Long–Run Determinants of Economic Performance and Institutional Quality: Evidence from Demographic and Political Change

Rainer Franz Kotschy

Long–Run Determinants of Economic Performance and Institutional Quality: Evidence from Demographic and Political Change

Inaugural-Dissertation

zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

an der Ludwig-Maximilians-Universität München

2018

vorgelegt von

Rainer Franz Kotschy

Referent: Prof. Dr. Uwe Sunde

Korreferent: Prof. Dr. Oliver Falck

Promotionsabschlussberatung: 11. Juli 2018

Tag der mündlichen Prüfung: 6. Juli 2018

Namen der Berichterstatter: Uwe Sunde, Oliver Falck, Joachim Winter

Acknowledgements

First and foremost, I would like to thank my co-author and first supervisor Uwe Sunde for his guidance and continuous encouragement over the last years. I am very grateful for his insightful comments, the productive and inspiring cooperation on joint projects, and his support by encouraging me to participate in international conferences. I am also grateful to my second supervisor Oliver Falck, who through numerous conversations and discussions inspired my research. In addition, I would like to thank my co-authors David Bloom, David Canning, Klaus Prettner, and Johannes Schünemann with whom I am collaborating on a joint project. I am particularly grateful to David Bloom, who gave me the opportunity of a research visit at *Harvard University*.

Furthermore, I wish to thank my wonderful colleagues at the *Chair for Population Economics*, who provided a stimulating and amicable work environment. It was a particular pleasure to work with Andreas Backhaus, Sebastian Link, and Gerrit Meyerheim, who accompanied me throughout my doctoral studies.

Moreover, I am grateful to the *Elite Network of Bavaria (ENB)*, which provided generous funding for my doctoral studies through the International Doctoral Program *Evidence-Based Economics (IDK-EBE)*. A special thanks goes to the program directors, Florian Englmaier and Joachim Winter, as well as to Silke Englmaier, Susan Fay, and Julia Zimmermann for their great help in handling administrative matters.

I owe gratitude to numerous people who contributed with their comments and suggestions to the various papers in this doctoral thesis.

For *"Can Education Compensate the Effect of Population Aging on Macroeconomic Performance? Evidence from Panel Data"*, the second chapter of this dissertation, which is joint work with Uwe Sunde, we thank four anonymous reviewers, participants of the VfS Annual Conference 2016 on "Demographic Change", the Second CREA Workshop on Aging, Culture, and Comparative Development in Luxembourg, the 2017 meeting of the VfS Population Economics Committee, the ESPE Annual Conference 2017, the EEA Annual Conference 2017, seminar participants at the Vienna Institute of Demography, the ifo Institute, and at LMU Munich, as well as David Bloom, Lukas Buchheim, David Canning, Christian Dustmann, Oliver Falck, Gustav Feichtinger, Alexia Fürnkranz-Prskawetz, Michael Grimm, Wolfgang Lutz, Klaus Prettner, Miguel Sanchez-Romero, Gesine Stephan, Joachim Winter, and Rudolf Winter-Ebmer for helpful comments and suggestions. Special thanks go to Alexia Fürnkranz-Prskawetz, Bernhard Hammer, and Elke Loichinger for sharing their labor force projection data.

For *"Health and Economic Growth: Reconciling the Micro and Macro Evidence"*, the third chapter of this dissertation, which is joint work with David Bloom, David Canning, Klaus Prettner, and Johannes Schünemann, we thank Ana Abeliansky and Alexander Khoury for helpful comments.

For *"Life Expectancy and Life-Cycle Wages: Evidence from the Cardiovascular Revolution in U.S. States"*, the fourth chapter of this dissertation, I would like to thank participants of the RES Annual Conference 2018, the International Conference on "Demographics, Immigration, and the Labor Market" in Nuremberg, the ESPE Annual Conference 2018, seminar audiences at the University of Munich, as well as David Bloom, Lukas Buchheim, David Canning, Oliver Falck, Uwe Sunde, and Joachim Winter for helpful suggestions.

For *"Population Aging and Income Inequality: Evidence from the Cardiovascular Revolution in U.S. States"*, the fifth chapter of this dissertation, I would like to thank Oliver Falck, Uwe Sunde, and Joachim Winter for helpful suggestions.

For *"Democracy, Inequality, and Institutional Quality"*, the seventh chapter of this dissertation, which is joint work with Uwe Sunde, we thank participants of the 24th Silvaplana Workshop in Political Economy, the EEA Annual Congress 2015, seminar audiences at the University of Munich and the University of St. Gallen, as well as Matteo Cervellati, Mark Gradstein, Florian Jung, and Frank Windmeijer for helpful comments and discussions.

Finally, I am indebted to my family for their support and encouragement. They kept up my morale and helped me to concentrate on this dissertation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

“Why are we so rich and they so poor?” (Landes, 1990, p. 1), or “Why are some countries so much more productive than others?” (Hall and Jones, 1999, p. 84). These questions stand at the core of research on the long-run determinants and processes of socioeconomic development. In their attempts to answer these questions, seminal contributions identified several fundamental and proximate factors as key determinants for macroeconomic performance and development in the long run. Specifically, these determinants include geography (Diamond, 1997; Gallup, Sachs, and Mellinger, 1999), institutional quality (North, 1990, 1991; Hall and Jones, 1999; Acemoglu, Johnson, and Robinson, 2001, 2005; Rodrik, Subramanian, and Trebbi, 2004), education (Benhabib and Spiegel, 1994; Glaeser et al., 2004), health (Bloom, Canning, and Sevilla, 2004; Weil, 2007), the demographic transition (Galor and Weil, 2000; Galor and Moav, 2002; Cervellati and Sunde, 2005) and the demographic dividend (Bloom, Canning, and Sevilla, 2003; Bloom, Canning, and Fink, 2010), as well as preferences (Dohmen et al., 2015; Falk et al., 2017).

Using state-of-the-art empirical techniques, this dissertation extends and complements this literature along several dimensions: First of all, this research sheds new light on the non-monotonic effects of population aging on macroeconomic performance and identifies education as a powerful force in compensating the negative consequences of demographic change. Moreover, this thesis demonstrates the pivotal role of health improvements for subsequent economic performance in terms of growth, aggregate wages, educational attainment, and income inequality. In particular, I establish a causal link between health in terms of adult life expectancy and aggregate wages as well as income inequality. Finally, this dissertation uncovers important interactions between economic performance, inequality, and political institutions in shaping institutional quality, which have gone largely unnoticed by the empirical literature. The analysis proceeds in six chapters which are self-contained and can be read independently. Nonetheless, they all share a common theme: the long-run interdependence of demographic change, economic performance, and institutional quality. The final chapter provides some brief personal thoughts regarding the research on population aging, macroeconomic performance, and institutional quality.

