" is the derivative of $T(y_{-1}) - B(y_{-1})$, given by $T'^{eff}_{-1} = T'(y_{-1}) - B'(y_{-1})$.[8]

[7] There is no tax levied on the parental leave benefit, but the benefit mildly affects the tax rate applied to the couple's taxable income (a peculiarity of the German tax system called *Progressionsvorbehalt*).

[8] For this thought experiment, we subtract from social security contributions and solidarity sur-

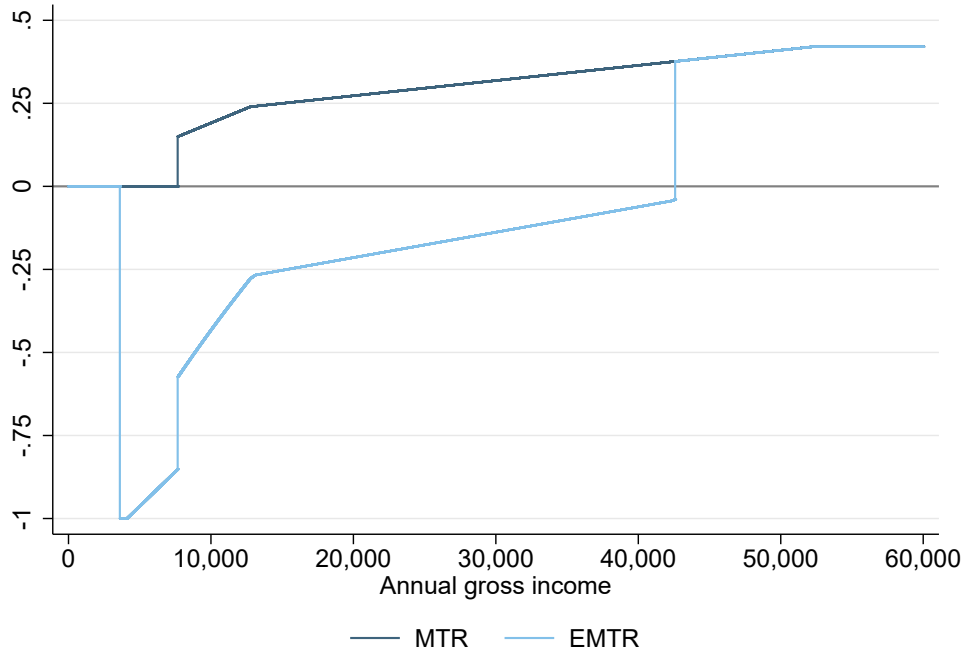


Figure 3.2: Effective Marginal Tax Rate

Notes: The graph shows the statutory marginal tax rate (MTR) and the effective marginal tax rate (EMTR) based on the tax code of 2007 and authors' derivation. Social security contributions and solidarity surcharge are excluded from the derivation.

The EMTR remains at zero for gross incomes up to 300 €, reflecting the tax allowance and the flat benefit of 300 €, irrespective of income. Beyond the 300 € threshold, the EMTR undergoes a significant drop to -1. This implies that each additional Euro of income results in an additional Euro of parental benefit. For instance, individuals with a gross income of 30,000 € (monthly gross income of 2,500 €) would experience a nearly 50 percentage point reduction in the marginal tax rate in the pre-birth year when factoring in the benefit they are eligible for in the subsequent year. The EMTR remains negative up to a gross (net) income of 43,200 € (32,400 €, monthly net income of 2,769 €), where the maximum benefit is reached. This reduction in the EMTR provides an incentive to increase net earnings in the pre-birth year for all individuals below the threshold of 43,200 € (32,400 €) gross (net) income relative to the counterfactual of no parental benefit. Notably, individuals below the lower threshold of 3,600 € gross income may also have an incentive to increase their pre-birth net income. For these individuals, crossing the lower income threshold of 3,600 € implies receiving a higher parental leave benefit than the minimum amount of 3,600 € in the subsequent year. Consequently, it also implies lower EMTRs in the current pre-birth year if they increase their net income.

For our identification, we exploit that the reform does not affect the EMTR for individuals with net incomes above the upper threshold of 32,400 €. We do not consider individuals below the lower threshold of 3,600 € for identification as they also face an

charge.

incentive to increase their pre-birth net income as mentioned above.

The implied shifts in EMTRs during the pre-birth year are substantial and much larger than the variation used in previous studies estimating income responses to income taxes (for recent papers, see, e.g. Bargain, Orsini, and Peichl 2014; Doerrenberg, Peichl, and Siegloch 2017; Kleven and Schultz 2014; Neisser 2021; Saez, Slemrod, and Giertz 2012). Furthermore, they are likely to be large enough that individuals who understood the policy had incentives to adjust their labor supply even though it may imply some non-trivial adjustment costs (Chetty 2012). It is important to note that these incentives are not permanent but are applicable only during the pre-birth year, akin to a temporary tax holiday (see e.g. Martinez, Saez, and Siegenthaler 2021). Also, they prevail for less then a year if couples do not plan ahead before getting pregnant.

3.2.4 Tax Planning of Spouses

As explained above, the parental benefit is increasing in the monthly pre-birth net income. Expecting parents can adapt the pre-birth net income not only by labor supply changes but also by exploiting tax planning.

Germany employs income tax withholding. Married couples have the choice to be taxed individually or jointly for the withholding income tax. The withholding tax a taxpayer faces is determined by the assigned tax class. Spouses can opt for either the individual withholding income tax schedule by assigning both spouses to tax class IV or the joint withholding income tax schedule by assigning one spouse (typically the primary earner) tax class III and the other spouse (typically the secondary earner) tax class V. In the latter scenario, the basic tax allowance from the spouse in tax class V is transferred to the spouse in tax class III. Consequently, the spouse in tax class V (usually the wife) experiences higher average and marginal withholding income tax rates, resulting in a higher individual withholding income tax burden compared to the individual withholding income tax schedule. On the other hand, the spouse in tax class III (typically the husband) faces lower average and marginal withholding income tax rates, resulting in a lower monthly withholding income tax burden. In general, the choice of the tax class primarily influences the taxation of income only on a monthly basis. The withholding income tax is offset against the final income tax when individuals or married couples file their tax returns the following year. Almost all German couples file their final income taxes jointly as only marriage bonuses and no marriage penalties exist in the German system of joint taxation. The parental leave benefit is dependent on the average monthly net income, which is influenced by the monthly withholding income tax. Consequently, the decision regarding the tax class can carry significant financial implications for the parental benefit and married couples encounter substantial incentives for strategic tax planning.

Let us look at an example: The primary earner of a married couple earns 32,000 €, the male median gross income, and is in tax class III. The secondary earner earns 24,600 €, the female median gross income, and is in tax class V. The couple is planning to have a child and decides that the secondary earner will be taking care of the child for the eligible 12 months. The couple has the flexibility to switch their tax classes during the 12 pre-birth months. Consequently, the primary earner would shift to tax class V, while the secondary earner would adopt tax class III. While this alteration would result in a higher withholding income tax burden for the primary earner and the couple

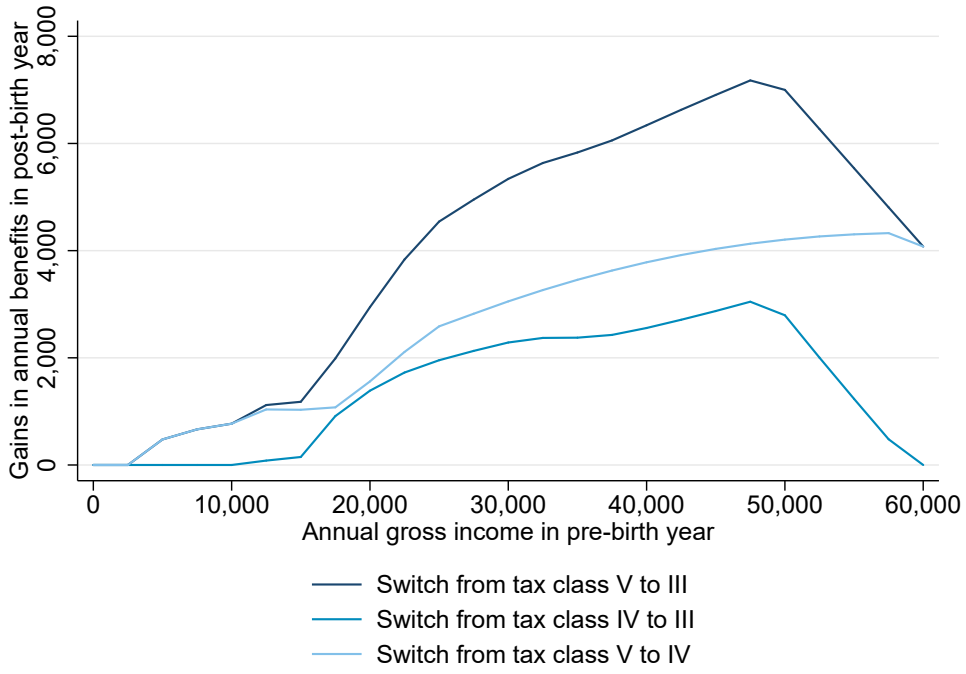


Figure 3.3: Gain in the Parental Benefit due to Tax Schedule Change

Note: This figure shows the potential gains in the parental benefit due to a tax schedule change of married couples in the pre-birth year. It is assumed that individuals change the tax schedule for the full pre-birth year and receive the parental benefit for 12 months. The graph is based on the parental benefit scheme, the tax code, social security contributions and solidarity surcharge from 2007. The dark blue graph shows the gains in the benefit when the tax schedule is changed from V to III, the light blue graph when the tax schedule is changed from V to IV and the bright blue graph when the tax schedule is changed from IV to III. The 32,400 € net income threshold translates to a gross income of approx. 50,000 € for tax class III, of approx. 60,000 € for tax class IV, and of approx. 70,000 € for tax class V.

throughout the year, there would be no changes to the total income tax due at the end of the year. Nevertheless, the secondary earner would experience lower withholding income taxes in the pre-birth year, thereby increasing the net income. Consequently, the eligible parental benefit increases by approx. 5,000 €, 20% of gross earnings (see Figure 3.3). Switches from tax class V to IV or from tax class IV to III would also be beneficial. Importantly, the costs associated with changing the tax schedule are low. Married couples can effectuate this change once a year by simply submitting a letter to their local tax administration, with the new tax class taking effect from the following month onward.

Figure 3.3 illustrates the increases in the annual parental benefit resulting from a switch in tax classes during the pre-birth year across the income distribution. The dark blue graph represents the gains when transitioning from tax Class V to III, the light blue graph from tax Class V to IV, and the bright blue graph from tax Class IV to III. These gains in the benefit increase with gross income, peaking at an annual gross income of approx. 50,000 €, which roughly relates to the net income threshold of 32,400 € of the

parental benefit for tax class III.[9] The most substantial gains are observed when shifting from tax Class V to III. The potential gains in the parental benefit can be significant, reaching up to 7,000 EUR for expecting parents. Even for lower incomes, the gains remain sizeable.

Following the introduction of the reform in 2007, the federal government realized the unintended incentives and promptly instituted guidelines in 2008, explicitly prohibiting the alteration of tax classes by expecting parents solely to maximize the parental benefit (BMFSFJ 2008). Legal challenges proceeded, with parents filing lawsuits against this prohibition, leading to divergent court decisions. The Federal Social Security Court (*Bundessozialgericht*) resolved the matter by ruling against the prohibition on changing tax schedules to increase the benefit in mid 2009 (B 10 EG 4/08 R). This legal discourse suggests a period of uncertainty among expecting parents regarding the legal framework surrounding the change of tax classes in the initial two years following the policy's introduction.[10] Hence, expecting parents may not have taken advantage of the tax planning channel in these early years. Google Trends do not reveal any striking changes in search requests for tax class changes around the reform introduction or for later years indicating that the salience of the tax planning channel was rather low (see Figure 3.A.2).

3.3 Data

In order to analyze whether expecting mothers adapt their pre-birth net income, we use administrative tax return data (*Taxpayer Panel*, TPP). It is provided and administered by the German Federal Statistical Office and the statistical offices of the states (Kriete-Dodds and Vorgrimler 2007). This administrative micro data is obtained by combining income tax returns and tax assessments by the tax authority for the years 2001 to 2018. The unit of observation is the income taxpayer, either a single individual or a married couple filing jointly.[11] In contrast to previous studies using an older version of the TPP data based on a 5% random sample (e.g., Doerrenberg, Peichl, and Siegloch 2017; Dolls et al. 2018), our data comprises the full population of German taxpayers.

We have precise information on taxable income (and its components in terms of income sources), final and withholding income taxes, the tax schedule, and basic socio-demographic characteristics (such as gender, marital status, filing status, year of birth, state of residence, religion, number and year of birth of children). The data set does not include information on the church tax, solidarity surcharge and social security contributions. Therefore, we compute these values ourselves (as we have all the necessary information) to construct the assessment basis for the parental benefit. The data set

[9] The 32,400 € net income threshold translates to a gross income of approx. 60,000 € for tax class IV and of approx. 70,000 € for tax class V.

[10] In 2013, the government amended the parental benefit law concerning changes in tax schedules. The relevant tax schedule for claiming the parental benefit is now the one that applied for the majority of months during the 12 pre-birth months. This implies that expecting parents need to adjust the tax schedule as early as the second month of pregnancy. This timeline may be too early for some couples who may not yet be aware of their pregnancy or have not explored information about the parental leave benefit.

[11] The majority of married couples file jointly in Germany since it is almost always financially beneficial.

includes the year of birth of children and all variables are reported annually. We then approximate the income in the 12 pre-birth months with the income in the pre-birth year. This approach is effective for births occurring earlier in the year. However, for births later in the year the pre-birth year coincides with the birth calendar year. Therefore, the estimated effect on female labor income will on average not cover the full response and should hence be seen as a lower bound.

Sample restriction We construct our main sample as follows. First, we focus on female labor income responses given that nearly 100% of women who give birth opt for parental leave, with more than 90% taking the maximum 12-month leave period (see Figure 3.A.1a and 3.A.1b). Second, our analysis is confined to first births, as the labor supply of women with higher-order births is likely influenced by the presence of a child and the inability to adjust their labor supply accordingly. Our primary focus is on first births occurring between 2005 and 2009, corresponding to the introduction of the parental leave benefit in 2007. Third, we focus on married women for several reasons. The majority of women are either already married or marry around the time of childbirth. Notably, we observe women only after they marry (when filing jointly with their husbands), as we cannot track both single and married spells of the same woman.[12] Additionally, tax planning is only an option for married couples. Our empirical strategy involves comparing trend differentials in the year preceding childbirth to two years before birth. Consequently, our sample comprises married women whom we observe at least two years before the birth of their first child.

Furthermore, our focus is on women with reported labor income (including zero labor income) and no income from self-employment. For self-employed individuals, the parental benefit is determined by income earned in the calendar year before the birth year, rather than the 12 months preceding the birth months, preventing the manipulation of income between years. This feature may make it challenging for self-employed individuals to adjust their income in the relevant pre-birth period.

Summary statistics Table 3.1 reports the summary statistics of a set of individual characteristics for our sample for the pre-birth year for the pre- and post-reform period. Table 3.B.1 reports the summary statistics for the treatment and control group separately. The characteristics are very similar across both periods. This finding suggests that there is at most very little selection into childbirth following the reform. Women are typically around 29 years of age, while men tend to be around 32 years old. The majority of these couples lives in West Germany. On average, the primary earner in a couple is male, and the female contributes to the couple's income with a share of 42%.

[12] Once a married couple files jointly, one spouse (typically the wife) is appended to the other spouse's (typically the husband's) spell. Hence, it is possible to observe both spouses of different-sex couples upon filing jointly, but not single and joint filing spells of the same women. Same-sex couples were only allowed to file jointly in 2013 (Herold and Wallossek 2023). Thus, we focus on different-sex couples.

	Pre-reform 2005	Post-reform 2007
Female age	29.09	29.26
	(11.57)	(4.64)
Male age	32.24	32.41
	(10.69)	(5.37)
Catholic	0.43	0.43
	(0.49)	(0.50)
West	0.91	0.91
	(0.29)	(0.29)
Gross female labor income	22,699	24,347
	(17,057)	(17,778)
Net female labor income	13,356	14,555
	(9,533)	(10,051)
Gross male labor income	33,218	35,336
	(29,909)	(31,638)
Net male labor income	20,485	21,102
	(17,275)	(17,621)
Female income share	0.41	0.42
	(0.28)	(0.27)
N	100,129	96,865

Table 3.1: Summary Statistics

Notes: Summary statistics for expecting parents in the pre-birth year. We show 2005 (pre-reform) and 2007 (post-reform) for the employed sample. 2006 is excluded as expecting parents could not respond yet due to the late passing of the reform. Net labor income is defined as net of withholding income tax. The table depicts the raw mean with standard errors in parentheses.

3.4 Descriptive Evidence on Female Pre-Birth Labor Supply

While it is well documented that women's labor market outcomes are significantly affected after childbirth (Kleven, Landais, and Søgaard 2019), we present descriptive evidence that changes often begin to manifest in the year leading up to childbirth, e.g. during pregnancy. This happens irrespective of the introduction of the paid parental leave benefit. Moreover, we document that this is especially the case for women at the bottom of the income distribution.

3.4.1 Fact 1: Expecting Mothers Decrease Their Income in the Pre-Birth Year, Especially Those at the Bottom

To document women's income trends along the income distribution prior to birth, we categorize women into 2,500 € income bins according to their net income two years prior to the birth of their child ($t = -2$). This includes a separate bin for women with zero income in $t = -2$. We then show how their income changes from this year to the year immediately preceding childbirth. To assess general behavior without any effects of the reform, we investigate the pooled pre-reform period including the years 2003 to 2005.

The average income growth in the general population is positive. However, we find a significant decline in labor income across the entire income distribution for women in their pre-birth year. The decline is much sharper at the lower end of the income distribution, where we find a reduction of almost 85% (see Figure 3.4a). The only exception are women in the first bin. This group shows a large positive income trend. As this bin contains women who report zero income two years before childbirth and then likely reenter the labor market, this is a pattern consistent with mean reversion. The income reductions become smaller in magnitude with increasing earnings up to approx. the median net income of 15,000 €. They remain constant at a negative rate of -11% above the median income. The larger confidence intervals for the higher income bins are due to the lower number of observations in these bins.

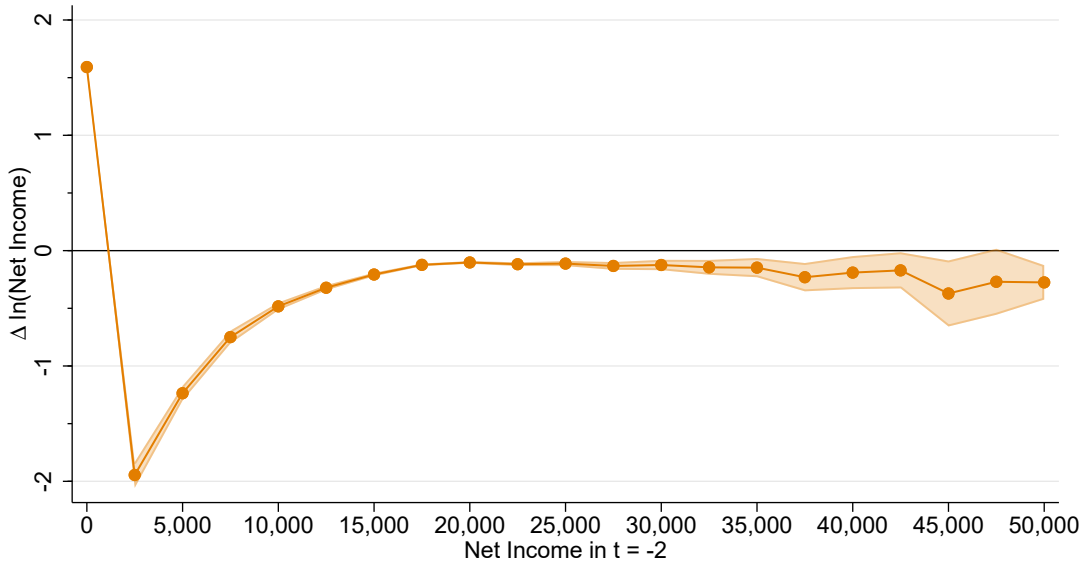
3.4.2 Fact 2: Expecting Mothers at the Bottom Leave the Labor Market

To show descriptive evidence on the share of women leaving the labor market during the pre-birth year, we employ a similar approach as for income trends before. The proportion of women leaving the workforce is defined as those with zero labor income in the pre-birth year before giving birth, despite having positive income in the year before. Figure 3.4b documents that a share of women generally leaves the labor market, finding again heterogeneity along the income distribution. About 37% of women in the second bin, i.e. net income in $t = -2 \in (0, 2500)$, leave the labor market in the pre-birth year, whereas only around 1% do so for incomes above the median net income of 15,000 €.[13]

Especially for the middle and upper part of the income distribution, both of these general trends may be explained by factors such as health considerations or occupational restrictions during pregnancy. Additionally, mandatory maternity leave could

[13] Figure 3.4b does not depict the first bin as this is not defined for the exit rate.

(a) Income Trends across the Income Distribution



(b) Exit Rate across the Income Distribution

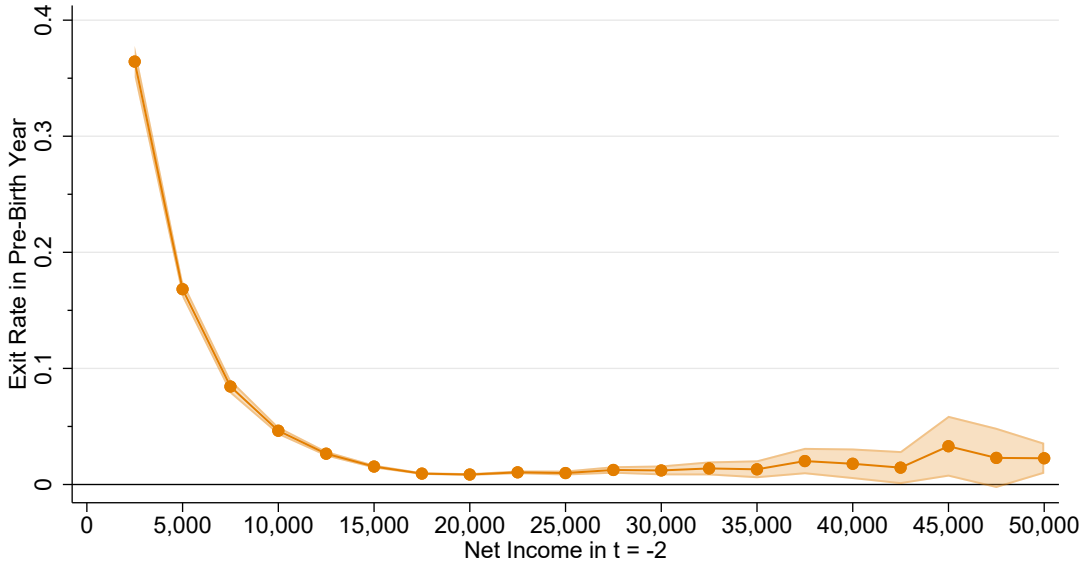


Figure 3.4: Pre-Reform Labor Supply of Expecting Mothers

Notes: Panel (a) shows income trends and panel (b) shows exit rates out of the labor market for pooled pre-reform years, namely 2003-2005. The exit rate is omitted for the lowest (i.e. zero) income bin as it is not defined. Estimates are shown with 95% confidence intervals. Our sample consists of married women in their pre-birth year ($t = -1$). Income is adjusted by the consumer price index.

be measured in the calendar year prior to the birth year if children are born early in the subsequent year due to the lack of birth months of children. But this can only explain a small reduction in income, not an exit. The differential responses observed among women across the income distributions suggest the presence of underlying factors beyond the general challenges encountered during pregnancy. Comprehending these general trends is important to later understand the reform responses to the introduction of paid parental leave.

3.4.3 Fact 3: The Income Change is Primarily Explained by Expecting Mothers Leaving the Labor Market

The described labor income response in Fact 1 includes both intensive and extensive margin responses. To quantify the share of the total income change attributable to the extensive margin, we do the following decomposition exercise:

$$\Delta ln\overline{y}_{kt-1} = \underbrace{-p_k \cdot ln(\overline{y}_{kt-2})}_{\Delta ln(\overline{y}_{kt-1,ext})} + \underbrace{(1-p_k) \cdot (ln(\tilde{y}_{kt-1}) - ln(\overline{y}_{kt-2}))}_{\Delta ln(\overline{y}_{kt-1,int})} \tag{3.1}$$

where $\Delta ln\overline{y}_{kt_{-1}}$ presents the total change in net income per income bin k from $t=-2$ to $t=-1$. We can decompose the total income growth into the extensive, $\Delta ln(\overline{y}_{kt_{-1},ext})$, and intensive margin response, $\Delta ln(\overline{y}_{kt_{-1},int})$. For this, we compute log income changes for both "leavers", who exit the labor market in $t=-1$ and "stayers", who remain in the labor market.[14] We then multiply the log income changes with the exit rate, p_k, and the rate of stayers, $1-p_k$ by income bin, respectively.

Figure 3.5 illustrates the decomposition of the total effect in an intensive and extensive margin response. Notably, the negative income growth during the pre-birth year is fully driven by the response at the extensive margin for expecting mothers at the bottom of the income distribution. Women who remain in the labor market experience positive income growth, aligning with the concept of mean reversion. Hence, this results in more than 100% of the income growth being attributed to the extensive margin for women at the bottom. A similar, yet, nuanced pattern emerges for higher-earning women. Most of the negative income growth can be attributed to women leaving the labor market in their pre-birth year. However, some women who remain in the labor force exhibit also minor income declines, albeit constituting only a small share of the overall response. Consequently, women exiting the labor market serve as the primary explanatory factor for the negative income growth we measure.

3.5 The First Channel: Labor Income Responses

In this section, we focus on the first channel, adjusting gross labor income, and show that the introduction of the parental leave benefit had an effect on expecting mothers' labor income in the year preceding childbirth.

[14] This shortens to - $ln(\overline{y}_{kt-2})$ for the leavers as their income in $t=-1$ is zero. $ln(\tilde{y}_{kt-1})$ is the average income conditional on being strictly positive.

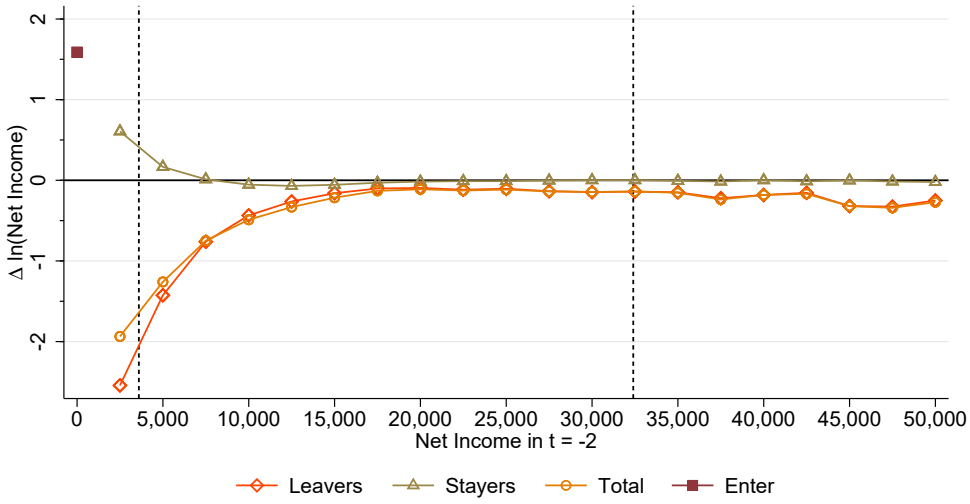


Figure 3.5: Decomposition of the Labor Income Effect

Note: The graph illustrates the change in income with respect to the previous year. Our sample consists of married women in their pre-birth year ($t = -1$) for pooled pre-reform years, namely 2003-2005. The total effect is decomposed into what is driven by the extensive margin (leavers) and intensive margin (stayers). The extensive margin effect that accounts for people entering the labor market is shown as the square. Income is adjusted by the consumer price index.