In the first part of this dissertation, which comprises Chapters 2 to 5, I examine how population aging, health improvements, and demographic change affect macroeconomic performance as measured by output levels, economic growth, aggregate wages, and—from a broader perspective—income inequality. This focus is motivated by the powerful effects of transformations of the demographic structure on economic development, which can occur along several dimensions: First of all, individuals become healthier and grow older thanks to medical advancement and improved living standards. Consequently, workers tend to exhibit a higher productivity resulting from raised physical capacity and cognitive functioning (Bloom and Canning, 2000; Feyrer, Politi, and Weil, 2017). Moreover, individuals invest more in educational attainment, if health improvements take place at a sufficiently young age (Ben-Porath, 1967; Cervellati and Sunde, 2013). This increase in investment in schooling transforms the life-cycle earnings trajectory of the average worker: On the one hand, individuals start to work later, thereby reducing their life-time labor supply; on the other hand, they gain a higher return for every additional year of work experience, more than compensating them for the foregone income. In addition, higher education better protects workers at older ages from income shocks, because they can select into occupations that are less physically demanding but require more human capital. Consequently, life-cycle earnings profiles increase more steeply at younger ages and flatten out more slowly at higher ages, thereby raising life-cycle income inequality (see Ashenfelter and Rouse, 1999, as well as Chapters 4 and 5 for details). Furthermore, a longer prospective life span creates a greater need for retirement savings (Bloom and Canning, 2000; De Nardi, French, and Jones, 2009). Finally, individuals usually reduce their fertility as a result of higher opportunity costs for child rearing (Becker, 1960) and better access to family planning (Pritchett, 1994). On the aggregate level, this fertility adjustment causes a change in the demographic structure to a transitory state with few young and few old people. This transformation creates the opportunity for a demographic dividend in terms of substantial gains in income per capita (Bloom, Canning, and Sevilla, 2003). Further shifts in the age composition of the population after the demographic dividend may also considerably slow down prospective economic growth as a result of dissaving, increased need for transfers, or life-cycle productivity patterns (Chapter 2). Hence, health gains, population aging, and demographic change constitute important determinants of macroeconomic performance in the long term, motivating the empirical investigations in this dissertation.

First, Chapter 2 focuses on the effects of aging on macroeconomic performance. This chapter is joint work with Uwe Sunde and presents our paper *"Can Education Compensate the Effect of Population Aging on Macroeconomic Performance? Evidence from Panel Data."* This chapter addresses three questions that have remained open in light of the existing literature: How do population aging and contemporaneous changes in aggregate human capital affect macroeconomic performance? Can investment in education offset the (potentially negative) effects of population aging? And, finally, what are the corresponding prospects of future economic development? In order to answer these questions, we proceed in three steps.

In a first step, we conduct an empirical development accounting exercise based on a model that encompasses the empirical frameworks used in the existing literature. Based on a cross-country panel of more than 130 countries for the period 1950 to 2010, we estimate how changes in either the age structure of the workforce or in the distribution of human capital affect macroeconomic performance in terms of GDP levels and growth rates.

In a further step, we explicitly consider the interactions between aging and changes in the skill composition. This analysis complements and extends the existing literature, which, with few exceptions, focused either on population aging or changes in the human capital endowment in isolation. Our results reveal that population aging has substantial consequences for economic performance, even when accounting for changes in the education composition, and that the demographic structure of the workforce and education both jointly affect economic performance. Moreover, the demographic structure affects economic performance non-monotonically, implying heterogeneous prospective development paths conditional on the extent of demographic change.

In a final step, we use our estimates to conduct quantitative exercises that shed light on the relative importance of the changes in the age and in the skill composition of the workforce that occur as consequence of the ongoing process of population aging. In particular, we project the macroeconomic performance in terms of GDP for several alternative scenarios that use the projected changes in age composition and education until 2050. We compare these projections to counterfactual scenarios that fix the age composition or human capital at current levels. According to these quantitative exercises, advancing population aging and a slowdown in educational attainment will dampen economic performance, particularly in developed economies where aging is especially pronounced and the population has already attained fairly high levels of education throughout all age cohorts. Investment in education constitutes a powerful force in compensating the negative consequences of population aging; however, our results also suggest that even enhanced investments in education are unlikely to completely offset the effects of population aging in the countries that face the greatest pressure of population aging. In contrast, for economies with a relatively stable demographic structure, aging is projected to have rather neutral effects on macroeconomic performance, while the projected increase in human capital implies a positive prospective performance.

Chapter 3 analyzes the association between population health and economic performance at the macro level. This chapter is joint work with David Bloom, David Canning, Klaus Prettner, and Johannes Schünemann and presents our paper *"Health and Economic Growth: Reconciling the Micro and Macro Evidence."* In particular, the evidence presented in this chapter aims at reconciling the micro-based with the macro-based approach of estimating the effect of health on economic growth.

On the one hand, the micro-based approach adds up the microeconomic effects of health to infer the implications for aggregate income. For example, Weil (2007) models output using an aggregate production function that includes the stock of health as measured by the adult survival rate in conjunction with micro-based returns to health. His calibration exercise suggests that health is a vitally important form of human capital deserving central attention in the development process.

On the other hand, the macro-based approach relies on growth regressions (see, for example, Sala-i-Martin, 1997; Durlauf, Johnson, and Temple, 2005). Strong cross-country correlations between measures of aggregate health and per capita income are well established (Preston, 1975; World Bank, 1993). Correspondingly, higher incomes promote access to many of the goods and services, such as a nutritious diet, safe water, sanitation, and good health care, which, in turn, improve health and longevity.

Our paper compares the size of micro-based estimates of the effect of health on wages with the macro-based estimates of the return to health on worker productivity. To this end, we estimate a production function model of economic growth, keeping our specification as close as possible to that of Weil (2007), thereby permitting a direct comparison between our estimates and his calibration. Estimating an aggregate production function using cross-country data is difficult, because reverse causality, omitted variables, and measurement error may bias the parameter estimates. We try to address these issues adequately.

According to our estimates, an increase in adult survival rates of ten percentage points is associated with a 9.1-percent increase in labor productivity. The corresponding 95-percent confidence interval includes the estimate derived by Weil (2007). Hence, our macro-based results conform with his micro-based results, thereby reconciling both approaches of estimating the effect of health on economic growth. Overall, our results suggest that public health measures might be an important lever for fostering economic development. As this work is still preliminary, however, we do not provide extensive robustness checks.

Importantly, reconciling the micro-based and the macro-based approach would imply that the return to health on the macro level could be derived from estimates based on micro-level data. This finding would be of practical importance for two reasons: First, low- and middle-income countries, which would profit most from health improvements, often lack reliable macro data on health outcomes and even economic performance. In this context, micro-level data—for example, based on surveys—might allow to gauge the potential benefits of health interventions. Second, micro-based econometric identification strategies might provide more reliable estimates of the causal effect of health on economic performance compared to macro-based strategies that require strong identification assumptions.

In Chapter 4, I investigate the effects of population aging induced by medical advancement on economic performance in terms of aggregate wages. This chapter is based on my paper *"Life Expectancy and Life-Cycle Wages: Evidence from the Cardiovascular Revolution in U.S. States."* I confine my perspective to the United States, which provide quasi-experimental variation for the econometric identification of a causal link from adult life expectancy to aggregate wages. In particular, the number of older but also healthier workers increased substantially following the introduction of novel treatment procedures for cardiovascular diseases. In this context, my research addresses three questions that complement and extend the existing literature. First, do improved health conditions, as measured by adult life expectancy, lead to more productive workers? Moreover, do health shocks affect the population homogeneously? And, finally, what are potential channels for a causal link?

In order to answer these questions, I exploit variation in the unexpected sharp decline in mortality rates from cardiovascular diseases among U.S. states beginning in the 1960s. This decline is used as an instrument for adult life expectancy in a balanced ten-year panel from 1940 to 2000 for the 48 contiguous U.S. states. The identification strategy exploits initial differences in mortality from cardiovascular diseases across U.S. states in 1960 when there existed little treatment possibilities for cardiovascular diseases. Between 1960 and 1970, a number of path-breaking innovations in the treatment of cardiovascular diseases were introduced and behavioral risk factors identified. The availability of these novel treatments as well as follow-up inventions and public education about risks helped to considerably reduce mortality from cardiovascular diseases between 1970 and 2000. The decline in mortality entailed a substantial increase in adult life expectancy, which varied across states, depending on the initial prevalence of cardiovascular diseases. Therefore, this quasi-experimental source of variation allows the estimation of a differences-in-differences model, where all states are treated but with varying treatment intensities. Economic performance as the main outcome is proxied by aggregate wages on the state level for different age cohorts and the workforce in total.