3.5.1 Empirical Strategy

Intuition of empirical approach To evaluate the impact of the introduction of the parental leave benefit on pre-birth earnings, we adopt the approach of Jakobsen and Søgaard (2022). The approach exploits that some income groups are not affected by the reform. It compares growth rates of an outcome before and after the reform along the income distribution, similar to a standard Difference-in-Differences approach. In the following, we discuss why such a comparison identifies a reform effect. For simplicity, we abstract from income effects when introducing the empirical approach's intuition, however, we discuss them later.

Figure 3.6 illustrates the intuition behind the approach, depicting the income growth across the income distribution for two distinct time periods: a pre-reform period characterized by stable marginal tax rates and a post-reform period marked by a reduction in marginal tax rates for expecting parents due to the introduction of the paid parental leave benefit. In both periods, we represent the income growth as a decreasing function of initial income. Specifically, income trends vary across the income distribution, with individuals at the bottom typically having larger income growth. Those with larger initial income experience lower or on average negative income growth from one year to another as they are more likely to have experienced positive temporary income shocks before.

Subsequently, we compare the trend differentials in the two periods. To that end, we define a validation region, where marginal tax rates remain stable in both pre- and post-reform periods, and an identification region, where significant changes in marginal

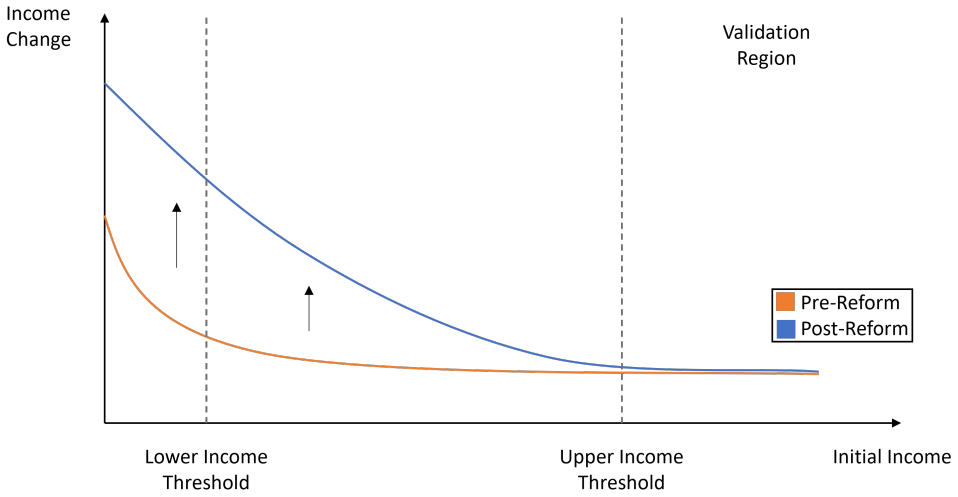


Figure 3.6: Illustration of the Expected Behavioral Response

Note: The graph illustrates the expected change in income in the pre-birth year before (orange line) and after the reform (blue line). The dashed lines indicate the lower and upper income threshold of the parental leave benefit. Individuals in the validation region face constant (effective) marginal tax rates before and after the reform.

tax rates occur post-reform. Figure 3.2 demonstrates stable marginal tax rates for net (and gross) income below 3,600 € and net (gross) income above 32,400 € (43,200 €). In contrast, there are significant changes in marginal tax rates for net incomes between 3,600 € and 32,400 €. Hence, we define individuals with net income above 32,400 € two years prior to birth of the first child as part of the validation region. Women with net income below that threshold are part of the identification region. Women with a net income below 3,600 € are a distinct group in our context: these individuals do not face a change in marginal taxes at their current incomes. But the reform substantially increases the gain in available income in case of surpassing the lower threshold of 3,600 €.

The untreated validation region can be understood as similar to the validation of parallel trends in Difference-in-Differences approaches. While there might be reasons for income changes from one year to another, this pattern in income changes should remain constant across time. Changes in trend differentials can be interpreted as behavioral responses to the reform and should only occur in the identification region due to the changes in EMTRs. In Figure 3.6, the trend differentials in the validation region mirror the same pattern in both the pre-reform and post-reform periods, since the reform does not change incentives for individuals above the income cutoff. Put differently, the growth rates in both periods evolve in parallel. In contrast, in the identification region the income changes are larger for the reform period than those of the pre-reform period. The lower marginal tax rate post-reform incentivizes individuals to earn higher labor income, shifting their line of income changes upward.

Now, we turn to the empirical implementation of our approach. Following Jakobsen and Søgaard (2022), we leverage the pre-reform period as a counterfactual for the post-reform period. To estimate reform effects, we compare the trend differential in the outcome during the post-reform period within the identification region to that in the

validation region and juxtapose it with the corresponding differential in the pre-reform period. This method is a *Difference-in-Income Trends* estimator and closely resembles a standard Difference-in-Differences approach. The Difference-in-Income Trends (Difference-in-Differences) estimator studies how time trends (outcomes) in the identification region (treatment group) and the validation region (control group) vary across time, from before the reform to after the reform. The identifying assumption demands that, without the reform, the growth rates of expecting mothers in both periods would have followed the same trend along the initial income. We can validate the identifying assumption by assessing if the growth rates in the validation region progress in parallel during both periods.

Estimation approach For our analysis, the outcome variable is the income growth from the second-to-last year before birth to the pre-birth year, $\Delta ln(y_{ic}) = ln(y_{ic-1}) - ln(y_{ic-2})$ for individual i and cohort c, indexed by the birth year. For the pre-reform cohort c equals 2006 and for the post-reform cohort c equals 2008. Formally, we firstly estimate the following flexible OLS-regression separately for each cohort:

$$\Delta ln(y_{ic}) = \alpha_c + \sum_{k=1,k\neq 13}^{20} \beta_{kc} \cdot B_{ki} + \epsilon_i \tag{3.2}$$

where $B_{ki} = \mathbb{1}[y_{ic-2} \in (Y_{k-1}, Y_k]]$ is the income bin indicator for bin k based on y_{ic-2}, measured in 2004 for the pre-reform cohort, and 2006 for the post-reform cohort. The bin size is 2,500 €. Women with y_{ic-2}= 0 are included in a separate first bin. We exclude the income bin of [32,500 €, 35,000 €) just above the cutoff. Consequently, α_c captures the income growth in the 32,500 € bin, and the β_{kc} coefficients quantify the differences in income growth for each bin relative to the 32,500 € bin. Subsequently, we can plot β_{kc} across initial net income for both the pre-reform and post-reform period as sketched in Figure 3.6. This plot serves not only to provide graphical evidence of reform effects but also to assess the credibility of the common-income-trend assumption. In the latter context, we examine whether the coefficients progress in parallel within the validation region.

Secondly, we directly estimate the changes in trend differentials between the pre- and post-reform periods, employing a pooled data set of both cohorts. We use the following flexible OLS-regression:

$$\Delta ln(y_i) = \gamma + \sum_{k=1,k\neq 13}^{20} \zeta_k \cdot B_{ki} + \eta \cdot R_i + \sum_{k=1,k\neq 13}^{20} \theta_k \cdot B_{ki} \times R_i + \epsilon_i \tag{3.3}$$

where $R = \mathbb{1}(c = 2008)$ is a reform indicator and the definitions of the other variables remain unchanged. In this estimation, we introduce a vector of income bin indicators interacted with the vector of reform indicators. The dummy for the 32,500 € income bin interaction is again excluded. θ_k is the coefficient vector of interest and presents the Difference-in-Income Trends estimator. It depicts the changes in income trend

differentials in bin k between the pre- and post-reform cohorts relative to the pre- and post-reform change in the baseline bin. Graphically, each θ coefficient illustrates the bin-specific difference between the income changes in the pre- and post-reform period.

3.5.2 Results

Figure 3.7 summarizes our results. The introduction of the parental leave benefit led expecting mothers to adjust their pre-birth net income. This key result emerges from three observations. First, we observe common trends in the validation region, supporting that our identification assumption holds. Second, the pre- and post-reform growth rates diverge in the identification region. Post-reform expecting mothers reduce their labor income in the pre-birth year less compared to their pre-reform counterparts in response to the lower effective marginal tax rates. Third, the effects are the largest for women at the bottom at the income distribution with an initial net income below 15,000 €, which roughly equals the median female net income.

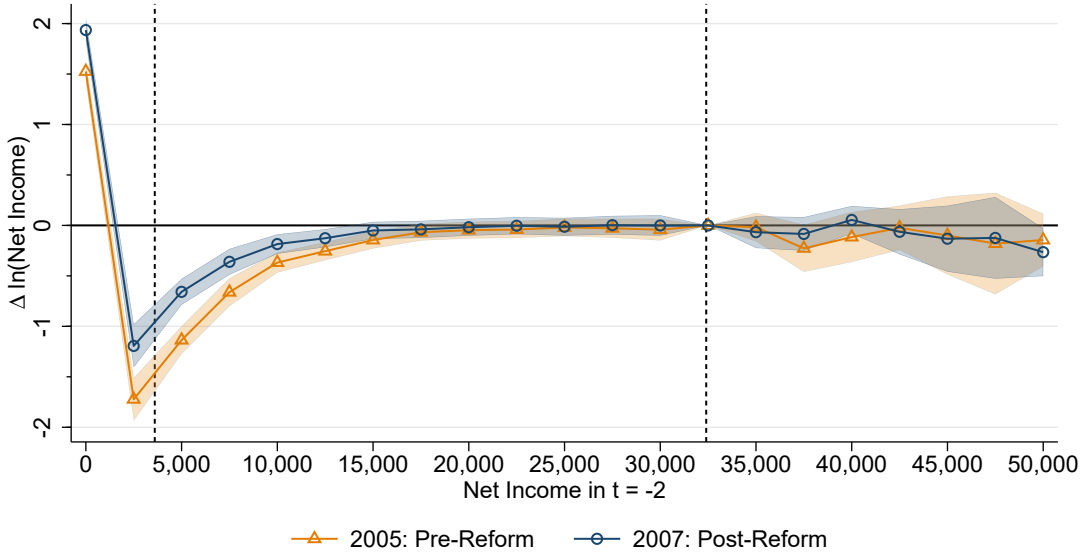
As explained in section 3.5.1, we do not expect the common-income-trend assumption to hold for women earning below 3,600 € net income since those individuals face an incentive to earn above the minimum cap due to the benefit's high replacement rate for low incomes. This is exactly what Figure 3.7a illustrates. Expecting mothers with an initial income of zero (first bin) are the only ones with a positive income trend in the pre-birth year relative to two years before birth. Hence, women with an initial income of zero increased their income by more or a higher share of women entered the labor market in the pre-birth year for the post-reform period. Expecting mothers in the second bin, (0 €, 2,500 €), have a large negative income trend. Notably, we estimate higher income trends in the post-reform period.

Figure 3.7b shows that labor income significantly increases by 65% for expecting mothers with initial net income of [0 €, 5,000 €) in the pre-birth year relative to two years before birth in the post-reform year compared to their counterparts pre-reform. The effect size decreases along the income distribution and becomes statistically insignificant for initial net incomes above 15,000 €. We will show that this is driven by fewer exits out of the labor market of these expecting women (see Section 3.5.3). The magnitude of the increase in labor income is large. The effect is still around 10% of an increase in labor income for expecting mothers with net incomes of 15,000 €. The confidence intervals in the validation region are larger, since there are less women earning net income above the threshold.

These results provide evidence that (low- and middle-income) expecting mothers seem to understand the incentives they are facing in the pre-birth year with respect to adapting their income. As mentioned earlier, the observed responses may be a lower bound since we cannot precisely identify the exact 12 pre-birth months.

To explore heterogeneous effects, we reestimate Equation 3.2 and 3.3 for two specific samples: married couples with a self-employed husband and married couples with the husband's income above the median net income (see Figure 3.C.1 and 3.C.2). We would expect these couples to be more aware of the incentives implied by the parental benefit as they might be advised by a tax consultant. However, we do not find differential responses by these couples.

(a) Trend Differentials across the Income Distribution



(b) Difference in Income Trends across the Income Distribution

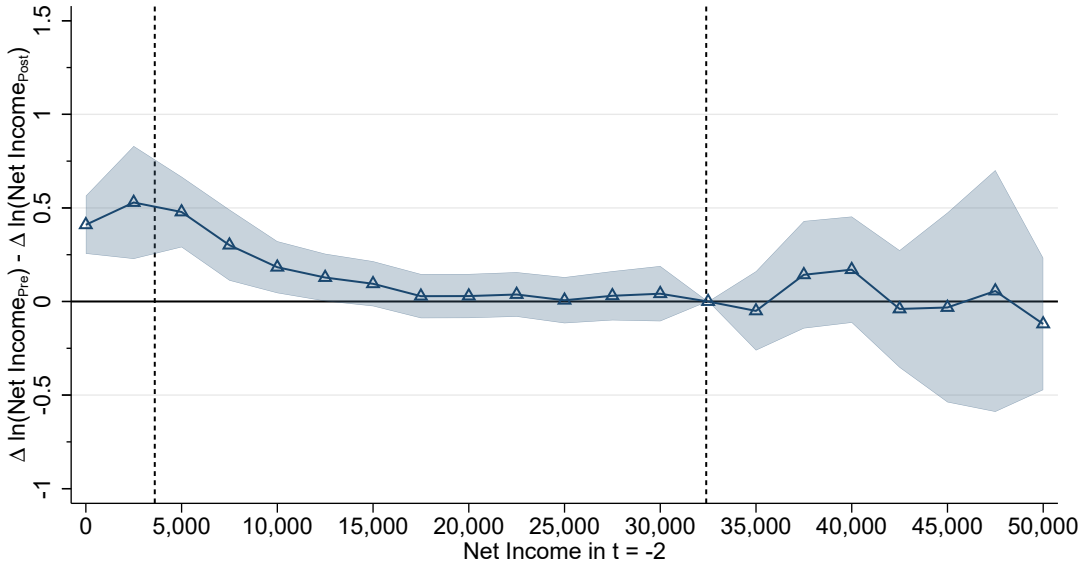


Figure 3.7: Income Changes and their Difference before and after the Reform

Notes: Panel (a) plots estimates of equations 3.2 and panel (b) plots estimates of equation 3.3. Estimates are shown with 95% confidence intervals. Our sample consists of married women in their pre-birth year ($t = -1$). Income is adjusted by the consumer price index.

3.5.3 Mechanism

Next, we explore which factors could drive these substantial effects for low- and middle-income women. In a first step, we assess changes in the extensive margin of labor supply. For this purpose, we analyze whether the probability of expecting mothers' to exit the labor market changes pre- to post-reform. We construct an indicator variable equal to 1 if an individual earns positive labor income in $t = -2$ but has no income in $t = -1$. Subsequently, we estimate Equations 3.2 and 3.3, employing the exit rate as the outcome variable. Our findings, illustrated in Figure 3.8, demonstrate that less women leave the labor market in the year prior to childbirth during the post-reform period compared to the pre-reform period. Figure 3.8 shows common trends in the exit rates in the validation region, supporting our identification assumption. The decrease in labor market exits accounts to up to 5%. Again, it is driven by responses at the bottom of the income distribution. This suggests that the lower EMTRs in the pre-birth year caused fewer expecting mothers at the bottom to discontinue employment. These women are likely to be able to flexibly adjust their labor supply as they may work in marginal employment, so-called *Mini-Job*.[15] Unfortunately, we cannot identify Mini-Jobs in our data and explore this further.

These findings lead to an interesting question: Does the observed effect persists beyond the parental leave period? Specifically, are post-reform mothers more likely to return to work earlier and/or with more hours? Investigating the persistence of these effects could provide valuable insights into the long-term implications of the reform on female labor supply.

3.5.4 Robustness Checks

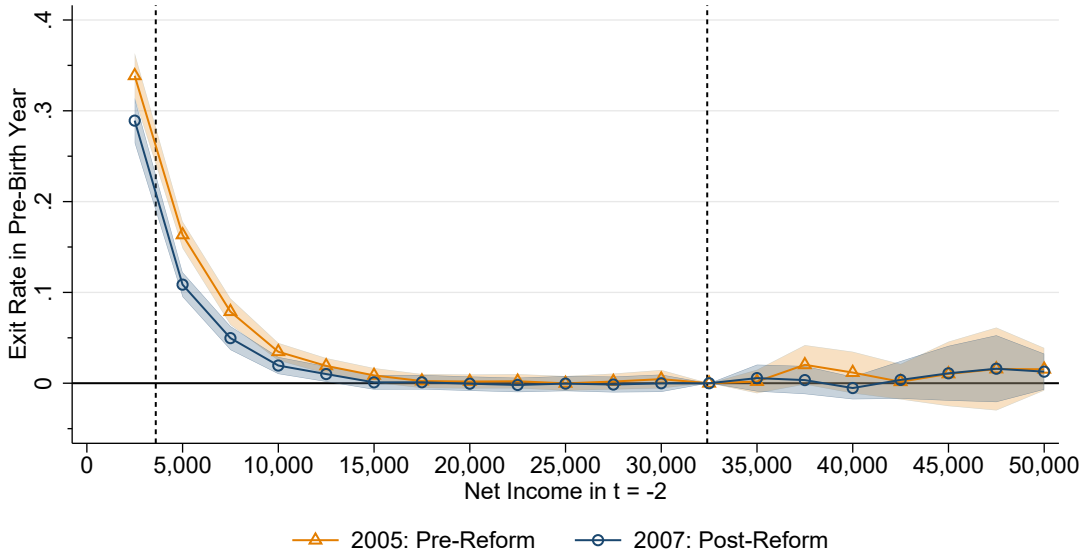
It is not just a time trend One concern that could invalidate our results is that the labor income responses could just reflect a general time trend. To test this, we estimate Equation 3.2 for earlier pre-reform and subsequent post-reform years to validate whether there is a discontinuous change around the reform. Studying additional pre-reform periods allows us to test if the growth rate patterns were constant across the income distribution before the reform. This would indicate that sudden changes in the reform period are unlikely the result of general patterns. Similarly, we examine growth rates in later post-reform periods. This allows us to examine if the reform effects are of a longer-run nature.

The trend differentials are consistent across various pre-reform years. Figure 3.C.3 effectively demonstrates this point. The same holds for both post-reform years 2007 and 2008. The effect in 2008 is even slightly larger than in our baseline year, 2007. This analysis confirms that the observed differences can indeed be attributed to the reform.

Income effects Until now, our analysis has disregarded income effects. Individuals in the validation region above the 32,400 € threshold, however, experience a substantial positive income shock after the introduction of the paid parental leave benefit. Before

[15] In our observed time period, a Mini-Job is an employment with a limited income of up to 400 € per month, subject to no income taxes and social security contributions.

(a) Trend Differentials across the Income Distribution



(b) Difference in Exit Rate across the Income Distribution

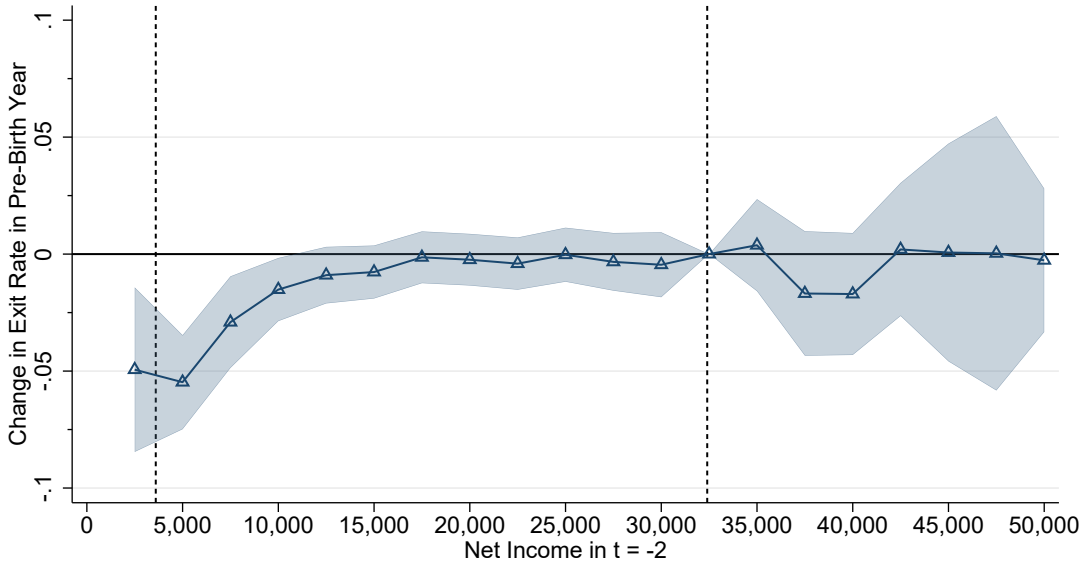


Figure 3.8: Exit Rate Changes and their Difference before and after the Reform

Notes: Panel (a) plots estimates of equations 3.2 and panel (b) plots estimates of equation 3.3 using the exit rate as the outcome variable. The exit rate is omitted for the lowest (i.e. zero) income bin as it is not defined. Estimates are shown with 95% confidence intervals. Our sample consists of married women in their pre-birth year ($t = -1$).

the reform, these women are likely to not have been eligible for the previous benefit due to their spouses' high income.[16] Post-reform, they receive the maximum monthly amount of 1,800 €, translating to up to 21,600 € in the annual benefit, thus signifying a sizeable increase.[17] Expecting mothers could reduce their labor income in anticipation of the forthcoming positive income shock.

Given that our estimated response for the labor supply channel are interpreted with respect to a reference bin, income effects within this bin could introduce bias to our results. If women in this bin respond to the anticipated positive income shock, we would expect them to have stronger declines in their income growth rates post-reform. As income in the pre-birth year is on average declining, the income growth rate would thus be shifted to a lower negative level. By referencing on one income bin in the validation region, we mechanically link the growth rates along the income distribution to that level and, hence, our estimates may be downward biased. The size of the income effect likely varies along the income distribution.[18] We would expect income effects to be largest right above the 32,400 € threshold in the validation region, our reference bin, and decrease in size the further one moves away from it. In our estimation, we thus substract an income effect that is larger than what would be expected in the identification region and the true substitution effects are even larger than what we estimate. Consequently, we underestimate the effects.

To assess whether this is the case, we analyze income growth over time. Income growth could differ across four margins: It can be bin-specific, differ by a woman's individual event-time (t), e.g. with respect to her timing of birth, possess calendar-year cyclicality, and be affected by income effects. Our identification approach addresses the first two factors by taking differences within income bins and by only comparing women in their pre-birth year. The third is addressed in the previous robustness check. To test for income effects, we explore log income changes for women above the 32,400 € threshold because, for this group, the parental leave benefit reform implied only income effects. For this exercise, we only include women we observe four years or more prior to having a first child to be able to observe income trend dynamics. For each calendar year, income changes for women in event-periods $t = -3$ and $t = -1$ are computed and averaged across individuals. These are changes with respect to the previous year and will be referred to as $\Delta_{-4/-3}$ and $\Delta_{-2/-1}$. Within the same calendar years, income changes will differ by event-time, as we documented an average decrease of income in the pre-birth year previously. However, the difference between $\Delta_{-4/-3}$ with respect to $\Delta_{-2/-1}$ should remain constant pre- and post-reform if there are no income effects. This assumes a comparable decrease of labor income when expecting a child pre- and post-reform.

Figure 3.9 presents $\Delta_{-4/-3}$ and $\Delta_{-2/-1}$ along with their 95% confidence intervals for the different birth cohorts of mothers. Income changes for women in their pre-birth year are slightly lower than those of women having children in three years, i.e. not yet

[16] As explained in Section 3.2, the relevant income for the pre-reform benefit was the sum of all positive incomes earned by the father in the calendar year prior to birth if the mother was not working during the benefit receipt period.

[17] In a related setting, Wingender and LaLumia (2017) document income effects among mothers in response to a child-related tax benefit post-birth in the US. Mothers receiving the benefit are less likely to return to work after childbirth.

[18] Our estimator would be unbiased if there were constant income effects over the income distribution.

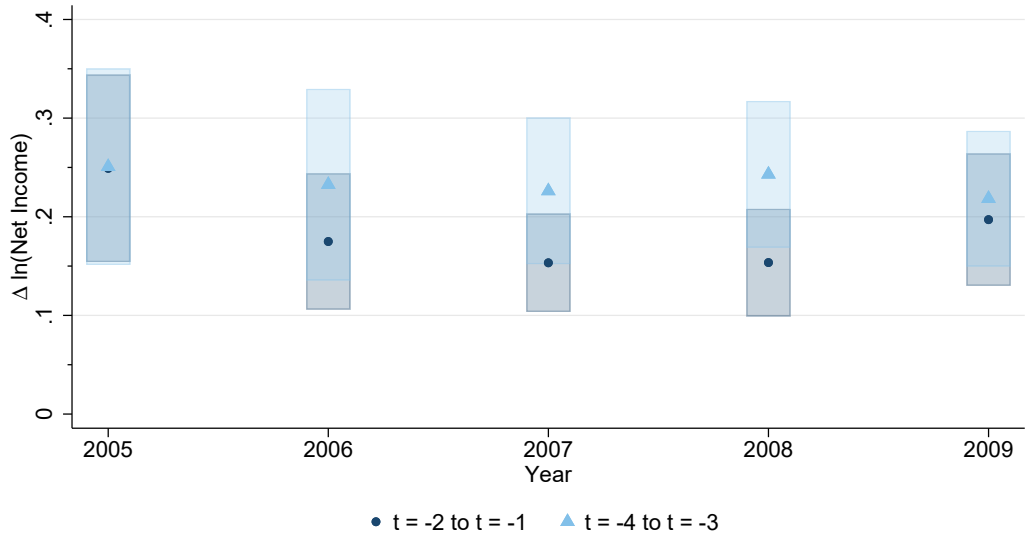


Figure 3.9: Income Changes for Different Cohorts of Expecting Mothers

Notes: The figure shows income changes for our sample of married women for pre- and post-reform cohorts. t measures relative event time. $t = 0$ is the birth year. The x-axis depicts calendar years. Income is adjusted by the consumer price index. Bar areas are 95% confidence intervals.

in their pre-birth year. Comparing the pre-reform years 2005 and 2006 to post-reform years 2007 and 2008 reveals a slight widening between the point estimates for $\Delta_{-2/-1}$ and $\Delta_{-4/-3}$ post-reform. This assessment of income growth rates shows, if anything, very small anticipatory responses to positive income shocks of expecting mothers in our setting. Consequently, we are confident that the potential downward bias implied by income effects on our estimates is minor. We plan to test the difference of the point estimates for statistical significance.

3.6 The Second Channel: Tax Planning

This section explores whether expecting parents engage in tax planning in the year preceding childbirth to increase their benefit and outlines our plans for the future.

3.6.1 Descriptive Evidence

For now, we examine this channel descriptively. Specifically, we investigate whether expecting mothers change their tax class in a beneficial way in the year prior to childbirth, $t = -1$. We define beneficial tax class changes as tax minimizing, i.e. changes into a lower withholding tax class (e.g. from tax class V to IV or III, or from tax class IV to III).

Notably, most married couples remain in their original tax class combination during the pre-birth year, irrespective of the reform (see Figure 3.C.4). Figure 3.10 illustrates the number of beneficial (light blue bars) and not beneficial (dark blue bars) tax class changes of parents expecting their first child for the pre-birth years 2005 to 2012. The

pre-birth years of 2005 and 2006 are both defined as the pre-reform period since the reform's implementation was only finalized in October 2006. Hence, expecting parents could only react from October 2006 onwards. Approx. 60% of tax class changes prior to the reform are beneficial changes. The year 2007, the reform year, records the highest share of couples changing their tax classes in the pre-birth year, both in a tax minimizing way and not - nearly 4,000 expecting parents, or close to 6% of all parents expecting their first child in the next year, adjust their tax classes. Hence, we can see an increase in tax class changes in the reform year, however, it is rather small in relative terms. Additionally, the number of tax class changes declines again in the subsequent post-reform years. Nevertheless, the share of non-beneficial tax class changes consistently decreases to around 30% of all tax class changes after the introduction of the parental benefit. Remarkably, we do not observe a higher number of beneficial tax class changes after 2009 when the legal uncertainties explained in Section 3.2 were resolved. To sum up, the introduction of the parental leave benefit seems to only have had a small impact on the choice of the tax class among parents expecting their first child in the subsequent year.

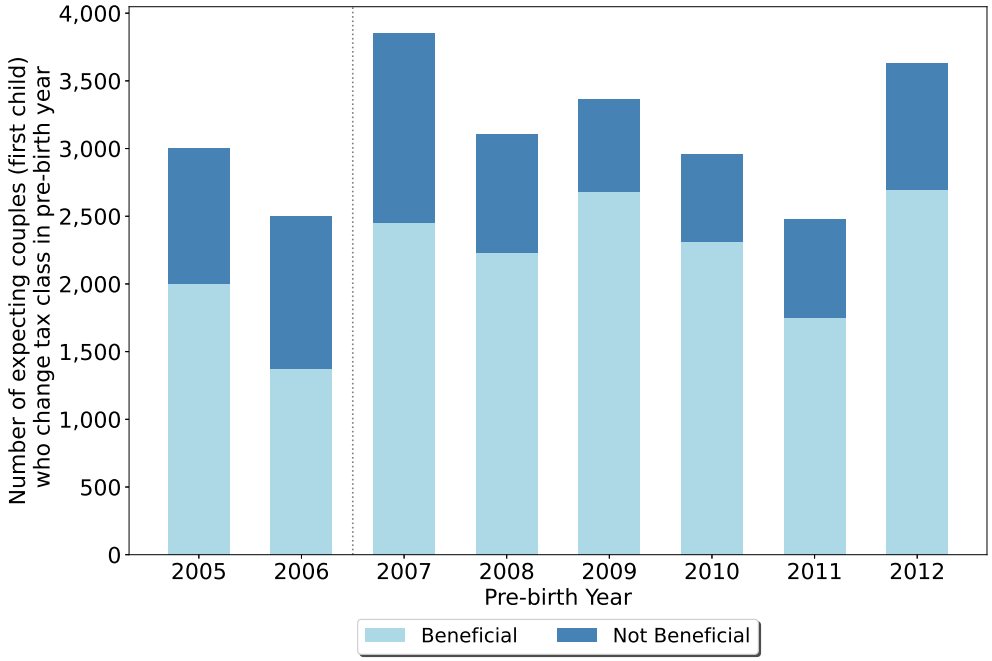


Figure 3.10: Tax Class Changes in Pre-Birth Year

Notes: The figure shows the number of beneficial and not beneficial tax class changes in the pre-birth year for our sample of expecting parents having their first child. The dotted vertical line illustrates the introduction of the paid parental leave benefit in 2007.

Discussion Behavioral biases or information frictions might offer some explanations for the small number of tax class changes we observe. The incentive structure of the parental leave benefit rule is rather complex. The parental leave benefit increases in pre-birth net income in a complicated non-linear way. As a result, the marginal replace-

ment rate for the parental leave benefit is less salient than e.g. the statutory marginal tax rate of the personal income tax. Hence, even expecting parents that generally understand that their benefit will increase in pre-birth income and change the tax schedule may misperceive the exact incentive structure. Abeler and Jäger (2015) and Feldman, Katuščák, and Kawano (2016) find that individuals underreact or even remain passive in complex tax systems. Moreover, Bayer, Simon, and Wegmann (2023) observe that the majority of surveyed individuals cannot distinguish between the withholding and the final income tax in Germany. Further explanations could be inattention as well as inertia of individuals (Jones 2012; Gabaix 2019).

Additionally, even if married couples do understand the complex incentives during the 12 pre-birth months, distributional issues within the household or gender norms might prevent them from acting benefit-maximizing. Buettner, Erbe, and Grimm (2019) show that couples do not engage in tax planning if this will induce larger losses in the after-tax income of the secondary earner. This holds regardless of which spouse has higher earnings. However, couples for which the wife is the primary earner are less likely to select the favorable tax treatment for the wife. Bertrand, Kamenica, and Pan (2015) argue that gender identity norms cause an aversion of the wife being the primary earner. This could also relate to the tax schedule setting - gender norms might induce an aversion of the wife receiving higher net income than the husband.

3.6.2 Next Steps

We plan to causally analyze the effect of the introduction of the parental benefit on tax class changes. The main challenge in determining effects on tax-class switching is that this is a rare event. We, hence, need to adapt an empirical approach that is able to deal with this property. While the approach of Jakobsen and Søgaard (2022) has many great properties (it, e.g., allows to identify effects along the income distribution), it is not optimal to study tax-class switching. Given that the events are rare, we only have very few switches per bin (resulting in very noisy data). We, therefore, plan to employ a classical Difference-in-Differences approach to examine impacts on tax planning. The approach's idea is straightforward: we plan to aggregate expecting mothers and compare the tax-class-switching behavior of those earning below vs. those earning above the upper *gross* income threshold pre- and post-reform. For this analysis, we cannot use the upper *net* income threshold to define the treatment and control group as the net income is affected by the choice of the tax class. Therefore, we use the 70,000 € gross income threshold which relates to the 32,400 € net income cap for individuals in tax class V. We cannot employ the lower gross income thresholds of tax class III (50,000 €) or IV (60,000 €) to define the treatment and control group since expecting mothers with the initial tax class V would then still have an incentive to switch to a lower tax class. If the reform affects behavior, we should observe diverging trends in switching behavior between those two groups after the reform. We will exclude women with income below 3,600 €, as they are unlikely to pay any withholding income tax and hence are not assigned to any tax class.

3.7 Conclusion

In this paper, we study the impact of paid parental leave benefits on pre-birth earnings. We exploit the introduction of a new parental leave benefit in 2007 in Germany as a quasi-experimental setting. The new parental benefit implies strong incentives to increase pre-birth net income as the amount of the benefit is explicitly linked to the net income in the 12 months before birth. Expecting parents can increase their pre-birth net income via two channels. Firstly, they can adapt their labor supply in the 12 months before birth to increase their net income. Secondly, even holding gross incomes constant, the German tax law is such that a change in spouses' tax schedule can increase monthly net income of one spouse without affecting the final income tax burden and thereby increase the parental benefit.

We use the universe of German taxpayers filing tax returns and employ a Difference-in-Income Trends estimator to identify the causal effect of the benefit on the first channel. For now, we explore the second channel descriptively. We show three main results: we first document novel stylized facts on the labor market behavior of women in the pre-birth year. Expecting mothers reduce their earnings in the pre-birth year, especially at the bottom of the income distribution. Second, the introduction of the parental leave benefit leads women at the bottom to decrease their pre-birth earnings less and this effect is largely driven by fewer women exiting the labor force in the year preceding childbirth. Lastly, when looking at the alternative channel to maximize the benefit (i.e. tax planning) we observe that a minority of expecting parents exploit this option. The benefit's complex incentive structure is likely to play a role here.

The findings of this paper hold crucial implications for shaping effective parental leave policies, particularly in the context of policymakers aiming to boost female labor supply. Recent studies emphasize the impact of parental leave policies on female labor market decisions. Given the considerable financial investment in these policies, understanding how parents respond is pivotal for optimizing the efficient use of public funds. A critical avenue for future research lies in exploring whether the increased labor supply observed in women during the pre-birth year extends beyond the parental leave period. Investigating the persistence of these effects can offer valuable insights into the long-term consequences of parental leave benefits on female labor supply.

Appendix to Chapter 3

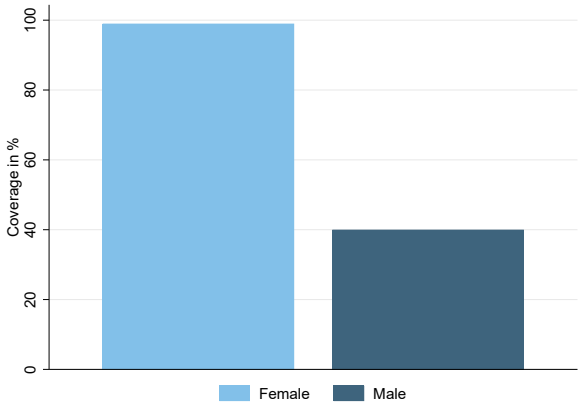
3.A Additional Institutional Details

Country	Paid PLB in place	Payment Base	Relevant Time
Australia	Yes	Adjusted taxable income	10-13 months
Austria	Yes	Average net income	3 months
Belgium	Yes	N/A	N/A
Bulgaria	Yes	Insurable earnings	24 months
Canada	Yes	Average insured earnings	52 weeks
Chile	Yes	Average net earnings	3 months
Colombia	Yes	Earnings	N/A
Costa Rica	Yes	Earnings	N/A
Croatia	Yes	Average insured earnings	6 months
Cyprus	Yes	Earnings	12 months
Czech Republic	Yes	Gross monthly earnings	N/A
Denmark	Yes	Earnings	3 months
Estonia	Yes	Earnings	12 months
Finland	Yes	Earnings	12 months
France	Yes	Net earnings	3 months
Germany	Yes	Average net earnings	12 months
Greece	Yes	Earnings	3 months
Hungary	Yes	Earnings	180 days
Iceland	Yes	Earnings	12 months
Ireland	Yes	N/A	N/A
Israel	Yes	Earnings	N/A
Italy	Yes	Earnings	N/A
Japan	Yes	Earnings	12 months
Korea	Yes	Earnings	N/A
Latvia	Yes	Earnings	N/A
Lithuania	Yes	Net earnings	12 months
Luxembourg	Yes	Earnings	12 months
Mexico	Yes	Earnings	12 months
Netherlands	Yes	Earnings	N/A
New Zealand	Yes	Earnings	N/A
Norway	Yes	Earnings	3 months
Poland	Yes	Earnings	12 months
Portugal	Yes	Earnings	6 months
Slovakia	Yes	Flat payment	N/A
Slovenia	Yes	Basic average earnings	12 months
Spain	Yes	Earnings	1 month
Sweden	Yes	Earnings	N/A
Switzerland	Yes	Earnings	N/A
Turkey	Yes	Earnings	N/A
United Kingdom	Yes	Earnings	8 weeks
United States	No	-	-

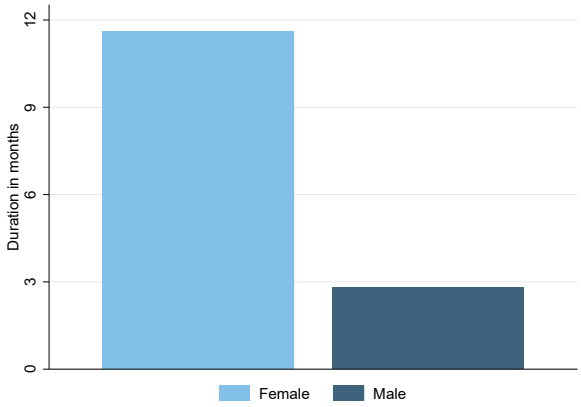
Table 3.A.1: Overview of OECD Countries' National Parental Leave Benefit Systems

Notes: Parental leave benefit is abbreviated with PLB. Relevant time refers to the time period for which the reference income is observed on which the benefits are based on.
Source: Based on OECD Family Database, Table PF2.1.E, December 2022.

(a) Take-Up



(b) Average Duration



(c) Average Amount

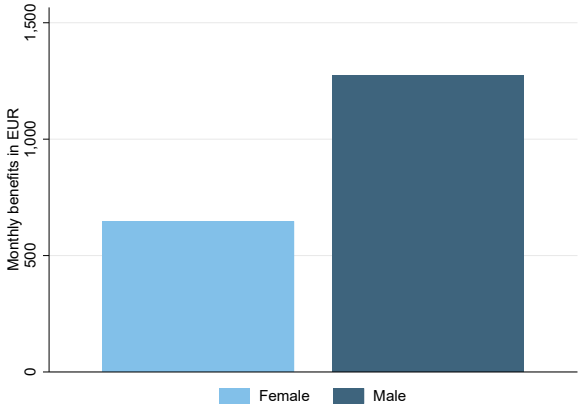


Figure 3.A.1: Descriptives on the Parental Leave Benefit

Source: Destatis, 2014.

	Positive labor income
-	withholding income tax
-	solidarity surcharge
-	church tax
-	social security contributions
=	net income

Table 3.A.2: Calculation of the Parental Leave Benefit's Net Income Base

Source: Parental leave benefit law (BEEG), 2007.

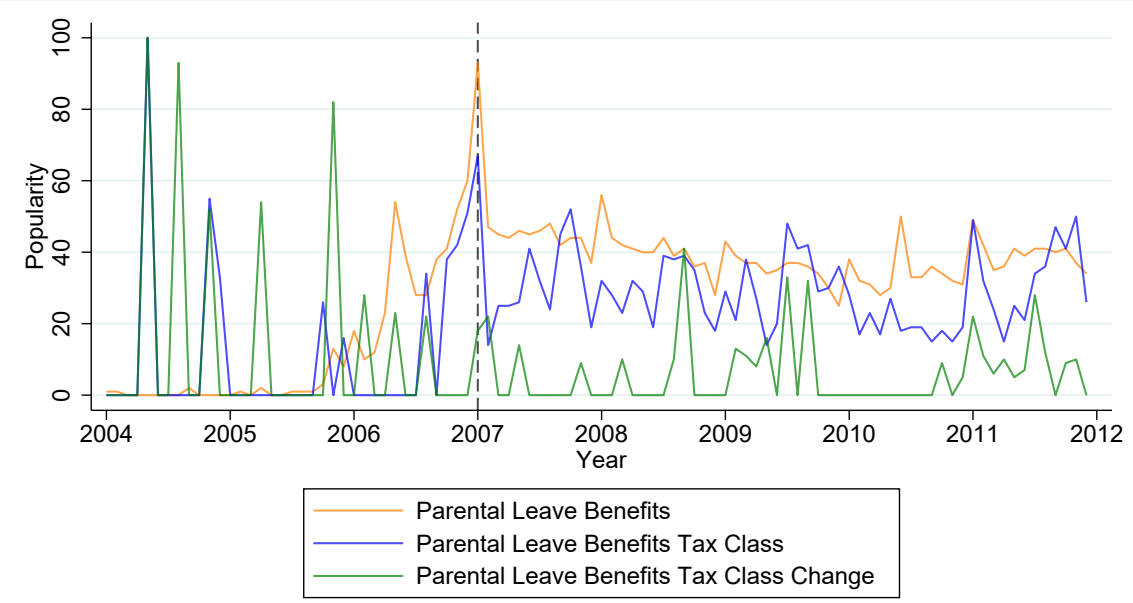


Figure 3.A.2: Google Trends for Keywords Related to the Parental Leave Benefit and Tax Class Changes

Notes: The Figure shows the interest for the three search requests as measured in Google searches. The search interest is measured relatively to the highest search interest from 2004 to 2024. For example, 100 measures the highest popularity of searches for a term, 50 means it is half as popular than at its peak. The searches refer to the German terms "Elterngeld", "Elterngeld Steuerklasse" and "Elterngeld Steuerklasse wechseln".
Source: Google Trends.

3.B Additional Descriptive Statistics

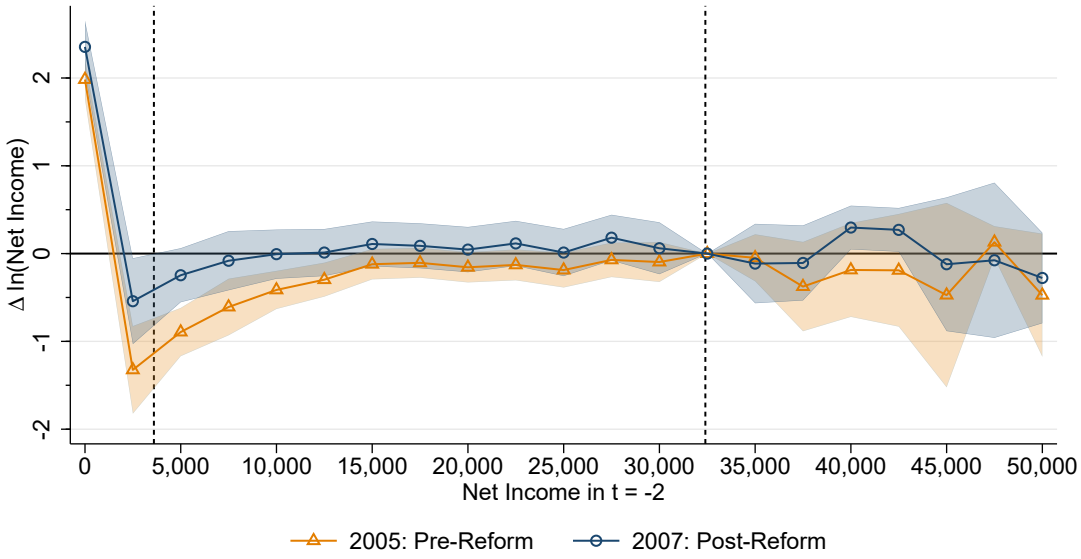
	Pre-reform (2005)		Post-reform (2007)	
	Treatment	Control	Treatment	Control
Female age	29.52	34.07	29.56	33.83
	(9.55)	(3.34)	(4.27)	(3.93)
Male age	32.28	36.15	32.39	36.04
	(8.90)	(4.71)	(5.04)	(4.70)
Catholic	0.47	0.44	0.47	0.43
	(0.50)	(0.50)	(0.50)	(0.50)
West	0.91	0.93	0.91	0.93
	(0.27)	(0.25)	(0.29)	(0.26)
Gross female labor income	27,465	75,270	28,159	74,380
	(11,221)	(26,217)	(11,214)	(25,784)
Net female labor income	16,210	42,156	16,949	41,947
	(5,906)	(14,646)	(6,020)	(14,256)
Gross male labor income	33,680	52,277	35,521	55,312
	(25,270)	(56,934)	(27,353)	(59,108)
Net male labor income	20,378	28,538	20,904	29,543
	(14,221)	(39,785)	(15,317)	(32,132)
Female income share	0.45	0.65	0.44	0.64
	(0.22)	(0.20)	(0.21)	(0.21)
N	76,255	2,143	74,859	3,095

Table 3.B.1: Summary Statistics

Notes: Summary statistics for expecting parents in the pre-birth year pre-reform (2005) and post-reform (2007) for our sample of married women in their pre-birth year ($t = -1$). The treatment group translates to the identification region of women earning between 3,600 €-32,400 €. The control group translates to the validation region of women earning above 32,400 €. 2006 is excluded as expecting parents could not respond yet due to the late passing of the reform. Raw mean, standard errors in parentheses.

3.C Additional Results

(a) Trend Differentials across the Income Distribution



(b) Difference in Income Trends across the Income Distribution

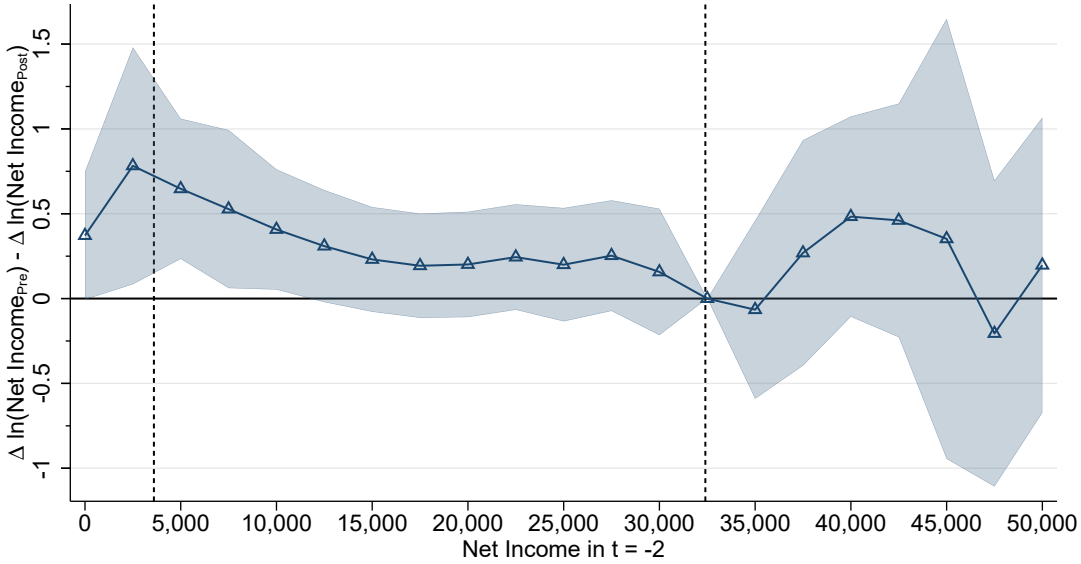


Figure 3.C.1: Income Changes and their Difference before and after the Reform for Women whose Husband is Self-Employed

Notes: Panel (a) plots estimates of equations 3.2 and panel (b) plots estimates of equation 3.3 for women whose husband is self-employed. Estimates are shown with 95% confidence intervals. Our sample consists of married women in their pre-birth year ($t = -1$). Income is adjusted by the consumer price index.

(a) Trend Differentials across the Income Distribution

(b) Difference in Income Trends across the Income Distribution

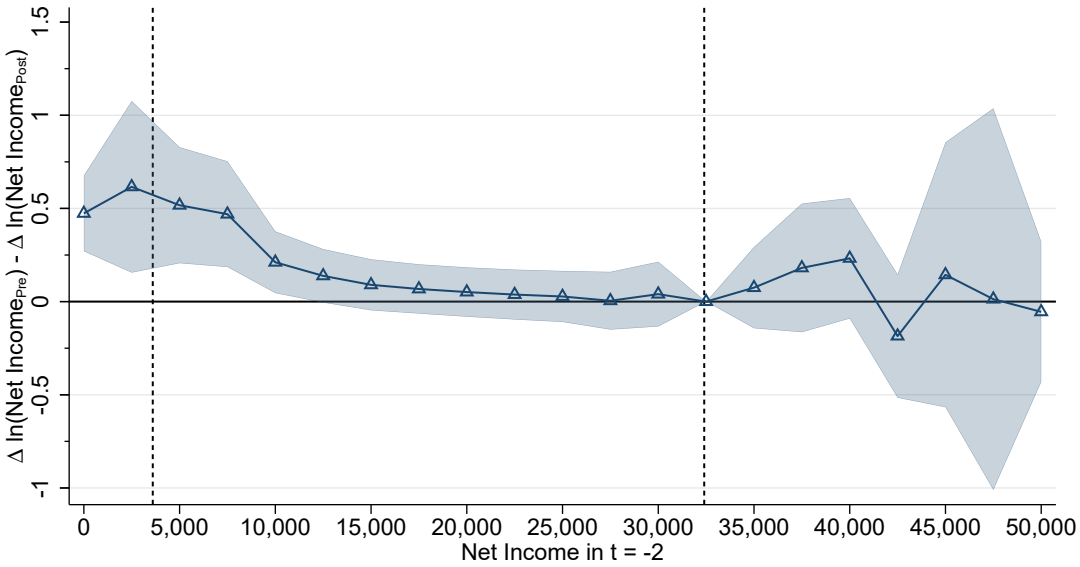


Figure 3.C.2: Income Changes and their Difference before and after the Reform for Women whose Husband's Income is above Median Income

Notes: Panel (a) plots estimates of equations 3.2 and panel (b) plots estimates of equation 3.3 for women whose husband's income is above median income. Estimates are shown with 95% confidence intervals. Our sample consists of married women in their pre-birth year ($t = -1$). Income is adjusted by the consumer price index.

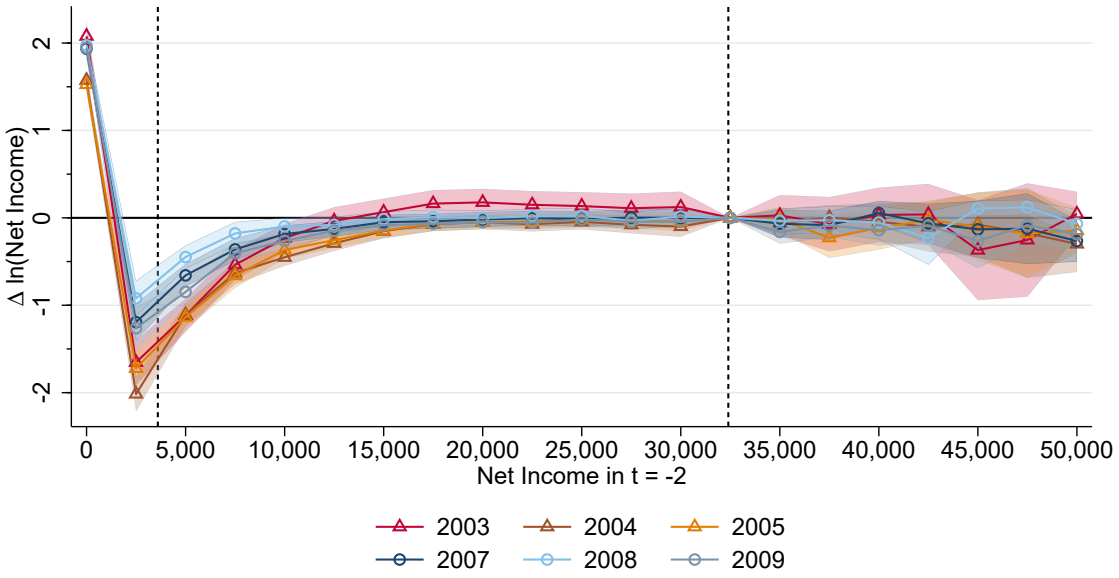


Figure 3.C.3: Trend Differentials across the Income Distribution - Estimated Separately by Years

Notes: Figure shows estimates of equation 3.2 for different pre- and post-refom years. Estimates are shown with 95% confidence intervals. Our sample consists of married women in their pre-birth year ($t = -1$). Income is adjusted by the consumer price index.

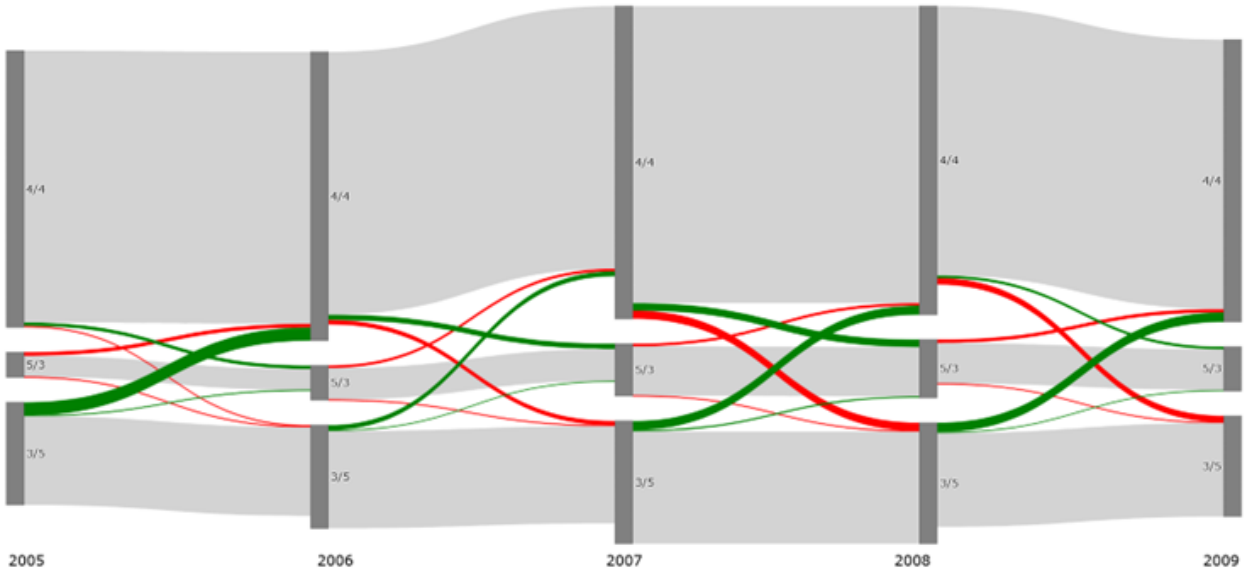


Figure 3.C.4: Tax class Changes in Pre-Birth Year by Initial Tax Class

Notes: The figure shows tax class changes in the pre-birth year for our sample of expecting parents having their first child. The green lines illustrate beneficial, the red line not beneficial tax class changes. The height of the bars and the thickness of the lines translate to the size of the shares.