The paper contributes to the literature in several ways. First, the empirical results establish a positive causal link between adult life expectancy and aggregate wages. In particular, the decline of mortality from cardiovascular diseases in the U.S. from 1968 onward led to an increase in life expectancy at 50 of approximately 3.16 years. According to the main results, this rise in life expectancy caused a wage increase for the 45- to 54-year-olds of roughly 9,762$. This wage hike corresponded to 47 percent of the wage change observed in the same time window. Furthermore, the results reveal that wage gains accrued to workers in the prime-age group between 25 and 54 as well as to old-age workers above 65. Compared to earlier generations, the life-cycle earnings profile of an average worker thus increases more steeply at younger ages, while it flattens out more slowly at higher ages. Overall, this pattern is consistent with a workforce that over time becomes healthier at any given age, and at higher ages in particular. An optimistic interpretation of this result suggests that health gains for prime-age and old-age workers might boost economic performance for aging societies and thus confine (some of the) potentially adverse effects of demographic change described in Chapter 2.

Another contribution is the focus on measurement of health conditions in the context of age-specific outcomes. Specifically, life expectancy as a proxy for average health may over- or understate the true health status of the population if measured at the wrong age. For example, consider the substantial increase of life expectancy at birth following the invention of vaccines and antibiotics. As important as this health shock was, it may grossly overstate the average health improvement of the median American who is around age 30 at this time. Overall, the findings suggest that mismeasurement leads to downward-biased estimates, if the change in average population health is overstated. Therefore, age-specific heterogeneity in the effect of health shocks and mismeasurement might be a reason for null results of life expectancy at birth on GDP per capita found by Acemoglu and Johnson (2007, 2014), Hansen (2014), and Bloom, Canning, and Fink (2014). This measurement problem cannot be mended by the instrumentation strategy employed in these papers, because the instrument itself suffers from the same conceptual shortcoming as the endogenous variable. Hence, the published estimates can be considered a lower bound for the causal effect of life expectancy on growth.

Lastly, this study makes progress in analyzing potential channels through which adult life expectancy affects aggregate wages. Specifically, the timing of wage hikes suggests that potential channels are direct health improvements, especially in the short run, and higher educational attainment as well as changes in individual behavior toward a more healthy lifestyle in the long run. In contrast, adjustments in labor supply cannot explain the wage increase, because labor force participation rates as well as usual working hours and weeks either declined or remained unchanged during the treatment period.

Chapter 5 examines the effect of population aging on income inequality. The chapter is based on my paper *"Population Aging and Income Inequality: Evidence from the Cardiovascular Revolution in U.S. States"* and refers to the setting of Chapter 4. In particular, medical advancement and demographic change led to substantial population aging in the United States during the twentieth century. This fundamental transformation of the demographic structure was accompanied by a substantial rise of investment in educational attainment, increased saving, better population health, a temporary preponderance of the working-age relative to the dependent population, and ultimately resulted in a (second) demographic dividend in terms of economic growth. While a plethora of work analyzed the beneficial effects of the demographic dividend on economic performance in terms of growth, little research examined the quantitative link between population aging and economic inequality. Hence, this chapter extends and complements the literature by providing a more rigorous empirical investigation of this link.

Based on the identification strategy from Chapter 4, this study quantifies the contribution of population aging to income inequality in the United States between 1940 and 2000. From a theoretical viewpoint, this interrelation is mechanical, though the direction and size of its effect are a priori ambiguous. As individuals age, income inequality evolves within age cohorts: Life-cycle earnings profiles suggest that income inequality rises between ages 25 and 64, when the return to educational attainment unfolds, while the income gap contracts again during retirement. Moreover, within-cohort inequalities tend to accumulate over the life-cycle with inequality being most pronounced among the 55- to 64-year-olds (OECD, 2017). Therefore, changes in the demographic structure cause a composition effect that may intensify or depress income inequality conditional on the relative size of age cohorts. For example, income inequality may first increase and later fall as large baby boomer cohorts work their way through the demographic structure. Moreover, population aging may contribute to transformations of life-cycle earnings profiles, which become steeper for young ages and flatten out more slowly at higher ages, thereby further reinforcing income dispersion over the life-cycle.

The empirical results in this chapter confirm a positive causal link from population aging to income inequality for a balanced panel of the 48 contiguous U.S. states between 1940 and 2000. In particular, the baseline estimate indicates that, at the margin, a one-percent increase in adult life expectancy, measured at the age of 30, leads to an increase in inequality of 0.9 Gini points, measured on a scale from 0 (perfect equality) to 100 (perfect inequality). The first-stage estimate implies that the decline in mortality rates from cardiovascular diseases between 1960 and 2000 led to an increase in life expectancy at 30 of 7.2 percent compared to 1960. Taken at face value, higher adult life expectancy thus raised pre-tax income inequality overall by 6.48 Gini points. Hence, population aging contributed considerably to the observed rise of pre-tax earnings inequality in the United States. Furthermore, an age-specific analysis reveals that the effect of population aging on income reaches its maximum when measured for young age groups between 20 and 30. In contrast, the effect becomes small and even vanishes for higher ages. Correspondingly, increased income dispersion results from health improvements and higher prospective longevity during working ages rather than from population aging per se. In particular, this finding is consistent with increasing income dispersion over the life-cycle due to increased investment into educational attainment (Cervellati and Sunde, 2013), life-cycle earnings profiles that flatten out more slowly at higher ages (Chapter 4), and wage polarization resulting from skill-biased technical change (Acemoglu and Autor, 2011; Autor and Dorn, 2013).

In the second part of this dissertation, which encompasses Chapters 6 and 7, I examine the role of economic performance in shaping institutional quality in the political as well as in the economic domain. This focus is motivated by the crucial role of institutions granting economic freedom and liberties for socioeconomic development (Acemoglu, Johnson, and Robinson, 2005). In the long run, however, institutional quality is not exogenously given but endogenously determined. In particular, potential determinants of stability and quality of democratic institutions comprise economic performance in terms of income shocks and distribution (Brückner and Ciccone, 2011; Brückner, Ciccone, and Tesei, 2012; Dorsch and Maarek, 2014) as well as demographic pressure in terms of a youth bulge

(Urdal, 2006). Moreover, the empirical literature has identified two key determinants of institutional quality: democratic institutions in terms of constraints on those in power (Acemoglu and Johnson, 2005; Acemoglu and Robinson, 2005; Acemoglu, 2008) and (re-)distributive pressure in terms of economic inequality that might erode institutional quality through influence activities and informality (Chong and Calderon, 2000; Chong and Gradstein, 2007a, 2007b). Chapters 6 and 7 extend and complement this literature by investing important interactions between economic performance and inequality, on the one hand, and political institutions and economic inequality, on the other hand, in shaping institutional quality. Even though these interactions seem to have been present in works by De Tocqueville (1835) and Lipset (1959), they have largely gone unnoticed in the empirical literature.

Chapter 6 investigates how economic performance in terms of income shocks and inequality affects democratic quality. This chapter is joint work with Uwe Sunde and presents our paper *"Income Shocks, Inequality, and Democracy."* Specifically, we test the hypothesis that income shocks trigger major changes in institutional quality, as reflected by transitions between autocracy and democracy, and that the corresponding effect depends crucially on the social environment, as reflected by economic inequality.