Chapter 4

How Important Are Local Knowledge Spillovers of Public R&D and What Drives Them?*

4.1 Introduction

There is a clear consensus that stimulating firm R&D is welfare increasing due to positive externalities and uncertainty (Arrow 1962; Nelson 1959), but the question about the most efficient way is still open to debate. In principle, there are two ways to do so. First, governments can stimulate private R&D by increasing the profitability of R&D investment, either by handing out cash payments or by providing a beneficial tax treatment. Second, governments can fund public R&D, which may stimulate firm R&D via local knowledge spillovers which are driven by personal interactions and workforce mobility. While a large body of literature has examined the impact of direct subsidies[1], empirical evidence on the magnitude of local knowledge spillovers of public on firm R&D and its determinants is limited and largely based on universities. Universities engage, however, in both, knowledge as well as "degree" production. Identifying the transmission channel at work is thus challenging. Since communication and travel costs have substantially fallen over the last decades, technological spillovers are less and less bounded by national borders (Griffith, Lee, and Van Reenen 2011) and local knowledge spillovers are a promising channel to solely support domestic firm R&D. To inform policy makers on whether public R&D should be part of the optimal firm R&D support strategy or not, our paper aims to quantify the magnitude of local knowledge spillovers of public R&D as well as its determinants by using patent, firm, and county level data for Germany between 1995 and 2015 and exploiting regional variation in public R&D undertaken by (publicly funded) research institutes and universities.

We believe that Germany provides an excellent set-up for our research question. First, Germany relies substantially on public R&D carried out by independent research institutes. Figure 4.1 illustrates the international comparison. While Germany, Japan, the UK, the US, and the average EU 28 country have similar R&D expenditures relative to GDP in the higher education sector, R&D spending in the government sector (including the research institutes) is highest in Germany among the selected countries. Second, the German firm R&D support strategy mainly consisted of funding public R&D dur-

*This chapter is based on a co-authored article in *Research Policy* (Koch and Simmler 2020).

[1]See David, Hall, and Toole (2000) and Zúñiga-Vicente et al. (2014) for a literature review on the impact of cash subsidies on firm R&D and Hall and Van Reenen (2000) and Guceri and Liu (2019) for a literature review on the impact of R&D tax credits on firm R&D. In general, prior literature has found that R&D tax credits are quite successful in stimulating firm R&D (e.g. Dechezlepretre et al. 2016; Guceri 2018; Lokshin and Mohnen 2012; Mulkay and Mairesse 2013; Rao 2016), while cash subsidies increase firm R&D in particular for smaller and younger firms (e.g. Bronzini and Iachini 2014; Howell 2017).

ing our sample period.[2] Germany did not offer super-deductions for R&D expenditures or a lower tax rate on returns from R&D investment in the form of a patent box as introduced in several European countries over the last two decades (see, for example, Alstadsaeter et al. 2018).[3] Thus, our focus on Germany mitigates concerns that the estimated spillovers of public R&D are driven by both, public R&D as well as generous direct subsidies. Third, German firms have comparably high R&D expenditures (see Figure 4.1). Thus, a no impact result cannot be explained by a generally low R&D intensity of firms.

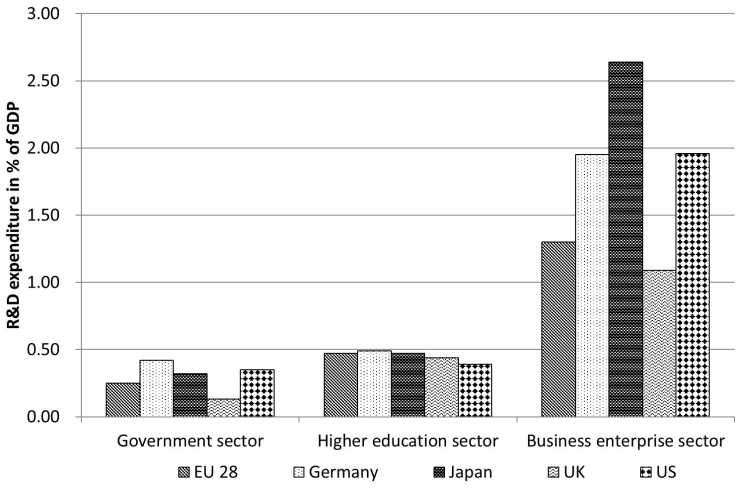


Figure 4.1: Public and Private R&D Expenditure as a Percentage of GDP

Source: Authors' calculations based on data from Eurostat, 2014.

Our paper consists of two empirical parts. In the first part, we aim to quantify the relative importance of three potential local public knowledge spillover channels. The first channel is firms' use of specific knowledge that is produced by a local public institution (i.e. public patent). It is based on the idea that public institutions advance (fundamental) scientific knowledge and that this has positive (technological) spillovers to applied researchers (e.g. Foray and Lissoni 2010). Prior work has shown that geographic proximity enhances these technological spillovers (e.g. Belenzon and Schankerman 2013), which is consistent with the argument that (public) knowledge has a tacit dimension and thus cannot be fully transmitted via written publications. The second channel is the collaboration (or joint ventures) of firms with (local) public institutions. This channel acknowledges that firms may not only use public knowledge but also work together with public institutions towards creating new knowledge. Firms benefit from collaboration as it allows them to access knowledge as well as technology of public institutions and to exploit synergy effects as, for example, documented for research consortiums in Japan (Branstetter and Sakakibara 2002). Prior work has shown that geographic proximity

[2]Germany introduced an R&D tax credit in 2019.

[3]Patent box regimes grant preferential tax treatment for corporate income derived from intellectual property (IP), especially patented inventions. Firms can declare the profits derived from IP and these profits are taxed at a lower rate relative to income derived from standard business activities. The design of patent box regimes differs with respect to e.g. the tax rate, the type of eligible IP, income or expenses across countries.

amplifies the likelihood of collaborations, potentially, as face-to-face interactions are still the preferred mode of communication (e.g. Rybnicek and Königsgruber 2019). The third channel is the use of non-specific public knowledge by local firms. This channel is motivated by the fact that innovation is a complex process, which consists of "weaving different types of knowledge into something new, different and unprecedented that has economic value" (Feldman and Kogler 2010, p. 384). In particular, new technologies do not emerge fully developed overnight but are rather fine-tuned over time. Existing local knowledge is likely to be key for the gradual adjustments as it is easily available to the entrepreneurs. While this knowledge includes the outcome of prior public R&D, it also includes the knowledge about the underlying process of these innovations, e.g. "the knowledge of what does not work, what approaches have been tried, and led to dead ends" (Feldman and Kogler 2010, p. 386). Since in particular the latter knowledge is highly tacit and entrepreneurs rely on existing local knowledge for the full development, formal or informal personal interactions (e.g. consulting projects) as well as workforce mobility are likely to be imperative for the transmission of it.[4]

To analyze the importance of the different spillover channels, we exploit regional variation in public R&D in Germany between 1995 and 2015. We proxy firm R&D with the number of patent applications and technological spillovers by using patent citation data. Collaboration spillovers are proxied by joint patent applications of firms and public institutions. Lastly, we proxy non-specific knowledge spillovers with the number of public patent applications within a region.

To assess the relevance of the three channels, we estimate count models of patent applications on the applicant-region level. To account for public R&D quality in our analysis, we use citation data to understand which public R&D, e.g. what type of public patents, is of higher relevance for local firm R&D, as measured by the number of citations by corporate patents. Moreover, we investigate whether public patents that are more relevant for local firm R&D cite more local firm patents to assess the role of reverse causality. Based on the results of the transmission channel analysis and using the characteristics of the average public patent (and taking into account that we capture only part of all collaborations) allows us then to quantify the relative importance of the different local knowledge spillovers channels. To deepen our understanding further, we estimate not only the direct effect on affected firms but also spillovers into other regions. Furthermore, we assess the evidence for impact heterogeneity related to the type of affected firm (low vs. high R&D intensity firms, measured by the number of patent applications filed before our sample period) as well as the type of corporate patent (low vs. high technological value, measured by the number of patent classes).

In the second part of the empirical analysis, we aim to quantify the overall magnitude of local knowledge spillovers of public R&D on local firm R&D within a county. We proxy public and firm R&D with the number of patent applications and use an instrumental variable (IV) strategy to account for a potential bias in the Ordinary Least Square (OLS) estimator due to measurement error, omitted variables and reverse causality. The excluded instrument we employ is the 4-year lagged institutional funding for research

[4]The existence of these non-specific knowledge spillovers may also explain why substantial local public knowledge spillovers and public knowledge tacitness do not contradict each other (see Foray and Lissoni 2010, p. 293): Non-specific public knowledge is not tacit because of the novelty or the strategy of the producers but as it is (usually) not included in written publications.

institutes.

Three aspects are worth highlighting with regard to our analysis. First, we do not aim to quantify the overall magnitude of knowledge spillovers of public R&D but only the magnitude of *local* knowledge spillovers and the underlying channels at work. We believe these two dimensions to be crucial to understand to what extent public R&D should be used to stimulate firm R&D as (national) governments care only about firms that undertake research in their region (country). Since communication and travel costs have substantially fallen over the last decades, technological spillovers of public R&D are less and less bounded by national borders (Griffith, Lee, and Van Reenen 2011), which means that local knowledge spillovers are potentially a key transmission channel for stimulating predominantly domestic firm R&D. Given our focus, we implement a relatively simple approach to capture local spillovers. In particular, we do not use a distance based approach as in Alcacer and Gittelman (2006) or Lychagin et al. (2016) but focus on spillovers within a county, which has in Germany an average size of around 1,000 square kilometers. This level of analysis is likely to capture a large part of local knowledge spillovers given the results of prior literature (e.g. Andersson, Quigley, and Wilhelmsson 2009; Belenzon and Schankerman 2013) and provides us with enough variation to assess the overall magnitude of local knowledge spillovers. Second, we focus on local *knowledge* spillovers. To remove time-invariant heterogeneity which could bias our estimates, we include region-fixed effects in all specifications. This means, however, that time-invariant effects of public institutions, e.g. related to the (unchanged) provision of R&D research infrastructure (e.g. facilities, resources, and services) or the provision of consultancy (see Ankrah and AL-Tabbaa 2015) are absorbed as well if they are unrelated to the knowledge production dimension of public R&D.[5] Third, local knowledge spillovers do include product market rivalry spillovers, e.g. negative spillovers that result from the competition of firms/producers in product markets (see Bloom, Schankerman, and Van Reenen 2013; Branstetter and Sakakibara 2002).[6] Their importance is, however, likely to be limited as public R&D producers compete with private firms in only very few markets. The market for contract research and for technical testing and analysis (e.g. metrology services) are probably most important. Moreover, the market share of public R&D producers in these markets is suggested to be small. The research institutes that belong to one of the four main research organizations considered in our analysis had own business income (i.e. turnover) of around 1.2 billion EURO (see Table 4.1) in 2010, while overall turnover in the sectors scientific research and development (NACE Rev. 2 Code 72) and technical testing and analysis (NACE Rev. 2 Code 71.2) amounted to 10 and 9 billion EURO in Germany in 2010, respectively.[7]

The data of our empirical analysis is - as already indicated - patent application and citation data. We choose this data as it allows an (also imperfect) approximation of the three channels outlined above. Moreover, it has the benefit that we are able to

[5]For example, if local public R&D affects firm R&D via the pure provision of research facilities and the provision is unrelated to public knowledge production, this would not be captured by our estimation strategy. If, however, public knowledge production increases the value of using public research facilities by local firms, we would capture the second effect partly as well.

[6]In addition, public R&D may by supporting or setting standards in a particular industry that influence product market competition between firms.

[7]Federal Statistical Office, Table 47415-009 (http://www.datenportal.bmbf.de/portal/de/K16.html, last accessed 16/03/2020).

study the patenting behavior of all firms in Germany which would not be the case when using financial statements data (and R&D expenses). The downside is that we only observe patent applications. As pointed out by Griliches (1998), this means that we potentially underestimate the extent of local knowledge spillovers since not all inventions are patentable and not all new innovations are patented. We address this concern by using the number of firm R&D employees as an alternative measure for firm R&D. Since this information is only observed on the county level, the heterogeneity results (which firms and patents benefit most from public R&D) should be interpreted with some caution as firms' propensity to "patent" varies between industries as well as with firm size (Bound et al. 1984; Scherer 1983). Moreover, we apply an IV strategy to address the measurement error in the number of public patents.

The patent data we use stems from the OECD RegPat and Citation database and covers all patents that are filed with the European Patent Office (EPO) and under the Patent Co-operation Treaty (PCT). It includes information on general patent characteristics, address information for inventors and applicants as well as citations of patent and non-patent literature. The home address of inventors is key for our analysis as we use it for the geographical mapping of patents. It allows a better approximation of where R&D takes place than using the location of the applicant (see, for example, Lychagin et al. 2016). The latter would in particular be misleading for large firms that consist of several establishments or for legally non-independent research institutes as for example the Fraunhofer institutes, for which the applicant is always the Fraunhofer Society located in Munich.

We obtain three main results from our empirical analysis. First, we find evidence that is consistent with the existence of all three types of spillovers. Based on the assumption that we are able to identify causal effects and that the proxies employed capture reasonably well the three different channels, our results suggest that the non-specific knowledge spillovers are most important as they account for around 2/3 of the overall local knowledge spillovers of public R&D. Moreover, we find that the non-specific knowledge spillovers are in particular relevant for low R&D intensity firms, which is consistent with the argument that non-specific knowledge is highly tacit and that tacit knowledge is in particular relevant for firms in their early stages (see Feldman and Kogler 2010). However, there is no difference between the channels with regard to what sort of public R&D matters the most (high quality public R&D) and which firm patents are affected the most (low quality firm R&D).

Second, assuming that our IV strategy allows a causal interpretation of the results and that the number of public patents is a reasonable proxy for public R&D production, we find evidence for substantial local knowledge spillovers of public R&D. For the median county, our estimates imply that one additional public patent generates around 3 additional firm patents. Consistent with the results of the transmission channels, we find that these additional corporate patents come from low R&D intensity firms. Since we do not find evidence that firms that benefit from local R&D in one county file less patent applications in other counties, our result suggests that local public R&D is an efficient way to stimulate (local) firm R&D. Using the average costs per EPO patent application (for UK firms) as reported by Dechezlepretre et al. (2016), our estimates imply public costs per additional firm patent of around 1.4 million EURO. This is less than 2/3 of the implied public costs for one firm patent application in the US or the UK

using R&D tax credits as firm R&D support strategy (see Dechezlepretre et al. 2016; Rao 2016).

Third, based on our estimation strategy and the proxies employed, we find that the local knowledge spillovers of public R&D on local firm R&D are increasing at a decreasing rate in the level of public R&D. While we are able to show that this effect is not driven by a potential correlation of region size with the level of public patents or a smaller number of low R&D intensity firms in regions with a large number of public patents, we are not able to explain *empirically* why the relationship is nonlinear. Theoretically, there are at least three potential explanations. First, the relevant non-specific knowledge created by additional public R&D could be decreasing in the level of public R&D. Second, public knowledge production could have diminishing returns as there may be only a limited number of firms within a region that can benefit from public R&D. Third, public and firm R&D producers may compete in local labor markets. In this case, more public R&D does not only lead to larger local knowledge spillovers but also to higher wages for R&D employees (due to the demand effect) and the latter effect dampens the former.[8]

Our work contributes to the literature on local knowledge spillovers of public R&D in several ways. First, we provide evidence in line with various other studies on positive local knowledge spillovers of public R&D. The seminal paper in this field is Jaffe (1989). Using US state level data, he finds an elasticity of corporate patents to university R&D expenditure of around 0.1. Andersson, Quigley, and Wilhelmsson (2009) find that openings of higher education institutions in Sweden increased the number of patents in the same region. Belenzon and Schankerman (2013) find - using citation data - that spatial proximity matters not only for private knowledge as shown by Audretsch and Feldman (1996) and Jaffe, Trajtenberg, and Henderson (1993) but also for the diffusion of public knowledge. Our analysis comes to the same conclusion but also suggests that citation data is unlikely to capture the full extent of local knowledge spillovers since a substantial part of them (e.g. non-specific knowledge spillovers) cannot be traced with citation data.[9]

Further, while most of the recent studies have used less aggregated data than Jaffe (1989), not all of them have addressed the potential endogeneity of public R&D. The most convincing approach employed in prior literature is by Kantor and Whalley (2014). They exploit that US universities usually invest a fixed amount of their endowments' market value every year and instrument university spending with university endowment in combination with variation in stock market returns. Our instrument follows this rationale. While the 4-year lagged institutional funding may not be fully exogenous, it should be noted that reverse causality is likely to be of minor importance in Germany as institutional funding is determined not only by the hosting state but by all German states as well as the federal state. In line with the results by Kantor and Whalley (2014), we also find a substantial downward bias when ignoring the endogeneity of local

[8]Two additional channels that have been suggested in the literature are that public R&D might directly replace private R&D, if firms substitute public R&D for their own R&D and that public R&D could distort the competition between firms by funding some firms at the cost of others, see David, Hall, and Toole (2000) and Zúñiga-Vicente et al. (2014) for a literature review.

[9]This problem may be less severe when using US patent data (which includes a more complete list of relevant existing patents), but is certainly true for EPO patent data as EPO patent applications must only contain the most relevant existing patents (see, for example, Maurseth and Verspagen 2002; Wartburg, Teichert, and Rost 2005).

public R&D.

Second, we contribute to the literature that investigates heterogeneity of local knowledge spillovers of public R&D. Link and Rees (1990) find that large firms are more likely to participate in university-based research but that small firms are better able to transfer the knowledge. Acs, Audretsch, and Feldman (1994) report that the elasticity of innovative activity with respect to corporate R&D is greater for large firms and that the elasticity with respect to public R&D is greater for small firms. Moreover, the authors show that geographic proximity between universities and corporate R&D has a larger impact on small firms. Our results for low (and likely small) and high R&D intensity (and likely large) firms are in line with these findings.

Which type of public R&D creates the largest spillovers has not yet been analyzed in great detail. Branstetter and Sakakibara (2002), who study the effectiveness of research consortiums in Japan, find that the more basic research (measured using survey data) is done by a consortium, the higher a firm's R&D after its inclusion in the research consortium. Work by Cassiman, Veugelers, and Arts (2012) suggests that basic public R&D is of higher importance for firm R&D in the case of industry-university collaborations. Our results show that patents that have a wider scope (i.e. patents that are filed for more patent classes) are more important for spillovers resulting from collaboration as well as technological and non-specific knowledge spillovers. Moreover, our work suggests that patents with a wider scope receive more forward citations and thus are likely to have a higher technological value.

Our two sets of heterogeneity results have implications for the transmission channel at work. A common argument for the existence of local knowledge spillovers is that geographic closeness facilitates the interaction between firm and public institution R&D employees, which allows them to learn from each others' research activities. Moreover, public knowledge may also be transferred to firms by employees that change jobs but are still employed in the same region (Almeida and Kogut 1999) or by university spinoffs, which locate close to their source university to capture the competitive advantage of the provision of skilled labor, specialized facilities, and expertise by the university (Bercovitz and Feldman 2006). While our results highlight the relevance of non-specific knowledge spillovers and thus are in line with the communication channel, they raise the question why high R&D intensity firms do not benefit from public R&D. One potential explanation for this finding is that there may be two separate R&D employee markets, one for low R&D intensity firms and public institutions and one for high R&D intensity firms.

Our third contribution to the literature is that we provide robust evidence that it is indeed public knowledge production that affects local firm R&D. One challenge regarding prior work is the identification of the driving force at work, as most of the studies focus on public R&D carried out by universities. Universities engage, however, in both "degree" as well as R&D production and both are likely to influence R&D activities by firms. Disentangling these two effects is challenging, as it necessarily relies on a parameterization. Abel and Deitz (2011), who proxy "degree production" with the number of degrees per 100 working-age people and knowledge production by R&D expenditures per enrolled student, find that it is the latter that increases local human capital since the impact of degree production is limited due to labor force mobility. Some of the results by Kantor and Whalley (2014) can be read in a similar vein. They find a larger effect of

university R&D on wages in counties with universities that have a higher research intensity and in counties with industries that cite university patents more often. While this seems conclusive, it should be noted that degrees also have a quality dimension, which is likely to be influenced by a university's knowledge production. We contribute to this stream of literature by applying an IV strategy that exploits variation in institutional funding of research institutes, which is substantially less related to "degree" production within a region compared to (institutional) funding for universities.[10] Moreover, our result suggests that technological and collaboration spillovers only account for a small share of the overall local knowledge spillovers of public R&D.

The rest of this article is structured as follows. In section 4.2 we explain the German research system. Section 4.3 describes the data used for our empirical analysis. Section 4.4 explains the methodology and reports the results for the transmission channels, and section 4.5 does this for the overall magnitude of local knowledge spillovers of public R&D. Section 4.6 concludes.

4.2 R&D in the Public Sector in Germany

Since launching of the Lisbon strategy in 2000, the EU member countries aim to invest 3% of GDP in R&D. Compared to other EU countries, Germany is close to this target with a sum of public and private R&D expenditures of 2.88% of GDP (based on 2014 data). Surprisingly, Germany is also one of the few EU countries that offered only cash subsidies to the private sector and did not grant tax credits for R&D expenses or a lower tax rate on the return of R&D investment in the form of a patent box over our sample period. Instead the German federal and the state governments heavily invest in public R&D carried out by independent research institutes and - similar to other countries - fund universities, which also engage in R&D activities. The majority of the independent research institutes belong to one of the following umbrella associations: the Max Planck Society, the Fraunhofer Society, the Leibniz Society and the Helmholtz Community.[11] As shown in Figure 4.1, R&D spending in the government sector (including the independent research institutes) is almost as high (12.5 billion EURO in 2014) as in the higher education sector (15.3 billion EURO).

The structure of the higher education sector in Germany is as follows. Overall, there are more than 400 higher education institutes, but only about 25% of them are "real" universities (see also Table 4.1). The rest are universities of applied sciences (*Fachhochschulen*). They engage in research as well but do so to a much smaller extent and are usually highly specialized in one field. While the number of universities of applied sciences has increased over time, the number of universities has been stable. The state, in which the higher education institute is located, is responsible for 75% of the basic funding, which is usually a function of the number of students, the number of graduates and the amount of third party funding obtained. These criteria vary from state to state. The federal government and the private sector contribute in the form of public

[10] Research institutes in Germany are not directly involved in the provision of higher education (except in the case of doctoral students). R&D production of research institutes may, however, impact (nearby) universities' degree production indirectly via, for example, collaborations of research institutes and universities as well as personal ties between the two.

[11] Governmental research institutes exist in Germany as well, but they are of minor importance and conduct mainly departmental research for federal and state administration.

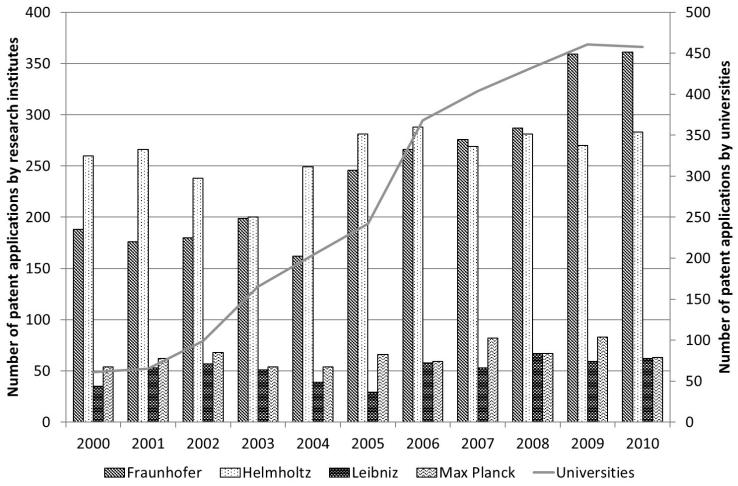


Figure 4.2: Evolution of Public Patent Applications by Type of Applicant

Notes: Patent applications include only applications filed with the European Patent Office or under the Patent Co-operation Treaty.
Source: Authors' calculation based on OECD RegPAT database, 2000-2010.

and private third party funding for specific research projects. Tuition fees are not a significant source of funding. The level of university R&D expenditures has increased in our sample period, from 8.4 billion EURO in 2000 to 12.7 billion EURO in 2010.[12] The main driver was an increase in third-party funding, which grew from 2.8 billion EURO (one fourth) in 2002 to 5.9 billion EURO (one half) in 2010. The number of university patent applications filed with the EPO or under the PCT also increased between 2000 to 2010 (see Figure 4.2). One factor which is likely to have contributed to the increase in university patent applications is a law change in 2002. Prior to 2002, professors and researchers at a university had "Professor's Privilege" which allowed them to file patents for their own inventions as the sole applicant. After the reform, universities were granted the intellectual property rights of their employees and the inventors received only a compensation for their discoveries. Employees at public research institutes never had these exclusive intellectual property rights to their inventions.[13]

The main players in the government sector are the research institutes that belong to one of the four research umbrella organizations named earlier. They account for roughly

[12]German Federal Ministry of Finance, Data portal, Table 1.6.2 (http://www.datenportal.bmbf.de/portal/de/K16.html, last accessed 19/12/2019).

[13]We believe the impact of the professor privilege for university staff until 2002 to be neglectable on our results. First, our sample for the analysis of the overall magnitude of local knowledge spillovers of public R&D only includes the years from 2003 to 2010. Moreover, we apply an IV strategy based on the funding for research institutes which should account for any remaining measurement error. Second, while we use the years from 2000 to 2010 for the transmission channel analysis, the resulting measurement error is likely to be small as university patents account for at most 40% of the considered public patents (based on 2010 data) and there was only a university privilege in 3 out of 11 years during our sample period. Moreover, the characteristics of the average public patent application are largely independent of the sample period considered (e.g. 2005 to 2010 vs 1995 to 2010).

75% of overall R&D expenditures in the government sector in Germany.[14] A large share of their income (60%) comes from the federal and the state governments in the form of institutional funding. There is, however, substantial heterogeneity between organizations (see Table 4.1). Further, the reader should note that the federal structure in Germany is also visible in public R&D funding as the state in which the institute is located does not bear the whole (state) burden, but all other states contribute to the funding as well.[15]

The Fraunhofer Society (17,000 employees) consists of 60 legally non-independent institutes (in 2010) and is the largest organization for application-oriented research in Europe. Two thirds of their R&D expenditures comes from project funding or own income which is relatively high compared to the other research organizations, the remainder is institutional funding. The latter comes to 90% from the federal government, 3.3% from the "home" state and 6.6% from all other states. In 2010, the overall budget of the Fraunhofer Society amounted to 1.62 billion EURO and the Fraunhofer institutes filed 361 patent applications with the EPO (or under PCT) in our data, out of a total of 502 patentable innovations in 2010.[16]

Researchers at the 75 Max Planck institutes (13,000 employees) conduct solely fundamental research. The majority of the organization's expenditures (1.54 billion EURO in 2010) is institutionally funded (around 80%) and stems equally from the federal government and the states. Half of the state share is funded by the "home" state and the rest by all other states. The Max Planck society registered 63 new patents with the EPO (or under the PCT) in our data and had in total 87 patent applications in 2010.[17] The relatively high expenditures per filed patent compared to the Fraunhofer institutes is due to a stronger focus on natural sciences (83% compared to 30%), humanities (11% compared to 2%) and fundamental research (100% compared to 5%).

The Helmholtz Community (32,000 employees) is the biggest research organization in Germany and consists of 17 research centers. They conduct fundamental research (71%) by employing large facilities. Institutional funding makes up around 75% of their total income (3.20 billion EURO in 2010) and public and private third party funding and own income around 25%. The institutional share is borne to 90% by the federal government, 5% by the "home" state and the rest by all other states. The Helmholtz institutes filed around 400 new patent applications in 2010, 283 are observed in our data.[18] While the ratio of expenditures per patent is closer to the one for the Fraunhofer institutes compared to the Max Planck institutes, it is still larger. The reason is the stronger

[14]German Federal Statistical Office, 2012, Finance and Taxation (Fachserie 14), Reihe 3.6, Table 1.3.

[15]The contribution of all non-home states follows the *Königsteiner Schlüssel*, which is based on tax revenue after fiscal equalization (67%) and population size of the respective state (33%).

[16]GWK - Pakt für Forschung und Innovation - Monitoring-Bericht 2018, Table 14 (https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/GWK-Heft-58_Monitoring-Bericht-2018.pdf, last accessed 19/12/2019).

[17]GWK - Pakt für Forschung und Innovation - Monitoring-Bericht 2018, Table 14 (https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/GWK-Heft-58_Monitoring-Bericht-2018.pdf, last accessed 19/12/2019).

[18]No exact number is available for 2010. In 2012, 409 patents have been filed. (GWK - Pakt für Forschung und Innovation - Monitoring-Bericht 2018, Table 14 (https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/GWK-Heft-58_Monitoring-Bericht-2018.pdf, last accessed 19/12/2019)).

	Higher Education	Fraunhofer	Max Planck	Helmholtz	Leibniz	Other
# Institutions	415 (106 Uni.)	60	75	17	86	./.
Share fundamental research	./.	5	100	71	75	30
Share natural sciences	30	30	83	52	57	27
Share engineering	20	64	0	34	10	22
Share humanities	22	2	11	1	20	31
# Overall patents	./.	502	87	approx. 400	approx. 110	./.
# EPO and PCT patents	458	361	63	283	62	./.
R&D budget (€bn)	12.7	1.62	1.54	3.20	1.41	5.2
Institutional funding (€bn)	6.8	0.55	1.23	2.04	0.91	./.
Own business income (€bn)	./.	0.5	0.07	0.6	0.01	./.
Institutional funding shares						
Federal gov.	25	90	50	90	50	./.
Home state gov.	75	6.7	25	5	25-50	./.
Other state gov.	0	3.3	25	5	50-75	./.

Table 4.1: Funding and Spending of Public R&D in 2010

Notes: Federal and state funding shares are on average. Contribution of other states depends on *Königsteiner Schlüssel*. Patent applications include only applications filed with the European Patent Office or under the Patent Co-operation Treaty. Humanities include social sciences.
Source: Authors' calculations based on German Federal Ministry of Finance, data portal, Table 1.6.1 and 1.2.2, German Statistical Office, 2012, Finance and Taxes (Fachserie 14), Reihe 3.6, Table 2.4, 3.8, 4; GWK - Pakt für Forschung und Innovation - Monitoring-Bericht 2018, Table 14, and OECD RegPat database, 2010.

focus on natural sciences (52%) and fundamental research (71%).

The Leibniz Society (13,500 employees) has 86 independent member institutions that widely vary from academic service facilities to fundamental research institutes. Around 65% of their budget (1.41 billion EURO in 2010) is institutionally funded. Half of the institutional funding comes from the federal government, and on average 25% from the "home" and 25% from all other states, but this varies between institutes. The member institutions of the Leibniz society filed 66 patent applications in our data, out of around 110 in total in 2010.[19] The relatively higher expenditures per filed patent result from a larger budget share for natural science (55%), humanities (20%) and for fundamental research (75%).

The funding for the research institutes is set for some years in advance in non-public committees, in which representatives of the umbrella organizations as well as of the federal and state governments take part. Institutional as well as overall funding increased over time, the latter from 5.4 billion EURO in 2002 to 7.8 billion EURO in 2010. Consistent with the budget increase, an upward trend in the number of patent applications by these research institutes in our data is observed between 2000 and 2010 (see Figure 4.2). Figure 4.3 summarizes the evolution of the overall number of patent applications by universities and our considered research institutes. Further, the figure shows the evolution of private patents, which also increased between 2000 and 2010 (although with a substantial drop during the financial crisis).

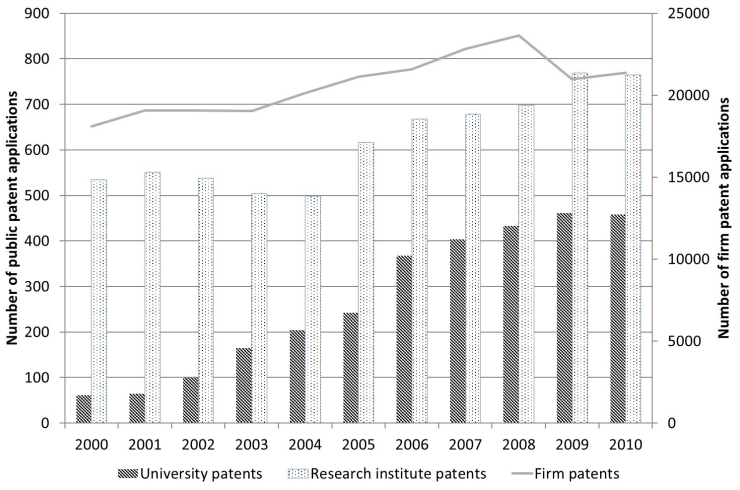


Figure 4.3: Evolution of University, Research Institute and Firm Patent Applications

Notes: Patent applications include only applications filed with the European Patent Office or under the Patent Co-operation Treaty.
Source: Authors' calculation based on OECD RegPAT database, 2000-2010.

[19]No exact number is available for 2010. In 2012, 121 patents have been filed (GWK - Pakt für Forschung und Innovation - Monitoring-Bericht 2018, Table 14 (https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/GWK-Heft-58_Monitoring-Bericht-2018.pdf, last accessed 19/12/2019)).

4.3 Patent Data

In the following, we describe the data used for our empirical analysis. All patent data stems from the OECD RegPat (Update March 2018) and Citation database (Update September 2017).[20] This database covers all patent applications filed with the European Patent Office (EPO) and under the Patent Co-operation Treaty (PCT) up to 2015 and includes general information about the patent as well as name, address and county codes of applicants and inventors.[21] Moreover, information on patent and non-patent literature citations are available.[22]

We identify public institutions as the applicant by using the applicant names in the data (as well as the harmonized applicant names provided by the so-called HAN database).[23] Corporate applicants are identified via the legal form using the applicant names. Moreover, we also assume a corporate applicant if the harmonized name is available (and if we do not identify a public institution as the applicant). The reason is that the harmonized names are based on financial statements data and thus should only include applicants that are corporations. Lastly, we also assume a corporate applicant if the applicant has filed more than 30 corporate patents.[24] To map patents into regions, we use the inventors' address information as this has been shown to capture local knowledge spillovers much better than using the headquarter location (see, for example, Lychagin et al. 2016).[25] For our study, this is in particular relevant for the Fraunhofer Society as well as the Max Planck society, which consists of legally non-independent institutes and thus the applicant location would always be the location of the Fraunhofer or the Max Planck society. The regional weight of a particular patent is based on the share of inventors that live in a particular region (in our case in a county). If a patent is filed by more than one applicant, we further weight the patent count by the number of applicants. Thus, a patent that is filed by two applicants, a public institution and a firm, and is invented by two inventors, one located in A and one located in B, has a regional public patent weight of 0.25 in A and of 0.25 in B and a regional corporate patent weight of 0.25 in A and of 0.25 in B. Figure 4.4a and Figure 4.4b illustrate the resulting distribution of public as well as firm patents in Germany in 2003. Counties with a darker color have more firm or public patent applications.

One potential disadvantage of our data is that it does not include all patents filed by firms, universities and research institutes. The resulting bias is, however, likely to be small. First, according to the German Patent Office, roughly 660.000 patents

[20] http://www.oecd.org/sti/intellectual-property-statistics-and-analysis.htm

[21] A patent application filed under the PCT is similar to a patent application filed with the EPO. It allows to file one patent application to protect an invention in several (potentially all PCT contracting) states.

[22] Since EPO or PCT patents may be cited via their national counterparts, we used the equivalent data sets provided by the OCED to account for that in the citation analysis. Moreover, the reader should note that the citation data includes both, applicant as well as examiner citations, but each cited patent is only counted once.

[23] http://www.oecd.org/sti/intellectual-property-statistics-and-analysis.htm

[24] The impact of the second and third selection criteria are small as we obtain similar results using only the legal form of the applicant as selection criteria.

[25] The use of the inventor address information is the main reason why we cannot use the patent data by the EPO (PATSTAT). While for some countries PATSTAT includes address information for inventors, it does not do so for Germany.

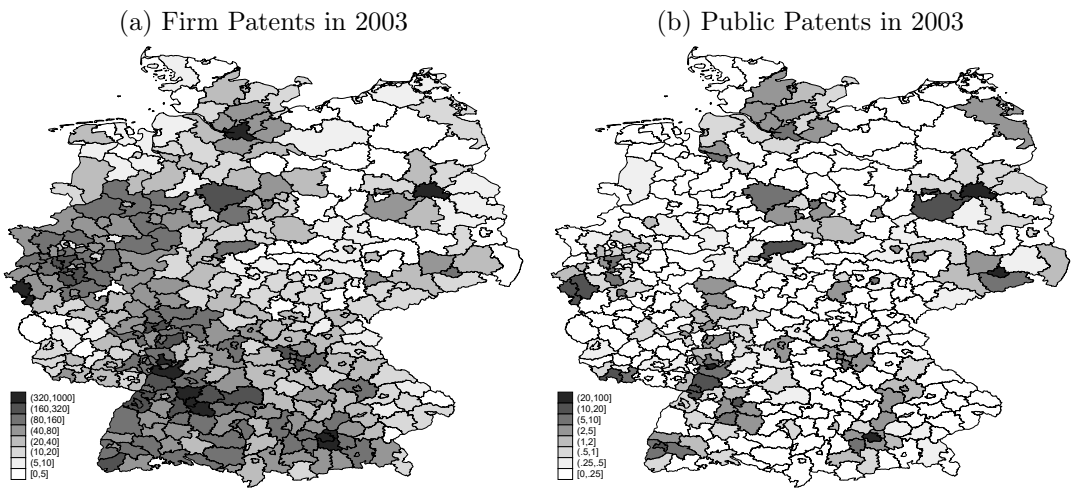


Figure 4.4: Number of Firm and Public Patents per County in 2013

Notes: Panel (a) shows the number of firm patent applications and panel (b) the number of public patent applications per county in 2003.
Source: Authors' calculations based on OECD RegPat database, 2003.

were valid in Germany in 2017, 80% of them were granted by the EPO.[26] This is very comparable to the ratio of EPO and PCT patent applications to overall patents for the research institutes, reported in Table 4.1. Second, Dechezlepretre et al. (2016) show that the impact of the UK R&D tax credit on the number of national and EPO patent applications is of similar magnitude.

4.4 Transmission Channels of Local Knowledge Spillovers of Public R&D

In the first part of the empirical analysis, we aim to quantify the relative magnitude of three different local public knowledge spillover transmission channels. Since this requires information about which public patents are relevant for firms, we start by investigating which observed public patent characteristics are a strong predictor for the number of (regional as well as non-regional) forward citations received, which is the most prominent measure for the technological value of a patent (see, for example, Trajtenberg, Henderson, and Jaffe 1997). Moreover, we use the citation data to assess the role of reverse causality for local public R&D with respect to local firm R&D.

4.4.1 Citation Data Analysis

Methodology To study the impact of public patent characteristics on the number of citations received by firm patents, we estimate a negative binomial regression model. The model explicitly addresses the problem of overdispersion in the data (e.g. in the number of citations received by a public patent) compared to the Poisson model. Our estimation model reads as follows:

[26] Press notice of the German Patent Office (https://www.dpma.de/dpma/veroeffentlichungen/statistiken/patente/index.html, last accessed 19/12/2019).

$$C_{i,r} = f(X_i, \lambda_r, \omega_t) + \epsilon_{i,r} \tag{4.1}$$

The dependent variable is the number of forward citations ($C_{i,r}$) of public patent i received by corporate patents that are produced in the same region as the public patent (r) and that cite the public patent within 5 years after publication of the public patent application.[27] We include only citations by corporate patents (i.e. patents with at least one firm applicant) and weight the citation by the share of firm applicants. Moreover, we focus on citations within 5 years to avoid that older public patents have a higher likelihood of receiving more citations by definition (since they exist for longer). Since we have patent data up to 2015, our sample covers public patents with application dates between 1995 and 2010.[28] To inform about differences of local and non-local technological spillovers, we construct also the number of non-regional citations by public patent i.

Our main explanatory variables of interest (X_i) are whether the public patent has been developed jointly with a firm (which we assume to be the case if the patent has been filed jointly) as well as three (public) patent characteristics that have shown to matter in prior literature (although not necessarily consistently). These are the (i) number of backward patent citations, (ii) the number of backward non-patent citations as well as (iii) the scope of a patent, i.e. the number of patent classes a patent is filed for. Trajtenberg, Henderson, and Jaffe (1997) report that university patents, which are supposedly more basic and thus are likely to have a higher technological value, have less backward patent citations and receive more forward citations. This is consistent with the argument that more backward citations indicate a more incremental innovation (Lanjouw and Schankerman 2001). When using a regression approach, Trajtenberg, Henderson, and Jaffe (1997) find, however, a positive correlation of the number of backward and forward citations similar to Harhoff, Scherer, and Vopel (2003). The number of non-patent citations has not been used in many studies yet, but Cassiman, Veugelers, and Zuniga (2008) argue that these citations might indicate a more complex and fundamental knowledge. Thus, patents that cite more non-patent literature should receive a larger number of forward citations, which is supported by the results by Branstetter (2005). Lastly, patents with a wider scope have been found to increase firm value (Lerner 1994), although they do not seem to have a higher monetary value (based on survey data) (Harhoff, Scherer, and Vopel 2003). Following prior literature, when entered into the regression, we scale all three measures by the maximum value of the respective variable in the same IPC class and year, and these measures are each labeled with the term "index" (see Table 4.2).

Our set of control variables includes time-fixed effects (ω_t, based on the application date of the cited public patent) and 3-digit IPC code (of the cited public patent) fixed effects in all regressions. In a sensitivity check, we also include main-inventor region-fixed

[27] We assume as outlined in Section 4.3 that a patent is produced in the same region if at least one inventor of the public patent lives in the same county as one inventor of the citing patent. Since inventors can live in different regions, a patent is not only produced in one region. We do not account for the strength of the regional link.

[28] We choose 1995 as the first year as universities and research institutes in East Germany had most likely completed their transmission process after the German re-unification.

	All Public Patents		Cited Public Patents	
	Mean	Median	Mean	Median
# Non-regional citations	0.49	0.00	2.02	1.00
# Regional citations	0.03	0.00	0.12	0.00
Index Patent Citations (PatCit)	0.14	0.00	0.13	0.00
Index Non Patent Citations (Non-PatCit)	0.26	0.00	0.22	0.00
Index Patent Classes (PatClasses)	0.31	0.27	0.37	0.27
Joint Venture with Firm (JV)	0.15	0.00	0.17	0.00
Patent Class A	0.15	0.00	0.14	0.00
Patent Class B	0.12	0.00	0.14	0.00
Patent Class C	0.27	0.00	0.28	0.00
Patent Class D	0.00	0.00	0.00	0.00
Patent Class E	0.01	0.00	0.01	0.00
Patent Class F	0.04	0.00	0.05	0.00
Patent Class G	0.24	0.00	0.21	0.00
Patent Class H	0.17	0.00	0.17	0.00
D(Cited = 1)	0.24	0.00	1.00	1.00
Observations	11,468		3,035	

Table 4.2: Descriptive Statistics for Citation Data Estimation Sample

Notes: Table shows descriptive statistics for public patent applications filed between 1995 and 2010 with respect to their type as well as the number of citations they received within 5 years by corporate patents. PatCit stands for the number of backward patent citations, Non-PatCit for the number of backward non-patent citations and PatClasses for the number of IPC classes. The index variables are the normalized versions of the variables, dividing each variable by the maximum value of the respective variable in the same IPC class and year. Patents (and thus patent applications) are grouped into patent classes. The broad sections are A: human necessities; B: performing operations, transportation; C: chemistry, metallurgy; D: textiles, paper; E: fixed constructions; F: mechanical engineering, lighting, heating, weapons, blasting; G: physics and H: electricity. For more information see www.wipo.int/edocs/pubdocs/en/wipo_guide_ipc_2019.pdf (last accessed 16/03/2020).
Source: Authors' calculations based on OECD RegPat and Citation database, 1995-2015.

effects (λ_r).[29]

Descriptive statistics for the full sample of public patents and the sample of public patents with at least one forward citation are shown in Table 4.2. The average (cited) public patent receives 0.03 (0.12) citations by regional firm patents and 0.49 (2.02) citations by non-regional firm patents within 5 years. While the absolute number of citations is low, the share of regional citations is with 6% substantial given that there are around 400 regions in Germany. Thus, the average regional share should - if all citations would come from within Germany and from all regions to the same extent - only amount to around 0.3%. The relatively low number of forward citations can be explained as follows: First, we consider only forward citations within 5 years by firm patents. When including all citations, the average number of citations increases to 1.3

[29]More precisely, we construct indicator variables for each region that are one if at least half of the inventors of a particular public patent are located in that region. When using dummy variables for each region in which at least one inventor is located, the model no longer converges.

(3.3) per (cited) public patent. Second, the number of cited patents by the average EPO patent application is less than 1/3 of the number of cited patents by the average US patent application (see Michel and Bettels 2001). Harhoff, Scherer, and Vopel (2003) report that granted patent applications by German inventors in 1977 have been cited by 0.6 EPO patents on average over 20 years. Taking into account that the number of cited patents by EPO patent application has fallen over time (see Webb et al. 2005), the average number of forward citations in our sample seems even relatively high. Third, it should be noted that only a small subset of patents receives a large number of forward citations. This is not only in our sample in which only one in every four public patents is at least cited once but also on average (see OECD 2015).

Comparing the average public patent and the average public patent that is cited at least once shows that the latter has slightly less backward citations (0.13 vs. 0.14), less non-patent citations (0.22 vs. 0.26) and is filed for more patent classes (0.37 vs. 0.31). Moreover, the average public patent that is cited at least once is more likely to result from a collaboration with a firm (0.17 vs. 0.15). The distribution of the main IPC class is, however, very similar.

Results The results of the citation analysis are reported in Table 4.3. In the first three columns, the dependent variable is the number of regional citations. Column (1) shows the results of our baseline specification which controls only for time and 3-digit IPC class fixed effects. In column (2) we include time and main-inventor region-fixed effects. The results are very similar and suggest that public patents that are filed jointly with firms, with a larger number of backward patent citations and with a wider patent scope receive more forward citations. The number of backward non-patent citations does not matter. Since the number of backward patent citations could simply pick up an additional effect for joint venture public patents, we add interaction effects with the two (significant) patent characteristics and the indicator variable for joint venture public patents in column (3). The result supports this presumption. While a wider patent scope increases the number of regional citations for all public patents, the number of backward patent citations matters only for joint venture patents.

To assess whether the two identified public patent characteristics matter only for regional or also for non-regional citations, we use the number of non-regional citations in column (4) and the share of regional citations in column (5) as dependent variable (and an OLS estimator). While joint venture public patents increase in particular regional citations, public patents with a wider scope receive both, more regional as well as non-regional citations. Thus, two insights emerge from the analysis. First, the scope of a public patent is a reasonable proxy for its technological value, as measured by the number of forward citations. We will use it to account for the quality of patents. A second insight is that joint ventures have a substantial local dimension, which is consistent with prior work on industry-university collaborations (see, for example, the overview article by Rybnicek and Königsgruber 2019).

We now investigate whether public patents that are of higher relevance for local firm R&D build more strongly on local firm knowledge. The methodology is the same as above, but instead of using the citations received we use the number of cited firm patents by a particular public patent as the dependent variable. The sample period thus shifts to 2000 to 2015. The average public patent cites 0.02 regional patents and

Model	Negative Binomial				OLS
Dep. Var.	# Regional Citations			# Non-Regional Citations	Share Regional Citations
	(1)	(2)	(3)	(4)	(5)
Index PatCit	0.595**	0.608**	-0.035	-0.307**	0.048**
	(0.239)	(0.263)	(0.325)	(0.156)	(0.023)
Index Non-PatCit	-0.056	-0.103	-0.038	-0.169*	0.010
	(0.175)	(0.168)	(0.163)	(0.102)	(0.014)
Index PatClasses	0.654**	0.731***	0.772**	0.732***	-0.015
	(0.277)	(0.247)	(0.329)	(0.140)	(0.020)
JV	1.468***	1.436***	1.351***	0.180*	0.112***
	(0.146)	(0.156)	(0.304)	(0.106)	(0.016)
Index PatCit*JV			1.275***		
			(0.479)		
Index PatClasses*JV			-0.220		
			(0.522)		
Observations	11,468	11,468	11,468	11,468	3,035
3digit-IPC-FE	x		x	x	x
Inventor-Region-FE		x			
Time-FE	x	x	x	x	x

Table 4.3: Results for Number of Citations *OF* Public Patents

Notes: Table shows point estimates of negative binomial model (column (1) to (4)) and OLS model (column (5)) for the impact of public patent characteristics on the number of regional forward citations (columns (1) to (3)), non-regional forward citations (column (4)) and the share of regional forward citations (column (5)). Citations include only citations by firms (applicant-weighted) within 5 years after the publication of the public patent application. All regressions include a full set of time dummies (based on the application date of the public patent). PatCit stands for the number of backward patent citations, Non-PatCit for the number of backward non-patent citations and PatClasses for the number of IPC classes. The index variables are the normalized versions of the variables, dividing each variable by the maximum value of the respective variable in the same IPC class and year. JV stands for joint venture between firms and public institutions. Bootstrapped standard errors in parenthesis. *,**,*** denote statistical significance at the 10, 5 and 1% level.
Source: Authors' calculation based on OECD RegPAT and Citation database, 1995-2015.

Model	Negative Binomial				OLS
Dep. Var.	# Regional Cited Patents			Non-Regional Cited Patents	Share Regional Cited Patents
	(1)	(2)	(3)	(4)	(5)
Index PatClasses	-0.331	-0.315	-0.241	-0.054	-0.003
	(0.389)	(0.416)	(0.612)	(0.099)	(0.013)
JV	2.094***	1.928***	2.139***	0.251***	0.066***
	(0.158)	(0.172)	(0.253)	(0.043)	(0.008)
Index PatClasses*JV			-0.151		
			(0.648)		
Observations	147,88	14,788	14,788	14,788	3,966
3digit-IPC-FE	x		x	x	x
Inventor-Region-FE		x			
Time-FE	x	x	x	x	x

Table 4.4: Results for Number of Citations *BY* Public Patents

Notes: Table shows point estimates of negative binomial model (column (1) to (4)) and OLS model (column (4)) for the impact of public patent characteristics on the number of cited regional (column (1) to (3)), cited non-regional (column (4)) and the share of cited regional corporate patents (column (5)). The number of cited patents includes only corporate patents (applicant-weighted) within 5 years after the publication of the corporate patent application. All regressions include a full set of time dummies (based on the application date of the public patent). PatClasses stands for the number of IPC classes. The index variable is the normalized version of this variable, dividing the variable by the maximum value of the variable in the same IPC Class and year. JV stands for joint venture between firms and public institutions. Bootstrapped standard errors in parenthesis. *,**,*** denote statistical significance at the 10, 5 and 1% level.
Source: Authors' calculation based on OECD RegPAT and Citation database, 1995-2015.

0.54 non-regional patents. When taking all citations into account and ignoring the time dimension the average public patent cites 1.8 patents. The regression results (see Table 4.4) suggest that public patents with a wider scope do not cite more regional firm patents, which mitigates concerns about reverse causality. Moreover, we find that joint venture public patents are more likely to cite regional firm patents, which is in line with the notion that collaborations are mainly local.[30]

4.4.2 Transmission Channel Analysis

We now turn to the transmission channel analysis in which we investigate the knowledge spillovers of public R&D related to (i) a firm collaborating with a public institution (collaboration spillovers), (ii) a firm citing a public patent (technological spillovers) and (iii) a firm locating in a region in which public patents are produced (non-specific knowledge spillovers).

Methodology To assess the impact of the three channels, we use a negative binomial model again. It reads as follows:

$$CP_{i,r,t} = f(c_{i,r}, JV_{i,r,t}, Cited_{i,r,t}, PP_{r,t}, X_{r,t}, \omega_t) + \epsilon_{i,r,t} \tag{4.2}$$

The dependent variable is the number of corporate patents filed by firm (applicant) i in region r at time t excluding joint venture patents with public institutions. The level of analysis is thus at the applicant-region level and includes the years 2000 to 2010 for all applicant-region combinations in which at least one corporate patent application has been filed between 1995 and 2015.[31]

Our three main explanatory variables capture the above mentioned transmission channels considered in our analysis. First, to capture collaboration spillovers we include an indicator variable that is one if applicant i in region r filed a patent with a public institution jointly ($JV_{i,r,t}$) in year t. Second, to capture technological spillovers we include an indicator variable that is one if the applicant i in region r cited a public patent in a patent application in t ($Cited_{i,r,t}$).[32] Third, we use the number of public patents in region r in year t to capture non-specific knowledge spillovers ($PP_{r,t}$) as this knowledge is likely to be linked to the number of innovations.[33] Since we expect that the non-specific knowledge spillovers require firms to have close links with public institutes in order to access the non-specific knowledge, we interact the number of public patents in a region with an indicator variable that is one if applicant i in region r has collaborated with a public institution in the past or cited a public patent.

[30] While it would be of interest to assess the link explicitly, the patent data does not allow us to do so. Since inventors cannot be matched to applicants, the true firm R&D location cannot be determined.

[31] We exclude applicant-region combinations in which less than one corporate patent application has been filed to increase the efficiency of the estimates.

[32] More precisely, we use the number of joint ventures as well as the number of citations of public patents. For most of the firms, however, only one collaboration or one citation is observed.

[33] One alternative indicator for the non-specific public knowledge suggested in the literature is the number of scientific publications (see, for example, Narin, Hamilton, and Olivastro 1997; Van Looy, Callaert, and Debackere 2006). While we observe scientific publications of research institutes and universities in the backward citations of patent application data, we refrain from using them in our analysis as this would induce a systematic measurement error (as only relevant publications are observed).