This hypothesis is rooted in the theoretical literature of democratic transitions under threats of revolutions (see, for example, Acemoglu and Robinson, 2000, 2005) and the alternative of elite-driven transitions to democracy (see, for example, Lizzeri and Persico, 2004). Accordingly, negative economic shocks might provide an opportunity to overcome autocratic institutions, especially in an environment with high inequality. Conversely, democracy might emerge for economic reasons in environments of low inequality and, thus, low redistributive conflict. In light of these predictions, we hypothesize that an appropriate empirical analysis of the income-democracy nexus should focus on economic shocks and non-marginal changes in democratic quality instead of exploiting continuous variation in income and institutional quality. Moreover, the theory suggests that the effects of income shocks crucially differ by the cohesiveness of society, as reflected by economic inequality. In economically highly unequal societies, negative income shocks are likely to trigger revolts and, thereby, open a window of opportunity for democratization, whereas positive income shocks tend to stabilize oligarchic structures. In economically equal societies, by contrast, positive income shocks do not generate much redistributive pressure, thereby helping to consolidate and improve democratic quality, whereas negative income shocks might erode democracy by creating tensions within the society.

Our results provide support for this hypothesis and document the crucial role of inequality for the effects of economic shocks on the quality and stability of political institutions. In particular, our findings show that negative income shocks unfold a negative effect on democracy in countries with low economic inequality but a positive effect in countries with high inequality. Therefore, our findings reconcile results for positive effects of income on democracy (for example, Benhabib, Corvalan, and Spiegel, 2013; Che et al., 2013) with evidence that negative income shocks have a positive effect on democratic improvements (for example, Brückner and Ciccone, 2011; Aidt and Franck, 2015). In particular, we document an important role of major income fluctuations and a significant asymmetry in interaction with economic inequality. Moreover, this finding also complements evidence from other contributions, suggesting that income unfolds vastly heterogeneous effects on democratic institutions (for example, Moral-Benito and Bartolucci, 2012, for heterogeneity across low and high income countries, and Cervellati et al., 2014, for heterogeneity with respect to colonial history). Finally, our empirical analysis provides some evidence for demographic pressure as another potential determinant of democratic transitions which, however, is too weak to be considered as conclusive.

Finally, Chapter 7 further investigates the determinants of institutional quality. This chapter is joint work with Uwe Sunde and is based on our paper *"Democracy, Inequality, and Institutional Quality."* Specifically, we test the hypothesis that political institutions in conjunction with economic (in-)equality shape institutional quality in the economic domain. This hypothesis implies a non-monotonic effect of democratic institutions on institutional quality conditional on economic inequality, which has not been documented in the empirical literature.

The starting point of our empirical analysis is the conceptual distinction between political and economic institutions: On the one hand, political institutions describe to what extent individuals can engage and participate in the political process via elections and referendums. On the other hand, economic institutions comprise aspects of de facto economic freedom as well as institutional features that directly affect the incentives for entrepreneurial activities and investment, such as bureaucratic efficiency and impartiality of the judiciary. There are important conceptual differences, which relate to the nature of institutions and their perception. Economic institutions are mostly implemented by laws, which have been passed by the government, and reflect de facto liberties of individual citizens in the economic domain. In contrast, political institutions, in terms of democracy, the constraints on the executive, or the ability to vote, reflect legally codified, constitutional rules. In this sense, political institutions can be seen as determinants of economic institutions but not vice versa (see, for example, Acemoglu, Johnson, and Robinson, 2005). Likewise, these types of institutions differ inherently in their nature: Political institutions serve as constraints for politicians and the government, whereas economic institutions enable private actors to interact and achieve their goals (see, for example, Voigt, 2013).

The main contribution of our study is the identification of a robust empirical interaction between democracy and equality in shaping the quality of the economic institutions. Our empirical strategy exploits variation in democratic quality and income equality within countries over the period 1970 to 2010, thereby conditioning on country-specific and time-specific unobserved heterogeneity that might influence institutional quality, and on controls for institutional quality in the past. We employ different estimation methods to account for the well-known problems in dynamic panels. Irrespective of the estimation method and the underlying identification assumptions, the results reveal a robust, negative direct effect of democracy but a significant positive interaction between democracy and equality in shaping the quality of economic institutions. In terms of size, this interaction term is large enough to render the marginal effect of democracy on institutional quality negative for high levels of inequality in the data set. Hence, our empirical model predicts institutional quality to be highest in democracies with low inequality. At the same time, however, our results imply that the beneficial effect of democracy on the quality of economic institutions may be eroded by excessive inequality, which constitutes a novel result in the literature.

Finally, we also contribute to an ongoing debate about how institutions should be measured, and what is comprised by different measures that are frequently used in the literature. In particular, while there is a common perception that different measures capture similar underlying institutional features, our findings document that there is substantial variation across measures of political and economic institutions, and that the correlation is lower than commonly thought, complementing the conceptual arguments by Voigt (2013). To our knowledge, our analysis is among the first to open this black box and test a theoretical prediction using a variety of different measures of political and economic institutions. The results indicate a very robust interaction between political institutions and (in-)equality within and across countries that has gone largely unnoticed in the existing literature and that holds across combinations of various measures.

Chapter 2

Can Education Compensate the Effect of Population Aging on Macroeconomic Performance? Evidence from Panel Data[1]

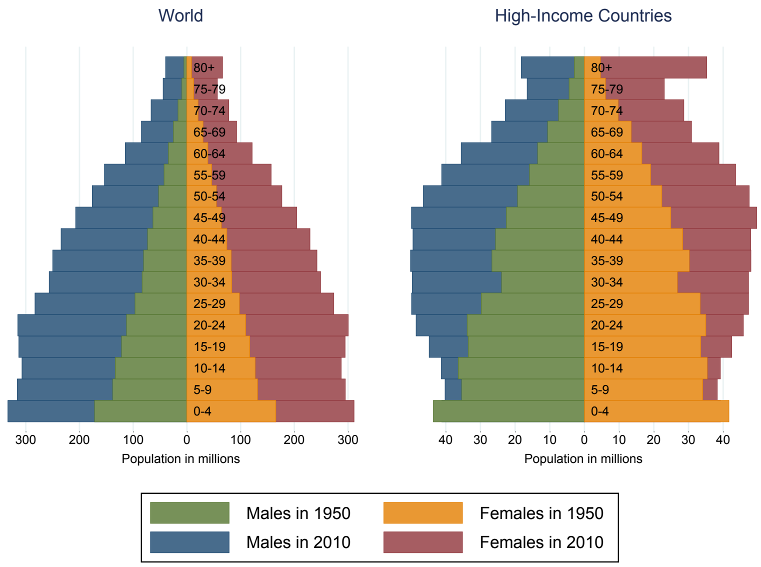
2.1 Introduction

Population aging is one of the most important economic and social challenges in the twenty-first century. With increasing life expectancy and falling fertility, the populations of most countries grow older, resulting in substantial shifts in the age composition of workforce and population at large. At the same time, the demographic transition and the associated shift in the age distribution imply substantial changes in the aggregate stock of human capital as well as its age distribution, as relatively large cohorts with low or moderate levels of formal education are replaced by relatively small cohorts with high levels of formal education.

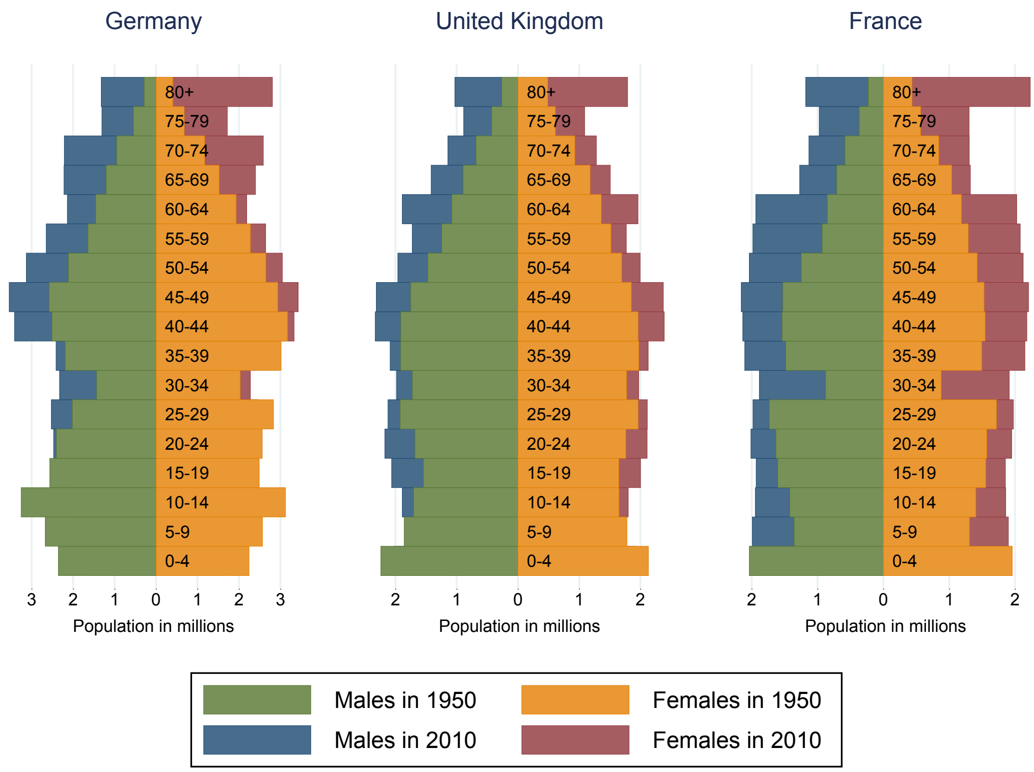
This can be illustrated by the changes in the age structure of populations over long period. Panel (a) of Figure 2.1 plots the age structure in the world and in high-income (OECD) countries in 1950 and 2010. Evidently, not only the size of the world population has changed over this period but, in particular, also the age composition. Whereas in a global perspective the population has increased rather uniformly across all ages, with a slowdown only visible for the youngest cohorts below 20 years of age, aging is much more pronounced among the high-income countries. However, even within the group of high-income countries, there are substantial differences in the demographic dynamics. Panel (b) of Figure 2.1 plots the corresponding patterns for Germany, the United Kingdom,

[1] A revised version of this paper is accepted for publication in *Economic Policy* and available online at `http://doi.org/10.1093/epolic/eiy011`. Please refer to the published version.

and France in 1950 and 2010. In Germany, the age composition of the population is most uneven, with the consequence of a stronger aging momentum than in the UK and, in



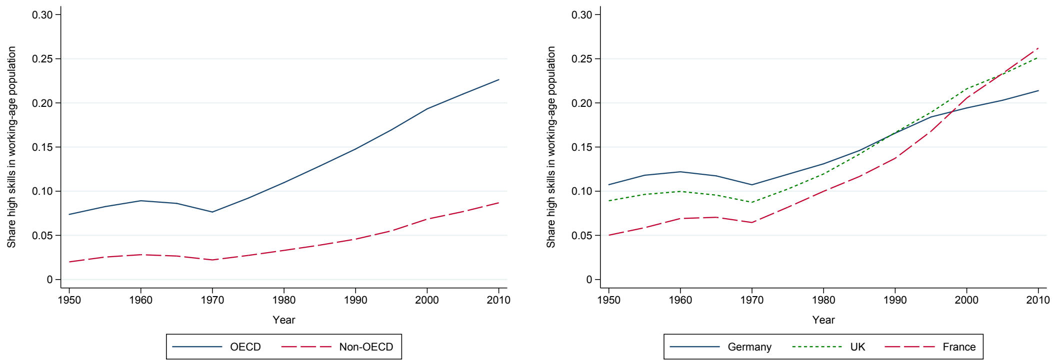
(a) World and High-Income Countries



(b) Germany, United Kingdom and France

Figure 2.1: Population Dynamics – Selected Regions

Data source: United Nations, Department of Economic and Social Affairs (2015). World Population Prospects: The 2015 Revision.



(a) OECD and Non-OECD Countries

(b) Germany, United Kingdom and France

Figure 2.2: Dynamics of Educational Attainment

particular, in France, where the age composition is fairly uniform at ages below 65.

These demographic changes have important consequences for productivity and human capital. The shift in the age composition has implications for the informal, experience-related human capital embodied in the population. This follows from the empirically well-documented age-experience profile from life-cycle models of human capital (Ben-Porath, 1967). At the same time, populations differ greatly in their formal education attainment, both across age groups and across countries. Younger cohorts typically exhibit much higher levels of schooling and formal training. Figure 2.2 documents the secular increase in the share of high skilled over the period 1950 to 2010 for high-income (OECD) and non-OECD countries in Panel (a), as well as for three developed countries, Germany, the UK, and France, in Panel (b).

Although these forceful demographic dynamics can be expected to have major implications for macroeconomic performance, the joint effects of population aging and of changes in the human capital endowment for macroeconomic performance are still not well understood. Whereas the economic consequences of aging and of changes in human capital have been investigated in isolation, their interactions have been largely neglected in the existing literature.

This paper addresses three questions that remain open in light of the existing literature: How do population aging and the contemporaneous changes in aggregate human capital affect macroeconomic performance? Can investment in education offset the (potentially negative) effects of population aging? And, finally, what are the corresponding prospects of future economic development?

Using data from a cross-country panel of more than 130 countries for the period 1950 to 2010, we investigate empirically how changes in the age structure of the workforce and in the distribution of human capital affect macroeconomic performance in terms of levels and growth rates. The investigation is based on an extended empirical development accounting model that encompasses the empirical frameworks used in the existing literature and that allows estimating the distinct effects of aging and human capital. These estimates can be used for a detailed analysis of the relative importance of aging and human capital dynamics for the projected development paths of countries around the world until 2050, and for a quantitative assessment of different scenarios of aging and education acquisition.

The analysis proceeds in three steps. The first step sets the stage by restricting the analysis to the effects of population aging and of changes in the aggregate human capital endowment for macroeconomic performance in isolation from each other, thereby replicating the existing evidence in the literature. The estimation results reveal that changes in the age composition of the work force significantly affect economic performance. The estimates mirror the well-known hump-shaped individual productivity patterns from micro studies, with the largest positive effects being associated with prime working ages and smaller effects for young and old population segments. Likewise, the levels and dynamics in aggregate human capital are shown to affect economic performance independently from the demographic structure.

In the second step, the empirical analysis explicitly considers the interactions between aging and changes in the skill composition. This analysis complements and extends the existing literature, which, with few exceptions, has largely been restricted to focusing either on population aging or changes in the human capital endowment in isolation. The results reveal that population aging has substantial implications on economic performance even when accounting for changes in the education composition, and that the demographic structure of the workforce and education both jointly affect economic performance. Moreover, the demographic structure affects economic performance non-monotonically, implying heterogeneous prospective development paths conditional on the extent of demographic change. At the same time, there is little evidence for eroding productivity of human capital attained in terms of formal education in older cohorts.