To assess the empirical support for the considered channels, we use a slightly different sample for each channel (but always account for all three channels). This is motivated by two reasons: First, when analyzing the citation channel for the dependent variable we do not use the overall number of corporate patents (applicant and inventor weighted) but only the corporate patents that do not cite public patents as otherwise there would be a mechanical link. Second, applicants that collaborate with public institutions or cite public patents might be very different from firms that do not (see, for example, the discussion in Scandura (2016) for industry-university collaborations). Thus, by choosing a particular sample of firms we aim to address at least partly a potential selection. Our sample to assess the evidence for collaboration/joint venture (citing a public patent) spillovers includes only firms that collaborated with a public institution (cited a public patent) in at least one region between 1995 and 2015 (JV and CITED sample). To assess the evidence for non-specific knowledge spillovers we use the full sample.[34] Moreover, to investigate the impact of public patent quality for the three channels, we include in additional specifications 1 less the quality indicator, based on the patent scope (number of IPC classes). The main advantage of this strategy is that it does not assume that we measure quality correctly.

Our set of control variables in all estimations includes applicant-region-fixed effects ($c_{i,r}$) as well as time-fixed effects (ω_t).[35] Further, we control for a wide range of region characteristics ($X_{r,t}$) following prior literature (e.g. Abel and Deitz 2011; Abramovsky, Harrison, and Simpson 2007). These include measures for the local labor market (population, unemployment rate, share of high skilled and share of low skilled workers, and number of students (per 1.000 inhabitants)) as well as measures for agglomeration (manufacturing employment) and urbanization (share of manufacturing employment). Moreover, we account for the tax burden on properties and business income in a region.[36] Lastly, we account for the fact that during our sample period, Belgium, the Netherlands and Luxembourg have introduced patent boxes in 2006 and 2007. We define two indicator variables that are one if a county belongs to the 25% closest counties to Belgium or the Netherlands (or Luxembourg). The indicator variables are interacted with the respective reform dummies.

Descriptive statistics for the three samples are reported in Table 4.A.1 in the Appendix. They suggest that applicants that collaborate with a public institution or cite a public patent are very similar and somewhat different from the average applicant. The average number of corporate patents in the JV and CITED sample is 0.51 but only 0.29 in the full sample. Further, the average county population of a collaborating as well as citing applicant is somewhat smaller than for the average applicant (270,000 vs. 300,000). The same is true for the average number of public patents (4.6 vs 5.2). In the JV sample, the mean of the number of joint ventures is 0.02 (compared to 0.01 in the CITED sample).

[34] Since the negative binomial model did not converge for the full sample, we report the estimates of a Poisson model with robust standard errors. For the negative binomial model, we report bootstrapped standard errors.

[35] The applicant-region-fixed effects are included by estimating a conditional negative binomial model since the number of fixed effects is too large to include them in a negative binomial model. While the conditional negative binomial is not a real fixed effect model, the results are very similar when using a fixed effect Poisson model. Results are available upon request.

[36] Since the tax rates are municipality-specific, we construct the county-averages using the municipalities' population in 1999.

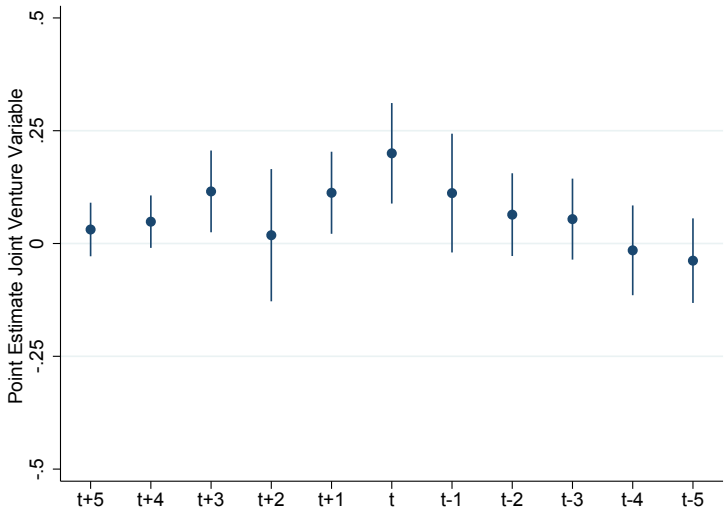


Figure 4.5: Estimates for JV - Not Quality Adj. and All Firms

Notes: Figure shows point estimates for 5 leads and 5 lags of the joint venture (JV) variable. The dependent variable is the number of corporate patent applications (applicant and region-weighted). The specification uses the same sample and includes the same control variables as the specification shown in column (1) in Table 4.5. More precisely, the sample includes all firms and the joint venture variable is not quality adjusted.
Source: Authors' calculations based on INKAR and OECD RegPat database, 2000 to 2010.

Results In the following we present the regression results. We start with the impact of joint venture public patents on the collaborating firm patenting behavior. To rule out that firms that collaborate with public institutions exhibit a different patent application pattern before the joint venture, we start by estimating a more flexible specification that includes 5 leads and 5 lags for the joint venture variable. The resulting point estimates highlight two aspects (see Figure 4.5). First, there is evidence that in the year of the joint application as well as the year before and after, the collaborating firm files more corporate patent applications in the region in which it collaborated. The anticipation effect is not surprising as the collaboration is likely to have started before the year of the joint application. Second, the effects for the years t+5 to t+2 are close to zero (with the exception of t+3). Thus, there is some evidence that collaborating firms have similar pre-trends compared to firms that collaborated with public institutions at a different point in time. However, this does not imply that our estimates represent causal effects as they could still be biased due to, for example, a self-selection of firms.

To assess the role of public patent quality, heterogeneity of the impact as well as possible spillovers to other regions, we rely on a regression analysis in the following and focus only on the same-year impact of joint ventures (acknowledging that this is likely to underestimate the overall impact of this particular channel). Column (1) in Table 4.5 shows the results of our baseline specification, which replicates the results reported in Figure 4.5: Collaborating with a public institution coincides with an increase in the number of corporate patent applications. The magnitude of the impact (around 20%) is very similar to the results by Scandura (2016), who reports an increase in R&D expenditure per employee of 17% for firms that collaborate with a university (at the

Model	Negative Binomial					
Dep. Var.	# Corporate Patents (CP)					
Firms	All	All	Low R&D	High R&D	All	All
Patent Quality	All	All	All	All	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
# JV	0.204***	0.465***	0.774***	0.288**	0.638*	0.478***
	(0.043)	(0.123)	(0.281)	(0.145)	(0.341)	(0.156)
# [JV*(1-QI)]		-0.383**	-0.598	-0.224	-0.526	-0.421*
		(0.182)	(0.425)	(0.210)	(0.494)	(0.233)
# JV OR	0.004***	-0.011	-0.126	-0.018	0.128	0.025
	(0.001)	(0.011)	(0.115)	(0.016)	(0.315)	(0.046)
# [JV OR*(1-QI)]		0.022	0.161	0.032	-0.136	-0.029
		(0.015)	(0.176)	(0.036)	(0.440)	(0.089)
Observations	73,395	73,395	20,685	52,710	73,395	73,045
Applicant-Region-FE	x	x	x	x	x	x
Time-FE	x	x	x	x	x	x
Control Variables	x	x	x	x	x	x
Dep. Var CP	0.89	0.89	0.51	1.23	0.28	0.62
CP OR	29.92	29.92	2.91	54.20	8.45	21.59
Absolute increase in the number of corporate patents if JV increases by one						
with average QI	0.18					
with QI = 1		0.42	0.39	0.35	0.17	0.30

Table 4.5: Results for Impact of JV of Firms with Public Institutions

Notes: Table shows point estimates of negative binomial model for the impact of collaborating with a public institution (measured by filing a joint public patent application) on the number of corporate patents (applicant and region-weighted). JV stands for joint venture, QI for quality index and OR for other regions. All regressions include time and applicant-region-fixed effects as well as the control variables described in the text. Column (1) shows the baseline specification. Column (2) accounts for the quality of the jointly filed public patent. Column (3) uses only low R&D intensity firms (less than 5 patent applications between 1995 and 1999) and column (4) only high R&D intensity firms. In column (5) the dependent variable is the number of high quality corporate patents (sum over corporate patents multiplied with the quality index) and in column (6) the number of low quality corporate patents (sum over corporate patents multiplied with 1 less quality index). The quality index is based on the number of patent classes a patent is filed for. For more information see the text. The absolute increase in the number of corporate patents if JV increases by one is calculated by multiplying the JV coefficient with the mean of the dependent variable (shown in the bottom of the Table as well). Bootstrapped standard errors in parenthesis. *,**,*** denote significance at the 10, 5 and 1% level.
Source: Authors' calculation based on INKAR and OECD RegPAT, 2000-2010.

end of the collaboration). In column (2), we account for the quality of the joint venture public patents, measured by their scope. In line with the results of the citation analysis, the point estimate for JV * (1-Quality Index) is negative and statistically significant. Thus, higher quality joint venture public patents have a higher value to the collaborating firm. In column (3) we only use low R&D intensity firms (firms with less than 5 patent applications between 1995 and 1999) and in column (4) only high R&D intensity firms. The results suggest that the absolute increase in the number of patent applications for one additional joint venture public patent with a quality index of one is very similar for low and high R&D intensity firms (0.39 vs 0.35, see bottom of Table 4.5).[37] Column (5) and (6) show that joint ventures have a larger impact on the absolute number of low quality firm patents, measured by their scope (0.30 vs. 0.17, see bottom of Table 4.5). A further insight relates to the extent of joint venture spillovers to other regions. We capture them by including the applicant-year specific number of joint ventures in all other regions. While our baseline specification suggests significantly positive spillovers to other regions, substantial heterogeneity seems to exist as in all other specifications the spillover effects are not statistically different from zero.

We turn to the results for the impact of citing a public patent on the citing firms' number of patent applications. As for joint ventures, we start presenting the results of the more flexible specification that includes 5 leads and 5 lags of the CITED variable (see Figure 4.6). The results suggest that firms that cite a public patent in a particular year file at the same time more patent applications (that do not cite a public patent). As for joint ventures, there is some evidence for positive effects in the year before and after. The point estimate for t+4 to t+2 are, however, very similar (although somewhat larger than zero), which suggests that the patenting behavior of firms before they cite a public patent is comparable to firms that cite public patents at a different point in time. However, as before this does not necessarily imply causality as self-selection could still bias our estimates.

The regression results for the (current year) impact of citing a public patent are reported in Table 4.6. The baseline specification (column (1)) replicates the finding reported in Figure 4.6. The point estimate is only smaller which is in line with the somewhat larger number of patent applications before the citation. Accounting for quality differences matters (see column (2)), the interaction effect is negative and statistically significant. Columns (3) and (4) show that the absolute increase in the number of corporate patents for citing one additional public patent is somewhat larger for high R&D intensity firms (1.1 vs. 0.7, see bottom of Table 4.6). Moreover, the results in columns (5) and (6) suggest that the absolute impact is larger for low quality patents (0.53 vs. 0.28). With respect to spillovers into other regions, we find evidence for positive ones (columns (1) and (2)). Surprising is, however, that the spillovers into other regions are significantly larger for low R&D intensity firms if the cited public patent has a lower quality (based on its patent scope, see column (3)). Thus, the spillovers within the firm from citing a low quality public patent are less localized compared to the spillovers within the firm from citing a high quality public patent. One potential adhoc explanation for this finding is that public R&D producers not only produce public knowledge but also provide important inputs for the firm R&D that is stimulated by high quality public R&D,

[37] The absolute increase is calculated by multiplying the point estimate with the mean of the dependent variable, shown in the bottom of Table 4.5.

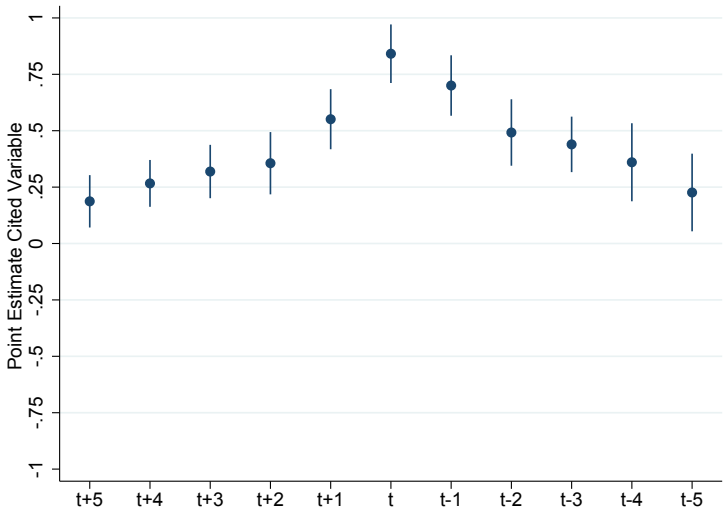


Figure 4.6: Estimates for CITED - Not Quality Adj. and All Firms

Notes: Figure shows point estimates for 5 leads and 5 lags of the CITED variable. The dependent variable is the number of corporate patent applications (applicant and region-weighted). The specification uses the same sample and includes the same control variables as in column (1) in Table 4.6. More precisely, the sample includes all firms and the CITED variable is not quality adjusted.
Source: Authors' calculations based on INKAR, OECD RegPat and Citation database, 2000 to 2010.

potentially due to the provision of research infrastructure or due to their non-specific knowledge.

Lastly, we explore whether there are non-specific knowledge spillovers of public R&D. As before, we focus on the current year impact.[38] The result, reported in Table 4.7, suggests so (additional to the collaboration and technological spillovers for which we control), assuming that the number of public patents is a reasonable proxy for them and that the additional firm patents are indeed caused by the additional public knowledge. While the effect is stronger for firms with close links to public institutions (CL), all other firms respond as well (1-CL). Accounting for quality does not seem to matter (column (2)). The reason is, however, that there is a strong difference between the impact on low (column (3)) and high R&D intensity firms (column (4)). While the first respond strongly, the latter do not seem to respond at all. Columns (5) and (6), which distinguish between high and low quality patents, also do not show meaningful results, which is again likely to be related to the striking difference between low and high R&D intensity firms. We thus estimate the specifications in columns (5) and (6) for low and high R&D intensity firms separately (not reported). The results suggest that both, the number of low and high quality patents by low R&D intensity firms, increases.

Finally, while there is no evidence for spillovers into other regions on average, there is evidence - similar to the citation channel - for positive spillovers into other regions for low R&D intensity firms (that have close links to public institutions) in the case of

[38] Unfortunately, estimating the more flexible specification is not possible as the model no longer converges.

Model	Negative Binomial					
Dep. Var.	# Corporate Patents (CP)					
Firms	All	All	Low R&D	High R&D	All	All
Patent Quality	All	All	All	All	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
# Cited	0.639***	0.909***	1.135***	0.730***	0.925***	0.741***
	(0.040)	(0.141)	(0.204)	(0.219)	(0.175)	(0.121)
# [Cited*(1-QI)]		-0.408**	-0.859***	-0.290	-0.489*	-0.215
		(0.192)	(0.292)	(0.289)	(0.250)	(0.164)
# Cited OR	0.033***	0.030**	0.008	0.019	0.041***	0.011
	(0.004)	(0.014)	(0.051)	(0.014)	(0.013)	(0.017)
# [Cited OR*(1-QI)]		0.003	0.164**	0.030	-0.018	0.021
		(0.024)	(0.075)	(0.022)	(0.021)	(0.027)
Observations	89,251	89,251	28,918	60,333	89,251	88,999
Applicant-Region-FE	x	x	x	x	x	x
Time-FE	x	x	x	x	x	x
Control Variables	x	x	x	x	x	x
Dep. Var. CP	1.026	1.026	0.631	1.527	0.299	0.727
CP OR	22.747	22.747	3.726	46.852	6.196	16.564
Absolute increase in the number of corporate patents if Cited increases by one						
with average QI	0.66					
with QI = 1		0.93	0.72	1.11	0.28	0.53

Table 4.6: Results for Impact of Citing a Public Patent

Notes: Table shows point estimates of negative binomial model for the impact of citing a public patent application (within 5 years after its publication) on the number of corporate patents (applicant and region-weighted). Cited stands for citing a public patent, QI stands for quality index and OR for other regions. All regressions include time and applicant-region-fixed effects as well as the control variables described in the text. Column (1) shows the baseline specification. Column (2) accounts for the quality of the public patents in a region. Column (3) includes only low R&D intensity firms (less than 5 patent application between 1995 and 1999) and column (4) only high R&D intensity firms. In column (5) the dependent variable is the number of high quality corporate patents (sum over corporate patents multiplied with the quality index) and in column (6) the number of low quality corporate patents (sum over corporate patents multiplied with 1 less quality index). The quality index is based on the number of patent classes a patent is filed for. For more information see the text. The absolute increase in the number of corporate patents if Cited increases by one is calculated by multiplying the Cited coefficient with the mean of the dependent variable (shown in the bottom of this Table as well). Bootstrapped standard errors in parenthesis. *,**,*** denote significance at the 10, 5 and 1% level.
Source: Authors' calculation based on INKAR, OECD RegPAT and Citation database, 2000-2010.

Model	Poisson					
Dep. Var.	# Corporate Patents (CP)					
Firms	All	All	Low R&D	High R&D	All	All
Patent Quality	All	All	All	All	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
PP*CL	0.012***	-0.004	0.029	-0.013	0.003	-0.008
	(0.005)	(0.017)	(0.025)	(0.020)	(0.017)	(0.018)
PP*CL*(1-QI)		-0.013	-0.033***	-0.010	-0.011	-0.012
		(0.008)	(0.010)	(0.015)	(0.009)	(0.009)
PP*(1-CL)	0.003*	0.013**	0.030***	0.005	0.011	0.014**
	(0.002)	(0.006)	(0.008)	(0.011)	(0.007)	(0.007)
PP*(1-CL)*(1-QI)		0.019	-0.038	0.029	0.008	0.025
		(0.022)	(0.031)	(0.026)	(0.022)	(0.023)
PP OR*CL	0.000	-0.000	-0.004	-0.002	-0.000	-0.001
	(0.000)	(0.002)	(0.006)	(0.002)	(0.002)	(0.002)
PP OR*CL*(1-QI)		0.001	0.019**	0.004	0.000	0.002
		(0.002)	(0.007)	(0.002)	(0.002)	(0.002)
Observations	564,698	564,698	410,735	153,963	564,643	561,805
Applicant-Region-FE	x	x	x	x	x	x
Time-FE	x	x	x	x	x	x
Control Var.	x	x	x	x	x	x
Dep. Var. CP (CL=1)	0.888	0.888	0.554	1.275	0.265	0.625
Dep. Var. CP (CL=0)	0.271	0.271	0.212	0.434	0.072	0.200
CP OR (CL=1)	23.785	23.785	3.270	47.538	6.647	17.198
Absolute increase in the number of corporate patents if public patents increases by one						
Firms with close links (CL=1)						
with average QI	0.01					
with QI = 1		0.00	0.02	-0.02	0.001	-0.005
Firms with no close links (CL=0)						
with average QI	0.001					
with QI = 1		0.004	0.006	0.002	0.001	0.003

Table 4.7: Results for Residing in a Location with Public Patents

Notes: Table shows point estimates of Poisson model for the impact of residing in a location with public patents on the number of corporate patents (applicant and region-weighted). PP stands for the number of public patent applications, QI stands for quality index, CL for close links to a public institution and OR for other regions. All regressions include time and applicant-region-fixed effects as well as the control variables described in the text. Column (1) shows the baseline specification. Column (2) accounts for the quality of public patents. Column (3) includes only low R&D intensity firms (less than 5 patent applications between 1995 and 1999) and column (4) only high R&D intensity firms. In column (5) the dependent variable is the number of high quality corporate patents (sum over corporate patents multiplied with the quality index) and in column (6) the number of low quality corporate patents (sum over corporate patents multiplied with 1 less quality index). The quality index is based on the number of patent classes a patent is filed for. The absolute increase in the number of corporate patents if the number of public patents increases by one is calculated by multiplying the PP coefficient with the mean of the dependent variable (shown in the bottom of this Table as well). Robust standard errors in parenthesis. *,**,*** denote significance at the 10, 5 and 1% level.
Source: Authors' calculation based on INKAR, OECD RegPAT and Citation database, 2000-2010.

low quality public R&D.[39] As mentioned above, one potential reason for this finding could be that public R&D producers provide important inputs for the production of firm R&D that is stimulated by high quality public R&D.

Discussion In the following, we compare the results for the different channels and quantify their relative importance. This exercise is based on the assumption that all three channels are correctly measured and that the estimated effects are causal. We find that all three channels have a stronger effect on firm R&D the higher the quality of public R&D and that all three channels have a larger impact on the absolute number of low quality corporate patents. The main difference between the channels is the type of firms affected. While the citation and collaboration channel seem to affect low and high R&D intensity firms largely to a similar extent, non-specific knowledge spillovers affect predominantly low R&D intensity firms.

Based on our empirical strategy and the proxies employed, the non-specific local knowledge spillovers seem most important in quantitative terms. While the average effect is small (semi-elasticity of 0.012 for firms with close links to public institutions and 0.003 for firms with no close links, see column (1) in Table 4.7), the number of firms affected in the average region with non-zero public patents is large; on average there are 8 firms with close links and 170 other firms. Thus, the average public patent generates via non-specific knowledge spillovers 0.08 additional public patents by firms with close links (absolute increase by 0.01 for 8 firms, see column (1) of Table 4.7) and 0.17 additional patents by firms with no links to public institutions (absolute increase by 0.001 for 170 firms, see column (1) of Table 4.7). The collaboration as well as the citing channel are suggested to be less important. The number of additional local corporate patents generated via the citation link amounts to 0.03 (calculated by multiplying the absolute increase of 0.66 per additional public patent, see column (1) in Table 4.6, with the number of regional applicants that cite a public patent on average, which is 0.046). The number of additional local corporate patents generated via the collaboration channel amounts to 0.04 (calculated by multiplying the absolute increase of 0.18, see column (1) of Table 4.5, the share of joint venture public patents, 0.15 - assuming that all collaborations are local, and the number of collaborating firms per joint venture, which is 1.3). Adding all (one year) effects up suggests that one additional average public patent in a region generates 0.32 local firm patents (at different points in time). This implies - based on the assumption outlined above - that the non-specific knowledge channel accounts for around 3/4 of the overall effect. Since Walsh, Lee, and Nagaoka (2016) show that only around 1/3 of all collaborations lead to joint patent applications, 3/4 should be considered as an upper bound. A lower bound may be derived by assuming that the collaboration channel is three times as important, while keeping the overall effect constant. In this case, the non-specific knowledge spillovers would account for around 1/2 of the overall effect.

[39] Since there is no independent variation in the data, we are not able to assess the spillovers into other regions for firms with no close links to public institutions.

4.5 Overall Magnitude of Local Public R&D Spillovers

Methodology In the second part of the empirical analysis, we aim to quantify the overall magnitude of local knowledge spillovers and whether they are - as suggested by the transmission channel analysis - dominantly driven by low R&D intensity firms. The analysis is at the county level and we employ an instrumental variable strategy since the OLS estimate may be biased due to measurement error (as the number of public patents is only an imperfect proxy for public knowledge production), reverse causality (as the funding of public R&D may depend on firm R&D in a region) as well as omitted variables (e.g. the emergence of new technologies or direct subsidies).

While the direction of the overall bias cannot be determined with certainty a priori given the different reasons, we believe that the OLS estimate is most likely downward biased. This is based on the rationale that the measurement error, which arguably causes a downward bias, is likely to be severe due to our focus on patents and unobserved patent quality but also as we exploit changes over time. Reverse causality and omitted variable bias, in contrast, are likely to be less important given that we exploit variation over time. Reverse causality is likely to be of minor importance as the institutional funding for a particular institute is not only determined by the hosting state but by all other German states and the federal states.[40] Moreover, direct subsidies, which are a strong candidate for an omitted variable bias, are relatively less important in Germany. While the federal and the state governments fund 19 billion EURO R&D expenditures in the government and higher education sector, they fund only 2 billion EURO R&D expenditures in the business sector (out of 47 billion EURO).[41]

Based on these considerations, we choose the 4-year lagged institutional funding for Fraunhofer and Leibniz institutes that engage in R&D activities as the excluded instrument.[42] This instrument is very likely to successfully address the bias resulting from measurement error as the institutional funding is a second measure of public R&D production and is unlikely to correlate with the measurement error in the number of public patents. Moreover, we believe our excluded instrument to at least partly mitigate

[40] In more detail, the argument is as follows: We believe that the funding shares of hosting and non-hosting states (as well as the federal state) reflect the (average) expected magnitude of local and technological spillovers. While the funding share of the hosting state increases with the expected local spillovers (as the hosting-state only benefits from them), it decreases in the latter (as all states benefit from this dimension). Assuming rational behavior of state governments then predicts that a state's (marginal) willingness to fund a research institute equals the (marginal) benefits of funding. This means that non-hosting states (as well as the federal state's) willingness to fund a particular research institute depends largely on the expected magnitude of the technological spillovers. Since there are 16 states and most institutes operate in only one state, for 15 (out of 16) states the decision about the level of funding of a particular research institute is independent of the expected magnitude of local spillovers.

[41] German Federal Ministry of Finance, Data portal, Table 1.1.1 (http://www.datenportal.bmbf.de/portal/de/K11.html, last accessed 19/12/2019).

[42] The data for Leibniz institutes have been obtained from annual publications of the Leibniz Society. The data for Fraunhofer institutes have been provided by the Fraunhofer Society. We also collected information for Helmholtz institutes and have been provided with data for Max Planck institutes by the Max Planck Society. For Max Planck institutes the data only starts in 2005 and for the Helmholtz institutes the data is less precise as Helmholtz institutes usually consist of several establishments and we were only able to collect institutional funding data on the institute level.

the bias that results from reverse causality and omitted variables. First, we focus on institutional funding instead of overall R&D expenditures to rule out an obvious reverse causality, which may exist as research institutes do carry out R&D on behalf of firms. Second, we use the 4-year lagged institutional funding since a time lag between the institutional funding decision and the number of firm patent applications reduces a potential correlation of institutional funding with the number of firm patent applications in a region, either directly or via omitted variables (e.g. the emergence of a new technology or direct subsidies).[43] While a longer period would certainly be better, our choice of the 4-year lagged institutional funding ensures that the instrument is still sufficiently strong. Moreover, the reader should note that in all specifications region-fixed effects are included to absorb a large part of potentially omitted variables and to account for the persistence in public as well as private R&D. To avoid that institute openings or closures affect our instrument, we construct it only for institutes that existed over our sample period.

Our estimation model reads as follows:

$$\ln CP_{r,t} = c_r + \beta_1 PP_{r,t} + \beta_2 PP_{r,t}^2 + \gamma X_{r,t} + t_t \epsilon_{r,t} \tag{4.3}$$

We use a log-linear model as it allows a straight forward implementation of the instrumental variable strategy and correct standard errors can easily be obtained. Moreover, count models with a large number of county-fixed effects tend to be somewhat fragile.[44] We report robust standard errors. Our sample period is - due to the availability of the excluded instrument - 2003 to 2010, which leaves us with 3,212 county-year observations for 402 counties (since for 4 observations the number of firm patents is zero).

Our dependent variable is (ln) number of corporate patents in region r at time t (both applicant and inventor weighted and including joint venture patents with public institutions). Our main explanatory variable of interest is the number of public patents in region r at time t. We use the quality unadjusted number of public patents as our instrument is unfortunately not strong enough to predict the quality adjusted number of public patents. To account for a potential nonlinear relationship we include the number of public patents linear and squared. The relationship could be nonlinear for a variety of reasons. First, there is evidence that the number of low R&D intensity firms is increasing with the number of public patents in a region (see Figure 4.7). Moreover, the number of firms that collaborate with public institutions in a region per public patent as well as the number of firms in a region that cite a public patent per public patent are decreasing with the number of public patents. The only dimension that does not seem to matter is the average quality of public patents.

Our set of control variables ($X_{r,t}$) is the same as for the transmission channel level analysis but with county (c_r) instead of applicant-region fixed effects. Descriptive statistics

[43] Direct subsidies provide a good illustration of this argument. While we find a positive and statistically significant correlation of the growth rate of overall direct subsidies and the growth rate of overall public funding for research institutes in the aggregate (correlation coefficient: 0.76, p-value: 0.03), no correlation exists between the growth rate of direct subsidies and the 4-year lagged growth rate of overall public funding for research institutes (correlation coefficient: -0.09, p-value: 0.82).