In the third and final step of the analysis, the estimation results are used to conduct quantitative exercises that shed light on the relative importance of the changes in the age and in the skill composition of the workforce that occur as consequence of the ongoing process of population aging. In particular, based on the empirical estimates, macroeconomic performance is projected under several alternative scenarios that use the projected changes in age composition and education. These projections are compared to counterfactual scenarios that fix the age composition or human capital at current levels. According to these quantitative exercises, aging and a slowdown in education attainment will dampen economic performance particularly in developed economies, where aging is especially pronounced and the population has already attained fairly high levels of education throughout all age cohorts. Investment in education turns out to be a powerful force in compensating the negative consequences of population aging. However, the results also suggest that even enhanced investments in education are unlikely to completely offset the effects of population aging in the countries that face the greatest pressure of population aging. In contrast, for economies with a relatively stable demographic structure, aging is projected to have rather neutral effects on macroeconomic performance, while the projected increase in human capital implies a positive prospective performance.

Furthermore, the results provide an estimate of the elasticity of substitution between the age composition and the human capital endowment of a country. This elasticity provides new insights into the change in the distribution of human capital that is needed in order to offset the effects of changes in the age composition of the workforce. The quantitative estimate for this elasticity suggests that aging-related shifts in the composition of the population require substantial increases in the education of young cohorts.

This paper contributes to the literature in multiple ways. Several contributions in macro-development have focused on the consequences of aging by focusing on the implications of variation in the young- and old-age dependency ratio for the demographic dividend (Bloom and Williamson, 1998; Bloom, Canning and Sevilla, 2003), and, more recently, Aiyar, Ebeke, and Shao (2016), and Acemoglu and Restrepo (2017) for productivity and technical change. Other contributions have analyzed the effects of aging and skills on growth. Feyrer (2007) finds that the age composition of the workforce affects macroeconomic performance, mainly through total factor productivity. Maestas, Mullen, and Powell (2016) use variation in aging across US states over the period 1980-2010 to estimate the growth effect of aging and find a substantial negative effect. However, these studies only indirectly account for the changes in human capital and its age composition. In contrast, Cuaresma, Lutz, and Sanderson (2014) investigate the joint effect of skills and aging. Instead of conducting a cohort-based analysis that accounts for the distribution of skills and aging as we do in this paper, they look at the role of labor force participation and dependency ratios. In contrast, Sunde and Vischer (2015) show that human capital affects output growth through the productivity of production factors and the potential to innovate (Lucas, 1988; Aghion and Howitt, 1992), or to adopt and diffuse new technologies (Nelson and Phelps, 1966). The approach taken in this paper incorporates these different contributions into a single coherent framework. This allows investigating the relative importance of changes in the age composition and in the skill composition of the population, shedding light on the robustness of earlier results. Thereby, we provide a systematic investigation and decomposition of aging effects through shifts in the demographic composition and changes in the human capital distribution which is missing in the existing literature. The findings indeed point to interactions between population aging and changes in the human capital composition, suggesting that restricting attention to only one dimension delivers an incomplete picture.

To our knowledge, the only two papers that go in a similar direction are by Lindh and Malmberg (1999) and Cuaresma, Loichinger, and Vincelette (2016). However, the analysis by Lindh and Malmberg (1999) is confined to using cross-country data for OECD countries, whereas Cuaresma, Loichinger, and Vincelette (2016) focus on European countries. Our analysis is based on a theoretically founded empirical framework that encompasses frameworks used previously and presents estimation and projection results for a long panel data set for more than 130 countries. Moreover, the estimates presented below contribute by allowing to conduct counterfactual simulations of economic performance under alternative scenarios of aging, human capital dynamics, labor force participation, and productivity. Another novelty are the estimates for an upper bound of the semi-elasticity between changes in the age structure and changes in human capital—and of changes in labor force participation and productivity improvements—which are required to offset the macroeconomic consequences of changes in the age composition in the most favorable case.

The analysis is also related to, and complementing, microeconometric work on age-education decompositions of labor earnings. Work by Card and Lemieux (2001) has used models with imperfect substitution between similarly educated workers in different age groups to study the dynamics of the college wage premium. More recent work by Acemoglu and Autor (2011) and Autor and Dorn (2013) shows for census data and tasks how skill-biased technological progress and changes in the supply of skill levels across cohorts has led to wage polarization in the United States. Vandenberghe (2017) investigates whether a better educated and more experienced workforce contributes to the recent rise in total factor productivity (TFP). Our estimation and projection results complement these studies by providing novel insights into the consequences of population aging and demographic change. In analogy to the approach popularized by Card and Lemieux (2001) and applied by Fitzenberger and Kohn (2006), we develop a decomposition that allows estimating elasticities of substitution between demographic aging and changes in the education structure. Our empirical findings also complement recent evidence for the effect of aging on productivity and wages. For instance, Göbel and Zwick (2013) find that productivity among employees is highest around 50 and only find modest declines in the productivity at older ages, while Börsch-Supan and Weiss (2016) find that there are (almost) no negative aging effects on productivity for production line workers before age 60. Complementing this, Mahlberg et al. (2013) find little evidence between productivity or wages at the firm level and the share of older employees in that firm. The findings for the aggregate level presented in this paper deliver macroeconomic age profiles that are consistent with these findings.

The quantitative analysis sheds new light on the potential implications of aging and education dynamics for growth. Recent work by Acemoglu and Restrepo (2017) suggests that directed technical change and a rapid adoption of automation technologies might provide a countervailing force to the negative growth effects of population aging, particularly in countries that undergo more pronounced demographic changes. Our findings allow quantifying how large, ceteris paribus, the productivity improvements of directed technical change would have to be in different countries in order to fully offset the effects of population aging and the associated education dynamics.

The remainder of this paper is structured as follows. Section 2.2 presents our methodology and empirical framework. A data description is provided in Section 2.3. Section 2.4 provides estimation results and Section 2.5 presents the implications of these estimation results for future economic performance, using different scenarios of aging and education projections. Section 2.6 concludes.

2.2 Methodology

The analysis is based on an aggregate production framework that underlies the standard development accounting model as in Benhabib and Spiegel (1994) and Hall and Jones (1999). Output Y is produced as a function of total factor productivity A, physical capital K, and human capital H of the form:

$$Y_{it} = A_{it} K_{it}^{\alpha} H_{it}^{1-\alpha} \tag{2.1}$$

Subscripts i and t denote cross-sectional units (countries) and time units (five-year intervals), respectively. Dividing by the labor force (working-age population) L_{it} delivers the output per worker in intensive form

$$y_{it} = \frac{Y_{it}}{L_{it}} = A_{it} k_{it}^{\alpha} \left(\frac{H_{it}}{L_{it}} \right)^{1-\alpha}$$

with $k_{it} = \frac{K_{it}}{L_{it}}$ being capital per worker.

The aggregate stock of human capital H_{it} is a function of human capital per worker h_{it} and the overall quality of the labor force Q_{it} as a function of the demographic structure of the workforce and cohort-specific productivity parameters. Quality of the labor force is assumed to be a simple size-weighted average

$$H_{it} := h_{it} Q_{it} = h_{it} \left[\pi_1 L_{it}^1 + \cdots + \pi_k L_{it}^J \right], \tag{2.2}$$

where $L_{it}^1, \ldots, L_{it}^J$ denote the labor force of each age cohort in the workforce and $\pi_1, \ldots, \pi_J$ the respective productivity of each group.[2] Age-related productivity differences can be related to differences in physical strength, or, more likely, correspond to differences in

[2] Alternatively, one could model the quality of the labor force more flexibly using a general constant elasticity of substitution (CES) form as in similar settings applied to different contexts, see, for example, Sato (1967), Hellerstein and Neumark (1995), Card and Lemieux (2001) or, more recently, Vandenberghe (2017). The CES specification would provide an even more flexible framework that allows for different productivity parameters across age groups and flexible substitution patterns between age groups and, potentially, with varying quality of human capital across age groups. Analyzing the simple case with substitution elasticity of one between physical and human capital and perfect substitution across age cohorts has the advantage of a straightforward derivation of a linear estimation framework. This assumption is inessential, however, and could be relaxed by working with a CES specification and conducting estimates using a non-linear estimation model.