[44] We assessed the robustness of our estimation strategy by estimating a fixed effect Poisson model coupled with a control function approach in addition. Results are very similar and available upon request.

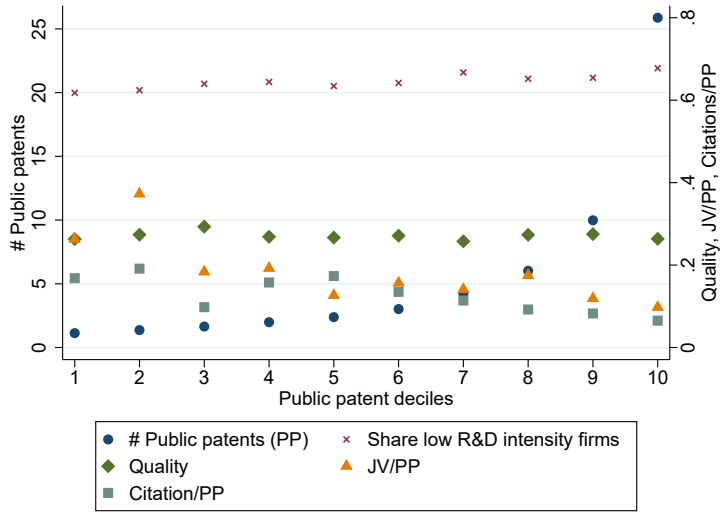


Figure 4.7: Reasons for a Potential Non-Linear Impact of Public on Firm R&D

Notes: Figure shows average number of public patents, average quality of public patents, share of low R&D firms, number of joint venture applicants per public patent and number of citing applicants per public patent within a region for public patent deciles (at least one public patent).
Source: Authors' calculation based on OECD RegPat and Citation database, 2003-2010.

for our sample, reported in Table 4.8, show that the average county has 200.000 inhabitants, around 64 corporate patents and 2 public patents, 30% of them are from universities and 70% from research institutes.

Results The regression results are reported in Table 4.9. Column (1) reports the OLS estimate and column (2) the IV estimate when including only the number of public patents. While the OLS point estimate is statistically significant, it is close to zero and would suggest an increase in the number of corporate patents by 0.3 for one additional public patent, which matches the result of the transmission channel analysis. The IV point estimate is substantially larger, but is not precisely estimated, despite the fact that the instrument is relevant, as suggested by the test-statistics, and has the expected sign (see bottom of Table 4.9). From column (3) onward, we include the number of public patents squared in addition. While in both, the uninstrumented (column (3)) and instrumented specification (column (4)), the squared number of public patents is negative and thus suggests a nonlinear relationship, only the IV point estimates are (marginally) significant. To test whether the impact of public on firm patents is indeed nonlinear, we employ the modified Ramsey's regression error specification test (RESET) by Pesaran and Taylor (1999), which tests for neglected non-linearities. The p-value of the test using 2nd and 3rd order polynomials are reported in the bottom of Table 4.9. While there are neglected non-linearities in the specification using only public patents (p-value below 10%, column (2)), this is not the case when using public patents linear and squared. Moreover, we assess whether the number of public patents squared simply picks up region-size heterogeneity as there are more public patents in larger regions. Column (5) shows the results when additionally including the interaction of

	Mean	P25	Median	P75	SD
# Corporate patents	63.86	13.39	32.73	79.35	96.85
# Public patents	2.02	0.00	0.39	1.58	5.63
# University patents	0.67	0.00	0.00	0.50	1.99
# Institute patents	1.35	0.00	0.17	0.91	4.06
Population in 1,000	202	106	148	239	231
Share skilled employees	4.03	2.10	3.00	4.70	3.23
Share unskilled employees	7.67	5.20	7.30	9.30	3.36
Students per 1,000 capita	27.83	0.00	0.70	36.60	50.19
Unemployment rate	7.04	4.50	6.20	8.90	3.37
Property tax multiplier	350	313	343	390	77
Business tax multiplier	359	329	349	399	61
Manufacturing employment in 1,000	37.75	16.51	26.77	46.14	33.46
Share of manufacturing employment	0.42	0.30	0.39	0.52	0.17
Close to Luxembourg	0.25	0.00	0.00	0.00	0.43
Close to Belgium or the Netherlands	0.34	0.00	0.00	1.00	0.47
Observations	3,216				

Table 4.8: Descriptive Statistics for County Level Estimation Sample

Notes: Table reports descriptive statistics for our county level estimation sample. Property tax and business tax multiplier are municipality-population weighted.
Source: Authors' calculation based on INKAR and OECD RegPAT database, 2003-2010.

the number of public patents with the population in 2002 (in million). While the size of a region seems to matter, more populated areas benefit more and not less from public R&D. Further, the point estimate for the squared number of public patents is largely unchanged. Thus, we conclude that the effect of public R&D is nonlinear and increases at a decreasing rate in the level of public patents.[45]

To assess the sensitivity of the results, we run several robustness checks (see Table 4.B.1 in the Appendix). First, we include state-year-fixed effects (column (1)). Second, we use (ln) number of R&D employees in a region as an alternative proxy for firm R&D. While the number of firm R&D employees is only available for 2003, 2005, 2007 and 2009, it is based on a full-assessment of firms' R&D activities in Germany. In both specifications point estimates are unchanged, only the precision is reduced. Our third robustness test excludes regions with an average number of corporate patents in the top 1% of the distribution (column (3)) to assess whether outliers are driving the results. The estimates are very similar and even more precisely estimated (significant at the 5% level), likely due to a higher instrument relevance. Lastly, we address whether agglomeration economies bias our results and include the number of corporate patent applications filed in a region between 1991 and 2000 interacted with year dummies (see column (4)). The results are basically unchanged.

Next we investigate which firms and which corporate patents respond to public R&D

[45]The effect of public R&D is even negative for 2% of the counties. Therefore the effect is not only nonlinear but non-monotone for these counties. Since the distribution of public patents is right-skewed (see Table 4.8), these counties represent a substantial share of all public patents in Germany.

Model	OLS	IV	OLS	IV	IV
Dep. Var.		ln(# Corporate Patents (CP))			
	(1)	(2)	(3)	(4)	(5)
# Public patents	0.004**	0.054	0.005	0.101*	0.115**
	(0.002)	(0.033)	(0.004)	(0.052)	(0.057)
# Public patents, sqrd.			-0.000	-0.002*	-0.003**
			(0.000)	(0.001)	(0.001)
# Public patents * Pop					0.073**
					(0.034)
Observations	3,212	3,212	3,212	3,212	3,212
Region-FE	x	x	x	x	x
Time-FE	x	x	x	x	x
Control Variables	x	x	x	x	x
P-value Underident.		0.002		0.027	0.039
F-Stat. Weak ident.		15		9	8
Dep. Var. CP	81.158	81.158	81.158	81.158	81.158
Point Estimate First Stage					
L4.ln(Inst. Funding)		1.687***		-23.673**	-17.764
		(0.438)		(11.597)	(11.218)
L4.ln(Inst. Funding), sqrd.				0.867**	0.608
				(0.396)	(0.412)
IV RESET 2nd Polynomial p-val.		0.313		0.680	
IV RESET 3rd Polynomial p-val.		0.044		0.890	

Table 4.9: IV Results

Notes: Table shows estimated coefficients for the impact of the number of public patent applications in a county on the number of firm patent applications in the same county. All regressions include a full set of county as well as time-fixed effects. The dependent variable is (ln) number of firm patent applications. In columns (1) and (2) we only include the linear number of public patents and in columns (3) to (5) also the number of public patents squared. In column (5) we additionally include an interaction effect between the number of public patents with the population in 2002 in million. In columns (2), (4) and (5) we address the potential endogeneity of public patent applications using an IV approach. The excluded instrument in column (2) is the 4-year lagged (ln) institutional funding for Fraunhofer and Leibniz institutes and in columns (4) and (5) the 4-year lagged (ln) institutional funding, linear and squared. In column (5) we additionally use 4-year lagged (ln) institutional funding interacted with population in 2002. Robust standard errors in parentheses. *,**,*** denote statistical significance at the 10, 5 and 1% level.
Source: Authors' calculation based on INKAR and OECD RegPAT database, 2003-2010.

in a region (see Table 4.10). In columns (1) to (3), we use the number of corporate patent applications filed by low R&D intensity firms and in columns (4) to (6) by high R&D intensity firms as dependent variable. Columns (1) and (4) show the baseline specification, in columns (2) and (5) we use only high quality corporate patents and in columns (3) and (6) only low quality corporate patents (based on their scope). The result suggests that it is in particular low R&D intensity firms that respond to local public R&D (column (1) vs. column (4)), the point estimates are larger in absolute terms and also more precisely estimated.[46] Moreover, the point estimates are similar for high quality (column (2)) and low quality patents (column (3)) of low R&D intensity firms. The absolute impact is, however, larger for low quality patents since low R&D intensity firms produce more low quality patents (see bottom of the table). Thus, our results are in line with the findings of the transmission channels.

Lastly, we investigate whether previously non-patenting firms start patenting, which could indicate university-spin offs, or whether previously patenting low R&D intensity firms patent more (see columns (5) and (6) of Table 4.B.1 in the Appendix). The result suggests that both low R&D intensity type firms respond to public patents but previously non-patenting firms are relatively less important as previously patenting firms have a 2.5 times higher average number of firm patent applications (column (6)).

Discussion While the IV results are qualitatively in line with the results of the transmission channel analysis, the effect size is substantially larger. This is consistent with a more severe measurement error (and/or a countercyclical reverse causality bias) and a less severe omitted variable bias, assuming that the latter would cause an upward bias. The IV estimates could, however, still be biased by reverse causality and omitted variables as the 4-year lag may be too short.

Based on the assumption of unbiased IV estimates, our results imply a semi-elasticity of 0.1 in the median county with non-zero public patents and thus that one additional public patent creates 3 additional corporate patents (based on the median number of firms in our sample).[47] Given that our transmission channel analysis does not suggest negative spillover effects into other regions by firms that benefit from public R&D within a region, 3.2 additional corporate patents within a region is a substantial effect. Whether this means that public R&D is an effective way to stimulate local public R&D depends, however, also on the cost of producing the (marginal) public patent. Since we only have institutional funding data for Fraunhofer and Leibniz institutes, we assume costs per public EPO patent of 4.5 million EURO based on the estimate by Dechezlepretre et al. (2016) for UK firms.[48] This implies that one additional corporate patent costs 1.4 million EURO using public knowledge production as firm R&D support strategy. To put these costs into perspective, we compare them to the public costs per firm patent application using another widely employed firm R&D support strategy, namely R&D tax credits. Two recent papers, one for the UK and for the US, estimate a ratio of value

[46] The fact that the point estimates have the opposite sign for high R&D intensity firms is unlikely to suggest that the additional patents of low R&D intensity firms come at the costs of high R&D intensity firms as the overall effect is positive.

[47] We focus on the median as the distribution of patent applications is skewed.

[48] This estimate seems reasonable compared to the average expenditures of research institutes per public patent in Germany (excluding own income and expenditures for research in humanities, see Table 4.1), which is 5.3 million EURO.

Model	IV					
Dep. Var.	ln(# Corporate Patents (CP))					
Firms	Low R&D			High R&D		
Patent Quality	All	High	Low	All	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
# Public patents	0.198**	0.235**	0.194*	-0.116	-0.109	-0.117
	(0.091)	(0.098)	(0.103)	(0.124)	(0.163)	(0.124)
# Public patents, sqrd.	-0.003**	-0.004**	-0.003*	0.002	0.002	0.002
	(0.001)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)
Observations	3,149	3,149	3,149	3,035	3,035	3,033
County-FE	x	x	x	x	x	x
Time-FE	x	x	x	x	x	x
Control Variables	x	x	x	x	x	x
P-value Underident.	0.028	0.028	0.028	0.028	0.028	0.028
F-Stat. Weak ident.	9	9	9	8	8	8
Dep. Var CP (PP $>$ 0)	35.670	8.743	26.915	40.684	10.763	29.929

Table 4.10: IV Results: Heterogeneity in the Impact of Public R&D

Notes: Table shows estimated coefficients of the impact of the number of public patent applications in a county on the number of corporate patent applications. All regressions include a full set of county as well as time-fixed effects and our set of control variables. Further, all columns show the results of the 2SLS estimation using 4-year lagged (ln) institutional funding for Fraunhofer and Leibniz institute, linear and squared, as excluded instruments. The dependent variable is in column (1) (ln) number of corporate patent applications by low R&D intensity firms, in column (2) (ln) number of high quality corporate patent applications by low R&D intensity firms, in column (3) (ln) number of low quality corporate patent applications by low R&D intensity firms. The dependent variable is in column (4) (ln) number of corporate patent applications by high R&D intensity firms, in column (5) (ln) number of high quality corporate patent applications by high R&D intensity firms and in column (6) (ln) number of low quality corporate patent applications by high R&D intensity firms. Low R&D intensity firms are firms that filed less than 5 patent applications between 1995 and 1999. Patent quality is based on patent scope (the number of IPC classes a patent is filed for). Robust standard errors in parentheses. *,**,*** denote statistical significance at the 10, 5 and 1% level.
Source: Authors' calculation based on INKAR and OECD RegPAT database, 2003-2010.

to money of 1.7 for R&D tax credits (see Dechezlepretre et al. 2016; Rao 2016). Thus, if we assume that a corporate patent by a German firm costs 4.6 million EURO, 2.7 million EURO in R&D tax credits need to be spent. It thus seems less expensive to stimulate firm R&D by creating knowledge outside the firm than by providing financial incentives to the firm.

An additional insight of the IV analysis is that the impact of public R&D on local firm R&D is decreasing in the level of public patents. Since our analysis shows that in particular low R&D intensity firms benefit from public R&D, the nonlinear relationship is unlikely to result from an uneven distribution of low R&D intensity firms. One explanation brought forward in the literature to explain a crowding out of private R&D by public R&D is the R&D employee wage channel: More public patents increase the costs for R&D employees in a region and thereby decrease R&D by firms. Since we do, however, not find evidence that high R&D intensity firms file less patent applications in response to public patents, this suggests that two R&D employee labor markets exist, one for low R&D intensity firms (and public institutions) and one for high R&D intensity firms. This would also explain why high R&D intensity firms do not benefit from local public R&D. The point estimates for our control variables (see Table 4.B.2) are in line with this presumption. While the share of skilled labor has a negative impact (also not statistically significant) on the number of corporate patents for low R&D intensity firms, the impact is positive (although again not statistically significant) for high R&D intensity firms.

4.6 Conclusion

In this paper, we aim to shed light on the overall magnitude of local knowledge spillovers of public R&D by universities and research institutes on firm R&D in Germany as well as its determinants. We proxy R&D activities by the number of patent applications observed in the OECD RegPat database. We find evidence that is consistent with the existence of three distinct transmission channels on the local level. First, firms that collaborate with a (local) public institution file more patent applications at the time of the collaboration. Second, firms that cite public patents also experience an increase in the number of patent applications around the time of citing the public patent. Third, firms that locate in a region in which public patents are produced file also more patent applications. While the results of our estimation strategy and the proxies employed suggest that the first and second channel have a substantial impact on the affected firm, their relative importance seems limited though. The number of regional citations of a public patent is very low and only a small fraction of all public patent applications result from a collaboration with a firm. The third channel is, however, suggested to be sizable and seems (compared to the other channels) to be in particular relevant for low R&D intensity firms. Given that the last transmission channel is, however, difficult to capture, our results should be interpreted with some caution.

Using an IV strategy that exploits variation in lagged institutional funding of research institutes, we find evidence for substantial local knowledge spillovers of public R&D. One additional public patent in the median county generates roughly 3 additional corporate patents if we assume our IV estimates to be unbiased. This suggests public costs of 1.4 million EURO per additional created corporate patent. Since we do not find evidence for less R&D in other regions by firms that benefit from local public R&D in one region,

our result indicates that local public R&D is a cost-efficient firm R&D support strategy. Moreover, our IV results are in line with the results of the transmission channel analysis as mainly low R&D intensity firms respond to local public R&D. While this seems conclusive, it should be stressed that our instrument, the 4-year lagged institutional funding, may not fully remove a potential omitted variable and reverse causality bias.

We believe that at least two conclusions can be drawn from our work. First, the German firm R&D support strategy, which relied dominantly on public R&D carried out by research institutes over our sample period, seems to be a successful one. If German firms are similar to, for example, US or UK firms, our result suggests that public R&D leads to more firm patent applications at the same public costs than using R&D tax credits to stimulate firm R&D. Second, public R&D seems to impact in particular R&D by local low R&D intensity (in other words small) firms. Since direct subsidies also impact in particular small firm R&D (Bronzini and Iachini 2014; González, Jaumandreu, and Pazó 2005; Howell 2017; Hyytinen and Toivanen 2005; Lach 2002), potentially due to alleviating financing constraints, our result suggests that non-specific knowledge spillovers may be very influential as they allow firms to focus on innovation strategies that are more likely to be successful.

Appendix to Chapter 4

4.A Additional Descriptive Statistics

Sample	JV		CITED		Public Patent	
	Mean	Median	Mean	Median	Mean	Median
# Corporate patents	0.51	0.00	0.51	0.00	0.29	0.00
# JV	0.02	0.00	0.01	0.00	0.00	0.00
# CITED	0.00	0.00	0.01	0.00	0.00	0.00
# Public patents	4.78	1.02	4.52	1.00	5.24	1.03
Population in 1,000	272	202	267	200	299	237
Share skilled employees	5.03	3.40	4.92	3.30	4.94	3.40
Share unskilled employees	8.49	8.10	8.46	8.10	8.61	8.30
Students per 1,000 capita	33.99	4.20	33.37	4.00	31.43	4.40
Unemployment rate	5.95	5.30	5.93	5.30	5.83	5.20
Property tax multiplier	357	344	358	343	359	344
Business tax multiplier	372	355	372	3543	3754	356
Manufacturing empl. in 1,000	53.87	40.34	53.24	40.17	61.29	45.45
Share of manufacturing empl.	0.42	0.39	0.42	0.40	0.44	0.41
Close to LUX	0.34	0.00	0.31	0.00	0.33	0.00
Close to BEL or NLD	0.36	0.00	0.34	0.00	0.38	0.00
Observations	73,395		89,251		564,698	

Table 4.A.1: Descriptive Statistics for Applicant-Region Level Estimation Samples

Notes: Table reports descriptive statistics for the three estimation samples for the transmission channels of knowledge spillovers of public R&D. JV stands for joint venture between firms and public institutions and CITED for a local corporate patent citing a public patent. The JV sample only includes applicants that collaborated with a public institutions between 1995 and 2015. The CITED sample only includes applicants that cited at least one public patent. The Public Patent sample includes all applicant-region observations. Applicant-region observations are only included if at least one corporate patent has been filed for this applicant-region combination between 1995 and 2015. The sample period is 2000 to 2010. Property tax and business tax multiplier are municipality-population weighted.
Source: Authors' calculation based on INKAR and OECD RegPAT and Citation database, 2000-2010.

4.B Additional Results

Model	IV					
Dep. Var.	ln(# Firm Patents)	(ln) R&D Employees	ln(# Corporate Patents)			
Patented before					Low R&D Yes	Low R&D No
	(1)	(2)	(3)	(4)	(5)	(6)
# Public patents	0.085	0.104	0.125**	0.093*	0.155*	0.218
	(0.056)	(0.078)	(0.061)	(0.052)	(0.085)	(0.155)
# Public patents, sqrd.	-0.002	-0.002	-0.002**	-0.001*	-0.003**	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	3,212	1,600	3,172	3,212	3,101	2,819
Control Variables	x	x	x	x	x	x
County-FE	x	x	x	x	x	x
State-Year-FE	x					
Time-FE	x	x	x	x	x	x
Outlier Excl.			x			
Agglomeration				x		
P-value Underident.	0.022	0.121	0.042	0.027	0.027	0.030
F-Stat. Weak ident.	8	11	12	8	9	9
Mean Dep. Var. (PP $>$ 0)	81.158		70.278	81.158	25.714	10.945

Table 4.B.1: Robustness IV Results

Notes: Table shows estimated coefficients of the sensitivity analysis for the impact of the number of public patent applications in a county on the number of firm patent applications in the same county. All regressions include a full set of county as well as time-fixed effects and our set of control variables. Further, all columns show the results of the 2SLS estimation using 4-year lagged (ln) institutional funding for Fraunhofer and Leibniz institutes, linear and squared, as excluded instruments. In column (1) we additionally include state-year fixed effects. In column (2) we use (ln) number of firm R&D employees as dependent variable. In column (3) we address a potential overdispersion of the data by excluding regions with a number of firm patents in the top 1% of the distribution. In column (4) we additionally control for the number of corporate patents in a region between 1991 and 2000 interacted with year dummies to account for the impact of agglomeration economies. In column (5) we only use the number of patents by firms that filed between 1 and 5 patent applications between 1995 and 1999 as dependent variable and in column (6) we only use the number of patents by firms that did not file patent applications between 1995 and 1999 as dependent variable. Robust standard errors in parentheses. *,**,*** denote statistical significance at the 10, 5 and 1% level.
Source: Authors' calculation based on INKAR and OECD RegPAT database, 2003-2010.

Model	IV					
Dep. Var.	# Corporate Patents					
Firms	Low R&D Intensity			High R&D Intensity		
Patents Quality	All	High	Low	All	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
Close to Lux * D(>2006)	-0.001	-0.020	0.002	0.095**	0.104**	0.091*
	(0.041)	(0.049)	(0.042)	(0.046)	(0.049)	(0.048)
Close to NLD, BEL* D(>2007)	-0.042	-0.014	-0.047	-0.055	-0.085*	-0.036
	(0.041)	(0.049)	(0.042)	(0.045)	(0.049)	(0.047)
Population	0.001	0.002	0.000	-0.006	-0.007	-0.006
	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)	(0.007)
Property tax multiplier	-0.001*	-0.002	-0.002*	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Business tax multiplier	0.004**	0.005**	0.004**	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Students per 1.000 capita	0.003	0.004	0.003	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Share unskilled employees	-0.057*	-0.061*	-0.058*	-0.006	-0.021	0.003
	(0.032)	(0.034)	(0.033)	(0.029)	(0.034)	(0.029)
Share skilled employees	-0.059	-0.046	-0.070	0.072	0.066	0.071
	(0.075)	(0.084)	(0.081)	(0.080)	(0.097)	(0.080)
ln Manufacturing empl.	1.292***	1.344**	1.210**	0.327	0.942*	0.011
	(0.474)	(0.574)	(0.492)	(0.488)	(0.541)	(0.500)
Share manufacturing empl.	-2.051	-1.527	-2.114	0.807	-0.530	1.440
	(1.365)	(1.603)	(1.421)	(1.386)	(1.570)	(1.412)
Unemployment rate	-0.020	-0.011	-0.024	-0.026	-0.022	-0.026
	(0.018)	(0.021)	(0.019)	(0.022)	(0.024)	(0.023)
Observations	3,149	3,149	3,149	3,035	3,035	3,033
County FE	x	x	x	x	x	x

Table 4.B.2: Results for Control Variables in Table 4.10

Notes: Table shows estimated coefficients for the control variables of the specification reported in Table 4.10. Robust standard errors in parentheses. *,**,*** denote statistical significance at the 10, 5 and 1% level.
Source: Authors' calculation based on INKAR and OECD RegPAT database, 2003-2010.

Bibliography

Abel, Jaison R and Richard Deitz (2011). "Do colleges and universities increase their region's human capital?" *Journal of Economic Geography*, 12 (3), 667–691.

Abeler, Johannes and Simon Jäger (2015). "Complex tax incentives." *American Economic Journal: Economic Policy*, 7 (3), 1–28.

Abramovsky, Laura, Rupert Harrison, and Helen Simpson (2007). "University Research and the Location of Business R&D." *The Economic Journal*, 117 (519), 114–141.

Acs, Zoltan J, David B Audretsch, and Maryann P Feldman (1994). "R&D spillovers and recipient firm size." *The Review of Economics and Statistics*, 336–340.

Adams, Paul, Darragh Kelly, Michael Grubb, and Matthew Osborne (2020). "Time to Act: A Field Experiment on Overdraft Alerts," FCA Occasional Paper (40).

Alcacer, Juan and Michelle Gittelman (2006). "Patent citations as a measure of knowledge flows: The influence of examiner citations." *The Review of Economics and Statistics*, 88 (4), 774–779.

Alesina, Alberto, Andrea Ichino, and Loukas Karabarbounis (2011). "Gender-based taxation and the division of family chores." *American Economic Journal: Economic Policy*, 3 (2), 1–40.

Almeida, Paul and Bruce Kogut (1999). "Localization of knowledge and the mobility of engineers in regional networks." *Management Science*, 45 (7), 905–917.

Alstadsaeter, Annette, Salvador Barrios, Gaetan Nicodeme, Agnieszka Maria Skonieczna, and Antonio Vezzani (2018). "Patent boxes design, patents location, and local R&D." *Economic Policy*, 33 (93), 131–177.

Altmann, Steffen, Armin Falk, Paul Heidhues, Rajshri Jayaraman, and Marrit Teirlinck (2019). "Defaults and donations: Evidence from a field experiment." *Review of Economics and Statistics*, 101 (5), 808–826.

Andersson, Roland, John M Quigley, and Mats Wilhelmsson (2009). "Urbanization, productivity, and innovation: Evidence from investment in higher education." *Journal of Urban Economics*, 66 (1), 2–15.

Andresen, Martin Eckhoff and Emily Nix (2022). "What causes the child penalty? Evidence from adopting and same-sex couples." *Journal of Labor Economics*, 40 (4), 971–1004.

Ankrah, Samuel and Omar AL-Tabbaa (2015). "Universities–industry collaboration: A systematic review." *Scandinavian Journal of Management*, 31 (3), 387–408.

Antecol, Heather and Michael D Steinberger (2013). "Labor supply differences between married heterosexual women and partnered lesbians: A semi-parametric decomposition approach." *Economic Inquiry*, 51 (1), 783–805.

Arrow, Kenneth J (1962). "The economic implications of learning by doing." *The Review of Economic Studies*, 29 (3), 155–173.

Asai, Yukiko (2015). "Parental leave reforms and the employment of new mothers: Quasi-experimental evidence from Japan." *Labour Economics*, 36, 72–83.

Attanasio, Orazio P and Valérie Lechene (2014). "Efficient responses to targeted cash transfers." *Journal of Political Economy*, 122 (1), 178–222.

Audretsch, David B and Maryann P Feldman (1996). "R&D spillovers and the geography of innovation and production." *The American Economic Review*, 86 (3), 630–640.

Auten, Gerald and Robert Carroll (1999). "The effect of income taxes on household income." *Review of Economics and Statistics*, 81 (4), 681–693.

Badgett, MV Lee, Christopher S Carpenter, and Dario Sansone (2021). "LGBTQ Economics." *Journal of Economic Perspectives*, 35 (2), 141–170.

Baker, Andrew C, David F Larcker, and Charles CY Wang (2022). "How much should we trust staggered difference-in-differences estimates?" *Journal of Financial Economics*, 144 (2), 370–395.

Bargain, Olivier, Kristian Orsini, and Andreas Peichl (2014). "Comparing labor supply elasticities in Europe and the United States." *Journal of Human Resources*, 49 (3), 723–838.

Barschkett, Mara (2022). *Age-specific Effects of early daycare on children's health.* DIW Working Paper 2028.

Bauer, Gerrit (2016). "Gender roles, comparative advantages and the life course: the division of domestic labor in same-sex and different-sex couples." *European Journal of Population*, 32, 99–128.