human capital that is acquired on the job in terms of experience. This is consistent with standard models of human capital acquisition over the life-cycle (Ben-Porath, 1967). The aggregate human capital stock per worker is thus given by

$$\frac{H_{it}}{L_{it}} = h_{it}\left[\left(\frac{1}{L_{it}}\right)\sum_{j=1}^{J}\pi_j L_{it}^j\right] = h_{it}\left[\sum_{j=1}^{J}\pi_j \frac{L_{it}^j}{L_{it}}\right] = h_{it}\left[\sum_{j=1}^{J}\pi_j S_{it}^j\right]$$

with S^j denoting the share of each age cohort in the total labor force such that $\sum_{j=1}^{J} S_{it}^j = 1$. In order to avoid multicollinearity in the empirical model, a reference category S_{it}^r is chosen such that

$$\frac{H_{it}}{L_{it}} = h_{it}\pi_r\left[S_{it}^r + \sum_{j\neq r}\frac{\pi_j}{\pi_r}S_{it}^j\right] = h_{it}\pi_r\left[(1-\sum_{j\neq r}S_{it}^j) + \sum_{j\neq r}\frac{\pi_j}{\pi_r}S_{it}^j\right].$$

The aggregate human capital stock per worker is then given by

$$\frac{H_{it}}{L_{it}} = h_{it}\pi_r\left[1 + \sum_{j\neq r}\lambda^j S_{it}^j\right], \tag{2.3}$$

with $\lambda^j := \frac{\pi_j}{\pi_r} - 1$ denoting the difference in relative productivity between an age cohort j and the reference category. Inserting the expression for the human capital stock per worker in (2.3) into the production function in (2.1) and taking logs yields

$$\begin{aligned}\ln(y_{it}) &= \ln(A_{it}) + \alpha\ln(k_{it}) + (1-\alpha)\ln\left(\frac{H_{it}}{L_{it}}\right)\\ &= \ln(A_{it}) + \alpha\ln(k_{it}) + (1-\alpha)\Big[\ln(h_{it}) + \ln(\pi_r)\Big] + (1-\alpha)\ln\left(1 + \sum_{j\neq r}\lambda^j S_{it}^j\right).\end{aligned}$$

The last term in parentheses can be expected to be close to unity, since the term for productivity ratios λ^j and the share of each age cohort in the total workforce is close to zero for a sufficiently large number of age groups, and correspondingly also their product. Hence, the last term in logarithms can be approximated by

$$\ln\left(1 + \sum_{j\neq r}\lambda^j S_{it}^j\right) \approx \sum_{j\neq r}\lambda^j S_{it}^j. \tag{2.4}$$

Human capital per worker h_{it} is assumed to be a function of an individual worker's skills which can either be high or low. Correspondingly, each skill group is assigned a skill-specific productivity $\{\pi_h, \pi_l\}$. Averaging over the entire economy, human capital per worker is thus the weighted average of the shares of each skill group $\{S_{it}^h,\, 1 - S_{it}^h\}$ multiplied by the respective productivity, or formally

$$h_{it} = \pi_h S_{it}^h + \pi_l(1 - S_{it}^h). \tag{2.5}$$

Taking logs and choosing the low-skill group as reference category, this expression can be rearranged to

$$\ln(h_{it}) = \ln\left[\pi_l\left(1 + \left(\frac{\pi_h}{\pi_l} - 1\right)S_{it}^h\right)\right],$$

which, using the same arguments as before, can be approximated by

$$\ln(h_{it}) = \ln(\pi_l) + \ln\left(1 + \lambda^h S_{it}^h\right) \approx \ln(\pi_l) + \lambda^h S_{it}^h \tag{2.6}$$

with $\lambda^h := \frac{\pi_h}{\pi_l} - 1$ denoting the difference in relative productivity between high-skilled and low-skilled workers. Log output is thus given by

$$\ln(y_{it}) \approx c + \ln(A_{it}) + \alpha \ln(k_{it}) + (1-\alpha)\lambda^h S_{it}^h + (1-\alpha)\sum_{j \neq r} \lambda^j S_{it}^j, \tag{2.7}$$

where $c = (1-\alpha)\big[\ln(\pi_l) + \ln(\pi_r)\big]$ is a constant. By taking first differences, the model is expressed in terms of growth rates:

$$\Delta \ln(y_{it}) \approx \Delta \ln(A_{it}) + \alpha \Delta \ln(k_{it}) + (1-\alpha)\lambda^h \Delta S_{it}^h + (1-\alpha)\left[\sum_{j \neq r} \lambda^j \Delta S_{it}^j\right]. \tag{2.8}$$

Because, in practice, total factor productivity is not observed, we model a country's total factor productivity as being determined by three components: An exogenous time trend ζ_t, which represents freely available technology from the world technological frontier in a given period t, allowing for a technology diffusion process across countries; the past level of output, which, by definition, comprises past TFP; and an idiosyncratic error component ε_{it}, which serves as the error term for the empirical framework. This modeling assumption for TFP is motivated by the strong correlation between initial productivity, reflected by output per worker, and subsequent growth rates (see, for example, Baumol, 1986).[3] Lagged output per worker therefore introduces persistence in the availability of technology within countries into the levels specification. This persistence may for example reflect capital-embodied technology that has been accumulated over time. Consequently, we posit

$$\ln(A_{it}) = \zeta_t + \gamma \ln(y_{it-1}) + \varepsilon_{it}. \tag{2.9}$$

Moreover, this specification implies a further straightforward extension of our estimation framework to long-run productivity differences across countries along other dimensions that might enter equation (2.9) as additional variables (for example, institutions).

Therefore, the empirical model which is used to estimate the effect of the demographic

[3]In an earlier version of this paper, we also included lagged skills in the levels equation. However, the respective variable was always insignificant in the empirical application and did not quantitatively change the overall effect of skills on output. Hence, the variable has been dropped from the specification.

structure of the workforce and the distribution of skills on output is given by

$$\ln(y_{it}) = \gamma \ln(y_{it-1}) + \alpha \ln(k_{it}) + (1-\alpha)\lambda^h S^h_{it} + (1-\alpha)\left[\sum_{j\neq r} \lambda^j S^j_{it}\right] + c_i + \zeta_t + \varepsilon_{it}, \quad (2.10)$$

where c_i allows for country-specific intercepts. This model is estimated with the within-transformation, removing the constant c and accounting for country-specific fixed effects.

In terms of dynamics, we assume that total factor productivity growth of a country is determined by four components: An exogenous time trend τ_t, which represents growth of freely available technology at the world technological frontier in a given period t, allowing for a technology diffusion process across countries; the economy's share of high skills in period $t-1$, which may facilitate the diffusion and adoption of already existing technologies (Nelson and Phelps, 1966) or foster novel innovation (Romer, 1990; Aghion and Howitt, 1992); the past level of output, which, by definition, comprises past TFP; and an idiosyncratic error component u_{it} serving as the error term for the empirical framework.