Bayer, Tim, Lenard Simon, and Jakob Wegmann (2023). *Withheld from Working More? Withholding Taxes and the Labor Supply of Married Women.* Accessed: 2024-03-12. https://conference.iza.org/conference_files/SummerSchool_2023/wegmann_j28978.pdf.

Beblo, Miriam and Denis Beninger (2017). "Do husbands and wives pool their incomes? A couple experiment." *Review of Economics of the Household*, 15, 779–805.

Becker, Gary S (1981). *A Treatise on the Family.* Harvard University Press.

Becker, Johannes, Jonas Fooken, and Melanie Steinhoff (2019). "Behavioral effects of withholding taxes on labor supply." *The Scandinavian Journal of Economics*, 121 (4), 1417–1440.

Belenzon, Sharon and Mark Schankerman (2013). "Spreading the word: Geography, policy, and knowledge spillovers." *Review of Economics and Statistics*, 95 (3), 884–903.

Bercovitz, Janet and Maryann Feldman (2006). "Entrepreneurial universities and technology transfer: A conceptual framework for understanding knowledge-based economic development." *The Journal of Technology Transfer*, 31 (1), 175–188.

Bergemann, Annette and Regina T. Riphahn (2022). "Maternal employment effects of paid parental leave." *Journal of Population Economics*, 36 (1), 139–178.

Bertrand, Marianne (2020). "Gender in the twenty-first century." *AEA Papers and Proceedings*. Vol. 110, 1–24.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan (2015). "Gender identity and relative income within households." *The Quarterly Journal of Economics*, 130 (2), 571–614.

Bick, Alexander and Nicola Fuchs-Schündeln (2017). "Quantifying the disincentive effects of joint taxation on married women's labor supply." *American Economic Review*, 107 (5), 100–104.

— (2018). "Taxation and labour supply of married couples across countries: A macroeconomic analysis." *The Review of Economic Studies*, 85 (3), 1543–1576.

Black, Dan A, Seth G Sanders, and Lowell J Taylor (2007). "The economics of lesbian and gay families." *Journal of Economic Perspectives*, 21 (2), 53–70.

Blau, Francine D and Lawrence M Kahn (2003). "Understanding international differences in the gender pay gap." *Journal of Labor Economics*, 21 (1), 106–144.

— (2017). "The gender wage gap: Extent, trends, and explanations." *Journal of Economic Literature*, 55 (3), 789–865.

Blaufus, Kay, Jonathan Bob, Jochen Hundsdoerfer, Christian Sielaff, Dirk Kiesewetter, and Joachim Weimann (2015). "Perception of income tax rates: evidence from Germany." *European Journal of Law and Economics*, 40 (3), 457–478.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen (2013). "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica*, 81 (4), 1347–1393.

Blundell, Richard, Pierre-Andre Chiappori, Thierry Magnac, and Costas Meghir (2007). "Collective labour supply: Heterogeneity and non-participation." *The Review of Economic Studies*, 74 (2), 417–445.

Blundell, Richard and Thomas MaCurdy (1999). "Labor supply: A review of alternative approaches." *Handbook of Labor Economics*, 3, 1559–1695.

BMFSFJ (2008). *Elterngeld als Teil nachhaltiger Familienpolitik*. Tech. rep. URL: https://www.bmfsfj.de/bmfsfj/service/publikationen/elterngeld-als-teil-nachhaltiger-familienpolitik/73822.

Borella, Margherita, Mariacristina De Nardi, and Fang Yang (2023). "Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labour Supply?" *The Review of Economic Studies*, 90 (1), 102–131.

Bound, John, Clint Cummins, Zvi Griliches, Bronwyn H Hall, and Adam B Jaffe (1984). "Who does R&D and who patents?" *R&D, patents and productivity*. Ed. by Zvi Griliches. University of Chicago Press, 21–54.

Branstetter, Lee (2005). "Exploring the link between academic science and industrial innovation." *Annals of Economics and Statistics*, 79 (80), 119–142.

Branstetter, Lee G. and Mariko Sakakibara (2002). "When Do Research Consortia Work Well and Why? Evidence from Japanese Panel Data." *The American Economic Review*, 92 (1), 143–159.

Bronzini, Raffaello and Eleonora Iachini (2014). "Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach." *American Economic Journal: Economic Policy*, 6 (4), 100–134.

Buettner, Thiess, Katharina Erbe, and Veronika Grimm (2019). "Tax planning of married couples and intra-household income inequality." *Journal of Public Economics*, 179, 104048.

Bundestag, Deutscher (1974). "Gesetzentwurf der Bundesregierung: Entwurf eines Dritten Steuerreformgesetzes vom 09.01.1974." *BT-Drs. 7/1470*.

— (2019). "Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Doris Achelwilm, Cornelia Möhring, Dr. Petra Sitte, weiterer Abgeordneter und der Fraktion – DIE LINKE – Drs. 19/12373 vom 30.08.2019." *BT-Drs. 19/12857*.

Callaway, Brantly and Pedro HC Sant'Anna (2021). "Difference-in-differences with multiple time periods." *Journal of Econometrics*, 225 (2), 200–230.

Canaan, Serena (2022). "Parental leave, household specialization and children's well-being." *Labour Economics*, 75, 102–127.

Carpenter, Christopher S (2007). "Revisiting the income penalty for behaviorally gay men: Evidence from NHANES III." *Labour Economics*, 14 (1), 25–34.

Cassiman, B., R. Veugelers, and S. Arts (2012). "How to capture value from linking to science-driven basic research: Boundary crossing inventors and partnerships." *2012 IEEE International Conference on Management of Innovation Technology (ICMIT)*, 619–623.

Cassiman, Bruno, Reinhilde Veugelers, and Pluvia Zuniga (2008). "In search of performance effects of (in) direct industry science links." *Industrial and Corporate Change*, 17 (4), 611–646.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019). "The effect of minimum wages on low-wage jobs." *The Quarterly Journal of Economics*, 134 (3), 1405–1454.

Chetty, Raj (2012). "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply." *Econometrica*, 80 (3), 969–1018.

Chetty, Raj, John N Friedman, and Emmanuel Saez (2013). "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review*, 103 (7), 2683–2721.

Chetty, Raj, Adam Looney, and Kory Kroft (2009). "Salience and taxation: Theory and evidence." *American Economic Review*, 99 (4), 1145–1177.

Chetty, Raj and Emmanuel Saez (2013). "Teaching the tax code: Earnings responses to an experiment with EITC recipients." *American Economic Journal: Applied Economics*, 5 (1), 1–31.

Cygan-Rehm, Kamila (2016). "Parental leave benefit and differential fertility responses: Evidence from a German reform." *Journal of Population Economics*, 29 (1), 73–103.

Dahl, Gordon B., Katrine V. Løken, Magne Mogstad, and Kari Vea Salvanes (2016). "What Is the Case for Paid Maternity Leave?" *Review of Economics and Statistics*, 98 (4), 655–670.

David, Paul A., Bronwyn H. Hall, and Andrew A. Toole (2000). "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence." *Research Policy*, 29 (4), 497–529.

De Chaisemartin, Clément and Xavier d'Haultfoeuille (2020). "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110 (9), 2964–2996.

Dechezlepretre, Antoine, Elias Einiö, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen (2016). *Do tax Incentives for Research Increase Firm Innovation? An RD Design for R&D*. NBER Working Paper 22405. National Bureau of Economic Research.

Deshpande, Manasi and Yue Li (2019). "Who is screened out? Application costs and the targeting of disability programs." *American Economic Journal: Economic Policy*, 11 (4), 213–248.

Destatis (2014). *Öffentliche Sozialleistungen. Statistik zum Elterngeld. Leistungsbezüge*. Tech. rep. Accessed:2024-03-12. https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00096487: Statistisches Bundesamt (Destatis).

— (2019). *Statistiken der Kinder- und Jugendhilfe: Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2018*. Tech. rep. Accessed: 2024-03-12. https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00123476: Statistisches Bundesamt (Destatis).

Dilmaghani, Maryam and Jason Dean (2023). "Sexual orientation and intra-household specialization before and after the legal recognition of same-sex marriage in Canada." *Applied Economics*, 1–25.

Doerrenberg, Philipp, Andreas Peichl, and Sebastian Siegloch (2017). "The elasticity of taxable income in the presence of deduction possibilities." *Journal of Public Economics*, 151, 41–55.

DOLLS, MATHIAS, PHILIPP DOERRENBERG, ANDREAS PEICHL, and HOLGER STICHNOTH (2018). "Do retirement savings increase in response to information about retirement and expected pensions?" *Journal of Public Economics*, 158, 168–179.

DRYDAKIS, NICK (2022). "Sexual orientation and earnings: a meta-analysis 2012–2020." *Journal of Population Economics*, 35 (2), 409–440.

FELDMAN, MARYANN P. and DIETER F. KOGLER (2010). "Chapter 8 - Stylized Facts in the Geography of Innovation." *Handbook of The Economics of Innovation, Vol. 1.* Ed. by BRONWYN H. HALL and NATHAN ROSENBERG. Vol. 1. Handbook of the Economics of Innovation. North-Holland, 381–410.

FELDMAN, NAOMI E, PETER KATUŠČÁK, and LAURA KAWANO (2016). "Taxpayer confusion: Evidence from the child tax credit." *American Economic Review*, 106 (3), 807–35.

FINKELSTEIN, AMY (2009). "E-ztax: Tax salience and tax rates." *The Quarterly Journal of Economics*, 124 (3), 969–1010.

FORAY, DOMINIQUE and FRANCESCO LISSONI (2010). "Chapter 6 - University Research and Public–Private Interaction." *Handbook of The Economics of Innovation, Vol. 1.* Ed. by BRONWYN H. HALL and NATHAN ROSENBERG. Vol. 1. Handbook of the Economics of Innovation. North-Holland, 275–314.

GABAIX, XAVIER (2019). "Behavioral inattention." *Handbook of Behavioral Economics: Applications and foundations 1.* Vol. 2. Elsevier, 261–343.

GEYER, JOHANNES, PETER HAAN, and KATHARINA WROHLICH (2015). "The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach." *Labour Economics*, 36, 84–98.

GIDDINGS, LISA, JOHN M NUNLEY, ALYSSA SCHNEEBAUM, and JOACHIM ZIETZ (2014). "Birth cohort and the specialization gap between same-sex and different-sex couples." *Demography*, 51 (2), 509–534.

GIOMMONI, TOMMASO and ENRICO RUBOLINO (2022). *The Cost of Gender Identity Norms: Evidence from a Spouse Tax Credit.* Dondena Working Papers 153.

GIRSBERGER, ESTHER MIRJAM, LENA HASSANI-NEZHAD, KALAIVANI KARUNANETHY, and RAFAEL LALIVE (2023). "Mothers at work: How mandating a short maternity leave affects work and fertility." *Labour Economics*, 84, 102364.

GLOGOWSKY, ULRICH, AMELIE GROSENICK, EMANUEL HANSE, LEONIE KOCH, ANDREAS PEICHL, and DOMINIK SACHS (2024). *The Impact of Parental Leave Benefits on Pre-Birth Earnings.* Unpublished manuscript.

GOLDBERG, ABBIE E (2013). ""Doing" and "undoing" gender: The meaning and division of housework in same-sex couples." *Journal of Family Theory & Review*, 5 (2), 85–104.

GOLDIN, CLAUDIA (2014). "A grand gender convergence: Its last chapter." *American Economic Review*, 104 (4), 1091–1119.

— (2021). *Career and Family: Women's Century-Long Journey toward Equity*. Princeton University Press.

González, Xulia, Jordi Jaumandreu, and Consuelo Pazó (2005). "Barriers to innovation and subsidy effectiveness." *The RAND Journal of Economics*, 36 (4), 930–950.

Goodman-Bacon, Andrew (2021). "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*, 225 (2), 254–277.

Griffith, Rachel, Sokbae Lee, and John Van Reenen (2011). "Is distance dying at last? Falling home bias in fixed-effects models of patent citations." *Quantitative Economics*, 2 (2), 211–249.

Griliches, Zvi (1998). "Patent Statistics as Economic Indicators: A Survey." *R&D and Productivity: The Econometric Evidence*. National Bureau of Economic Research, Inc, 287–343.

Gruber, Jon and Emmanuel Saez (2002). "The elasticity of taxable income: evidence and implications." *Journal of Public Economics*, 84 (1), 1–32.

Guceri, Irem (2018). "Will the real R&D employees please stand up? Effects of tax breaks on firm-level outcomes." *International Tax and Public Finance*, 25 (1), 1–63.

Guceri, Irem and Li Liu (2019). "Effectiveness of fiscal incentives for R&D: Quasi-experimental evidence." *American Economic Journal: Economic Policy*, 11 (1), 266–91.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura (2012). "Taxation and household labour supply." *The Review of Economic Studies*, 79 (3), 1113–1149.

Hall, Bronwyn and John Van Reenen (2000). "How effective are fiscal incentives for R&D? A review of the evidence." *Research Policy*, 29 (4-5), 449–469.

Hansen, Mary Eschelbach, Michael E Martell, and Leanne Roncolato (2020). "A labor of love: The impact of same-sex marriage on labor supply." *Review of Economics of the Household*, 18, 265–283.

Harhoff, Dietmar, Frederic M Scherer, and Katrin Vopel (2003). "Citations, family size, opposition and the value of patent rights." *Research Policy*, 32 (8), 1343–1363.

Hermle, Johannes, Elena Herold, and Nikolaus Hildebrand (2024). *Preferences over Relative Income within the Household*. IZA Working Paper DP No. 16803.

Hermle, Johannes and Andreas Peichl (2018). *Jointly optimal taxes for different types of income*. CESifo Working Paper No. 7248.

Herold, Elena, Leonie Koch, and Carina Neisser (2024). *Disentangling Gender Norms and Tax Incentives - Analyzing the Introduction of Joint Taxation to Same-Sex Couples*. Unpublished manuscript.

Herold, Elena and Luisa Wallossek (2023). *The Marriage Earnings Gap*. Accessed: 2024-03-12. https://www.dropbox.com/scl/fi/4ffpmqtalsmu37exdwf5k/

Herold_Wallossek_Marriage_Earnings_Gap.pdf?rlkey=9k3uqnr6djy19bg0me30xl70e&e=1&dl=0.

HOFMARCHER, THOMAS and ERIK PLUG (2022). "Specialization in same-sex and different-sex couples." *Labour Economics*, 77, 101995.

HOWELL, SABRINA T. (2017). "Financing Innovation: Evidence from R&D Grants." *The American Economic Review*, 107 (4), 1136–1164.

HYYTINEN, ARI and OTTO TOIVANEN (2005). "Do financial constraints hold back innovation and growth? Evidence on the role of public policy." *Research Policy*, 34 (9), 1385–1403.

ISAAC, ELLIOTT (2023). "Suddenly married: Joint taxation and the labor supply of same-sex married couples after US v. Windsor." *Journal of Human Resources*.

JAFFE, ADAM B (1989). "Real effects of academic research." *The American Economic Review*, 79 (5), 957–970.

JAFFE, ADAM B, MANUEL TRAJTENBERG, and REBECCA HENDERSON (1993). "Geographic localization of knowledge spillovers as evidenced by patent citations." *The Quarterly Journal of Economics*, 108 (3), 577–598.

JAKOBSEN, KATRINE MARIE and JAKOB EGHOLT SØGAARD (2022). "Identifying behavioral responses to tax reforms: New insights and a new approach." *Journal of Public Economics*, 212, 104691.

JEPSEN, CHRISTOPHER and LISA JEPSEN (2022). "Convergence over time or not? US wages by sexual orientation, 2000–2019." *Labour Economics*, 74, 102086.

JOHNSON, ERIC J and DANIEL GOLDSTEIN (2003). "Do defaults save lives?" *Science*, 302 (5649), 1338–1339.

JONES, DAMON (2012). "Inertia and overwithholding: explaining the prevalence of income tax refunds." *American Economic Journal: Economic Policy*, 4 (1), 158–185.

KALÍŠKOVÁ, KLÁRA (2014). "Labor supply consequences of family taxation: Evidence from the Czech Republic." *Labour Economics*, 30, 234–244.

KANTOR, SHAWN and ALEXANDER WHALLEY (2014). "Knowledge spillovers from research universities: Evidence from endowment value shocks." *Review of Economics and Statistics*, 96 (1), 171–188.

KEANE, MICHAEL P (2011). "Labor supply and taxes: A survey." *Journal of Economic Literature*, 49 (4), 961–1075.

KIRCHHOF, PAUL (2011). *Bundessteuergesetzbuch: Ein Reformentwurf zur Erneuerung des Steuerrechts*. Vol. 4. CF Müller GmbH.

KLEVEN, HENRIK, CAMILLE LANDAIS, and JAKOB EGHOLT SØGAARD (2019). "Children and gender inequality: Evidence from Denmark." *American Economic Journal: Applied Economics*, 11 (4), 181–209.

Kleven, Henrik Jacobsen and Esben Anton Schultz (2014). "Estimating taxable income responses using Danish tax reforms." *American Economic Journal: Economic Policy*, 6 (4), 271–301.

Kluve, Jochen and Sebastian Schmitz (2018). "Back to work: Parental benefits and mothers' labor market outcomes in the medium run." *ILR Review*, 71 (1), 143–173.

Kluve, Jochen and Marcus Tamm (2013). "Parental leave regulations, mothers' labor force attachment and fathers' childcare involvement: Evidence from a natural experiment." *Journal of Population Economics*, 26 (3), 983–1005.

Koch, Leonie (2024). *Income Taxation of Couples and Gender (In)equality in Earnings*. Unpublished manuscript.

Koch, Leonie and Martin Simmler (2020). "How important are local knowledge spillovers of public R&D and what drives them?" *Research Policy*, 49 (7), 104009.

Kriete-Dodds, Susan and Daniel Vorgrimler (2007). "The German Taxpayer-Panel." *Schmollers Jahrbuch*, 127 (3), 497–509.

Küpper, Beate, Ulrich Klocke, and Lena-Carlotta Hoffmann (2017). *Einstellungen gegenüber lesbischen, schwulen und bisexuellen Menschen in Deutschland: Ergebnisse einer bevölkerungsrepräsentativen Umfrage*. Nomos.

Lach, Saul (2002). "Do R&D subsidies stimulate or displace private R&D? Evidence from Israel." *The Journal of Industrial Economics*, 50 (4), 369–390.

Lalive, R., A. Schlosser, A. Steinhauer, and J. Zweimuller (2013). "Parental Leave and Mothers' Careers: The Relative Importance of Job Protection and Cash Benefits." *The Review of Economic Studies*, 81 (1), 219–265.

Lalive, Rafael and Josef Zweimüller (2009). "How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments." *The Quarterly Journal of Economics*, 124 (3), 1363–1402.

LaLumia, Sara (2008). "The effects of joint taxation of married couples on labor supply and non-wage income." *Journal of Public Economics*, 92 (7), 1698–1719.

Lanjouw, Jean O. and Mark Schankerman (2001). "Characteristics of Patent Litigation: A Window on Competition." *The RAND Journal of Economics*, 32 (1), 129–151.

Lau, Charles Q (2012). "The stability of same-sex cohabitation, different-sex cohabitation, and marriage." *Journal of Marriage and Family*, 74 (5), 973–988.

Leppel, Karen (2009). "Labour force status and sexual orientation." *Economica*, 76 (301), 197–207.

Lerner, Joshua (1994). "The importance of patent scope: An empirical analysis." *The RAND Journal of Economics*, 25 (2), 319–333.

Liebman, Jeffrey B and Richard Zeckhauser (2004). *Schmeduling*. Accessed: 2024-03-12. http://www.econ.yale.edu/~shiller/behmacro/2004-11/schmeduling-zeckhauser.pdf.

LINK, ALBERT N and JOHN REES (1990). "Firm size, university based research, and the returns to R&D." *Small Business Economics*, 2 (1), 25–31.

LOKSHIN, BORIS and PIERRE MOHNEN (2012). "How effective are level-based R&D tax credits? Evidence from the Netherlands." *Applied Economics*, 44 (12), 1527–1538.

LUNDBERG, SHELLY J, ROBERT A POLLAK, and TERENCE J WALES (1997). "Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit." *Journal of Human Resources*, 463–480.

LYCHAGIN, SERGEY, JORIS PINKSE, MARGARET E SLADE, and JOHN VAN REENEN (2016). "Spillovers in space: Does geography matter?" *The Journal of Industrial Economics*, 64 (2), 295–335.

MADRIAN, BRIGITTE C and DENNIS F SHEA (2001). "The power of suggestion: Inertia in 401 (k) participation and savings behavior." *The Quarterly Journal of Economics*, 116 (4), 1149–1187.

MANSER, MARILYN and MURRAY BROWN (1980). "Marriage and household decision-making: A bargaining analysis." *International Economic Review*, 31–44.

MARTINEZ, ISABEL Z, EMMANUEL SAEZ, and MICHAEL SIEGENTHALER (2021). "Intertemporal labor supply substitution? evidence from the swiss income tax holidays." *American Economic Review*, 111 (2), 506–546.

MAURSETH, PER BOTOLF and BART VERSPAGEN (2002). "Knowledge Spillovers in Europe: A Patent Citations Analysis." *Scandinavian Journal of Economics*, 104 (4), 531–545.

MICHEL, JACQUES and BERND BETTELS (2001). "Patent citation analysis: A closer look at the basic input data from patent search reports." *Scientometrics*, 51 (1), 185–201.

MULKAY, BENOIT and JACQUES MAIRESSE (2013). "The R&D tax credit in France: Assessment and ex ante evaluation of the 2008 reform." *Oxford Economic Papers*, 65 (3), 746–766.

NARIN, FRANCIS, KIMBERLY S HAMILTON, and DOMINIC OLIVASTRO (1997). "The increasing linkage between US technology and public science." *Research Policy*, 26 (3), 317–330.

NEISSER, CARINA (2021). "The elasticity of taxable income: A meta-regression analysis." *The Economic Journal*, 131 (640), 3365–3391.

NELSON, RICHARD R (1959). "The simple economics of basic scientific research." *Journal of Political Economy*, 67 (3), 297–306.

OECD (2015). "Measuring the technological and economic value of patents." *Chapter 2 in Enquiries into intellectual property's economic impact*.

— (2016). *Taxing Wages 2016*. Tech. rep. Accessed: 2024-03-12. https://www.oecd-ilibrary.org/taxation/taxing-wages-2016_tax_wages-2016-en: OECD.

— (2022). *OECD Family Database*. Tech. rep. Acessed: 2024-03-12. https://www.oecd.org/els/family/database.htm: OECD.

Olivetti, Claudia and Barbara Petrongolo (2016). "The evolution of gender gaps in industrialized countries." *Annual Review of Economics*, 8, 405–434.

— (2017). "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries." *Journal of Economic Perspectives*, 31 (1), 205–230.

Oreffice, Sonia (2011). "Sexual orientation and household decision making: Same-sex couples' balance of power and labor supply choices." *Labour Economics*, 18 (2), 145–158.

Oreffice, Sonia and Dario Sansone (2023). "Commuting to work and gender norms by sexual orientation." *Labour Economics*, 85, 102451.

Pesaran, M. Hashem and Larry W. Taylor (1999). "Diagnostics for IV Regressions." *Oxford Bulletin of Economics and Statistics*, 61 (2), 255–281.

Petrongolo, Barbara and Maddalena Ronchi (2020). "Gender gaps and the structure of local labor markets." *Labour Economics*, 64, 101819.

Rao, Nirupama (2016). "Do tax credits stimulate R&D spending? The effect of the R&D tax credit in its first decade." *Journal of Public Economics*, 140, 1–12.

Raute, Anna (2019). "Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits." *Journal of Public Economics*, 169, 203–222.

Rees-Jones, Alex and Dmitry Taubinsky (2020). "Measuring "schmeduling"." *The Review of Economic Studies*, 87 (5), 2399–2438.

Roantree, Barra (2023). *Income tax and the careers of women*. Accessed: 2024-03-12. https://github.com/barraroantree/papers/blob/gh-pages/incometaxes.pdf.

Romer, Paul M (1990). "Endogenous technological change." *Journal of Political Economy*, 98 (5, Part 2), 71–102.

Rybnicek, Robert and Roland Königsgruber (2019). "What makes industry–university collaboration succeed? A systematic review of the literature." *Journal of Business Economics*, 89 (2), 221–250.

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz (2012). "The elasticity of taxable income with respect to marginal tax rates: A critical review." *Journal of Economic Literature*, 50 (1), 3–50.

Sanders, Anne EH (2016). "When, if not now? An update on civil partnership in Germany." *German Law Journal*, 17 (3), 487–508.

Scandura, Alessandra (2016). "University-industry collaboration and firms' R&D effort." *Research Policy*, 45 (9), 1907–1922.

Scherer, Frederic M (1983). "The propensity to patent." *International Journal of Industrial Organization*, 1 (1), 107–128.

Schönberg, Uta and Johannes Ludsteck (2014). "Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth." *Journal of Labor Economics*, 32 (3), 469–505.

SELIN, HÅKAN (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), 894–922.

SPIESS (2011). "Vereinbarkeit von Familie und Beruf – wie wirksam sind deutsche "Care Policies"?" *Perspektiven der Wirtschaftspolitik*, 12, 4–27.

SUN, LIYANG and SARAH ABRAHAM (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225 (2), 175–199.

TEBALDI, EDINALDO and BRUCE ELMSLIE (2006). "Sexual orientation and labour supply." *Applied Economics*, 38 (5), 549–562.

TRAJTENBERG, MANUEL, REBECCA HENDERSON, and ADAM JAFFE (1997). "University Versus Corporate Patents: A Window On The Basicness Of Invention." *Economics of Innovation and New Technology*, 5 (1), 19–50.

VALFORT, MARIE-ANNE (2017). *LGBTI in OECD countries: A review*. OECD Social, Employment and Migration Working Papers, No. 198. OECD.

VAN DER VLEUTEN, MAAIKE, EVA JASPERS, and TANJA van der LIPPE (2021). "Same-sex couples' division of labor from a cross-national perspective." *Journal of GLBT Family Studies*, 17 (2), 150–167.

VAN LOOY, BART, JULIE CALLAERT, and KOENRAAD DEBACKERE (2006). "Publication and patent behavior of academic researchers: Conflicting, reinforcing or merely co-existing?" *Research Policy*, 35 (4), 596–608.

WALSH, JOHN P., YOU-NA LEE, and SADAO NAGAOKA (2016). "Openness and innovation in the US: Collaboration form, idea generation and implementation." *Research Policy*, 45 (8), 1660–1671.

WARTBURG, IWAN von, THORSTEN TEICHERT, and KATJA ROST (2005). "Inventive progress measured by multi-stage patent citation analysis." *Research Policy*, 34 (10), 1591–1607.

WEBB, COLIN, HÉLÈNE DERNIS, DIETMAR HARHOFF, and KARIN HOISL (2005). *Analysing European and International Patent Citations*. Accessed: 2024-03-12. https://www.oecd-ilibrary.org/content/paper/883002633010.

WINGENDER, PHILIPPE and SARA LALUMIA (2017). "Income Effects on Maternal Labor Supply: Evidence from Child-Related Tax Benefits." *National Tax Journal*, 70 (1), 11–52.

WORLD BANK (2024). *Women, Business and the Law 2024*. Tech. rep. Accessed: 2024-03-12. https://wbl.worldbank.org/en/reports: World Bank.

ZÚÑIGA-VICENTE, JOSÉ, CÉSAR ALONSO-BORREGO, FRANCISCO FORCADELL, and JOSÉ GALÁN (2014). "Assessing the effect of public subsidies on firm R&D investment: A survey." *Journal of Economic Surveys*, 28 (1), 36–67.

Zurla, Valeria (2022). *How Should We Design Parental Leave Policies? Evidence from Two Reforms in Italy*. Accessed:2024-03-12. https://vzurla.github.io/papers/jmp_draft_nov22.pdf.

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, 14. März 2024

Leonie Koch