Consequently, the growth rate of total factor productivity is assumed to take the form

$$\Delta \ln(A_{it}) = \tau_t + \theta S^h_{it-1} + \psi \ln(y_{it-1}) + u_{it}. \quad (2.11)$$

This modeling of technological progress again accommodates for the strong correlation between initial productivity and subsequent growth (Baumol, 1986) and has been widely applied in models that study economic growth in general or the demographic dividend in particular (Fagerberg, 1994; Dowrick and Rogers, 2002; Bloom, Canning, and Sevilla, 2004; Cuaresma, Lutz, and Sanderson, 2014). Specifically, this modeling assumption implies conditional convergence in productivity across countries. In contrast to other models of conditional convergence such as Mankiw, Romer, and Weil (1992), however, this modeling of TFP growth allows for long-run differences in productivity even after the diffusion process is complete. Such differences may enter the estimation model through other variables in equation (2.11). Correspondingly, the estimation equation in growth rates is given by

$$\begin{aligned}\Delta \ln(y_{it}) = {} & \psi \ln(y_{it-1}) + \alpha \Delta \ln(k_{it}) \\ & + (1-\alpha)\lambda^h \Delta S^h_{it} + \theta S^h_{it-1} + (1-\alpha)\left[\sum_{j\neq r} \lambda^j \Delta S^j_{it}\right] + \tau_t + u_{it}. \quad (2.12)\end{aligned}$$

Estimating the model in terms of growth rates also accommodates for the possibility of a unit root in the error term, if income follows a random-walk. Correspondingly, the series will be stationary. As will become clear below, coefficient estimates do not differ substantially between both models, but, unsurprisingly, the levels model is more efficient and explains a larger fraction of the variation. Results for both versions of the model are reported in Section 2.4.

This specification of the estimation framework is very flexible and can be adjusted to obtain the regression models of important other contributions of the literature. For example, the estimation model of Feyrer (2007) is obtained by assuming human capital to be the exponential of a piece-wise linear function of human capital savings and imposing no further assumptions on the structure of TFP growth apart from a common time trend across. Given this set of assumptions, the effect of the demographic structure is contained by total factor productivity.

The specification of Cuaresma, Lutz, and Sanderson (2014) can be obtained under the following assumptions: human capital per worker takes an exponential form as described above, GDP is expressed in terms of per capita instead of per worker terms, and the demographic structure of the workforce is neglected. In this case, the demographic structure enters output through the labor force participation rate and the share of the working-age population in the total population.

Finally, the specification of Sunde and Vischer (2015) is derived by assuming that human capital enters both, productivity and output, in logarithms instead of shares. Further control variables can be included by extending either the TFP residual by lagged level controls or the output by additional terms as a multiplicative or exponential function.

2.3 Data

Data for output and physical capital are from Penn World Tables (PWT) by Feenstra, Inklaar, and Timmer (2015). The main dependent variables are log output per worker and the corresponding growth rate. In robustness analysis, we also use output per capita.

Data for the demographic structure are taken from different sources. The primary source of information about the working-age population for age cohorts in five-year intervals from 15 to 69 as well as for human capital and the corresponding projections is the IIASA-VID database by Lutz et al. (2007).[4] We define age cohorts of the workforce as cohort shares of the total working-age population in brackets 15–19 (below 20), 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 55–59, 60–64, 65–69 (65+). This reflects the potential workforce of a cohort in a given period in the estimation; that is, we refrain from an adjustment for hours worked or employment shares to avoid endogeneity problems that might bias the estimates. In some specifications, the cohorts are collapsed to ten-year intervals in order to reduce the number of parameters to be estimated. The microeconometric evidence on age-productivity profiles discussed in the introduction indicates that the cohort 50–54 represents the most productive cohort. In light of this, we take this cohort as the reference group. Different classifications do not affect the results qualitatively. An alternative data source for population counts and human capital by age is Barro and Lee (2013).

[4]The IIASA-VID projection data are available at `http://www.iiasa.ac.at/web/home/research/researchPrograms/WorldPopulation/Projections_2014.html`.

For this data set, no population projections are available, which is the reason for using the IIASA-VID data as baseline. Alternative population counts and projections as well as young- and old-age dependency ratios are obtained from the United Nations World Population Prospects.[5] Data on life expectancy are obtained from the World Development Indicators provided by the World Bank.[6]

Human capital per worker is proxied by the share of high- and low-skilled individuals in the working-age population. The share of low-skilled workers is defined as the sum of the respective shares of individuals with either no formal education, or primary or secondary schooling only. Correspondingly, the share of high-skilled corresponds to those workers who have received formal tertiary education or equivalent vocational skills. The respective shares are taken from the IIASA-VID database by Lutz et al. (2007). As described in Section 2.2, the share of high-skilled human capital is chosen as reference category. Data are available for up to 139 countries in five-year intervals from 1960 to 2010 (13 time periods in total). Table A.1 in the Appendix reports descriptive statistics.

In the projection analysis, we also make use of data for hours and labor market participation provided by the International Labor Organization (International Labour Organization, 2011).[7] In addition, we use projections for hours and labor market participation for 26 European countries that have been constructed recently by Fürnkranz-Prskawetz, Hammer, and Loichinger (2016).

2.4 Estimation Results

This section reports the estimation results regarding the effect of the age structure and the distribution of skills on economic performance. In a first step, both effects are investigated in isolation, thereby reproducing the analysis conducted in the existing literature. In a second step, we provide evidence for a model that combines both dimensions. The section ends with results from robustness checks and alternative estimation frameworks.

The empirical models are estimated either in levels as in equation (2.10) or first differences as proposed in equation (2.12). Lagged levels of output per worker and the share of high-skilled workers in the population enter all estimation models in levels to control for convergence dynamics of output and technological diffusion, respectively. If not stated otherwise, specifications are estimated for a baseline panel of 120 countries in five-year intervals for the time period 1950–2010.[8]

[5]Data from United Nations World Population Prospects are from the 2015 Revision and available at `http://esa.un.org/unpd/wpp/Download/Standard/Population/`.

[6]The data can be retrieved at `databank.worldbank.org/wdi`.

[7]The data can be obtained online at `http://www.ilo.org/ilostat`.

[8]The robustness material contains results for 139 countries when using alternative data sources.

2.4.1 Demographic Structure

Estimation results for the effect of the age structure of the workforce on output are reported in Column (1) of Table 2.1 for the model in levels. The reference age group in the levels model is the cohort aged 50–54 years. The results are obtained from a specification of the estimation framework with country fixed effects, period fixed effects, and controls for lagged output per worker and capital per worker.

The results reveal that all coefficients for the cohort-specific workforce shares are negative and significant. Therefore, shifting population mass out of the reference cohort 50–54 into another cohort has a negative impact on output. This effect is particularly pronounced for a relative increase of the population group aged 60–64, revealing a negative effect due to population aging. An increase in the population share of this cohort by one-percentage point at the cost of the reference group of age 50–54 implies a decrease in output per worker of roughly 5.5 percent. Population shifts of such size are no exception in the data. Across all workforce shares, around 25 percent of all out-shifts of a cohort are roughly equal to a unit percentage point shift or even larger. The same pattern holds for 25 percent of all in-shifts into a cohort. Furthermore, the estimated negative point estimates are largest for the age cohorts that are either at the very beginning or at the end of their work lives. These patterns are consistent with estimates from disaggregate data mentioned in the Introduction, which suggest that productivity is highest for individuals around age 50, when they have acquired sufficient work experience and on-the-job training. In particular, the results are also in line with a hump-shaped pattern as predicted by standard human capital theory over the life-cycle (Ben-Porath, 1967): For middle-aged cohorts additional productivity gains become smaller as the marginal return from more experience decreases and ultimately declines to zero as the benefits of additional education investments deteriorate with the lower amortization period. At some point, the depreciation rate of human capital outweighs additional gains by experience such that individual productivity decreases in many cases toward the end of the work life. Taken together, the results largely confirm earlier micro-level findings on the effects of population aging (for example, Göbel and Zwick, 2013). Moreover, the joint Wald test on the coefficients of all workforce shares confirms that the overall demographic structure has a significant impact on output.

Column (1) of Table 2.2 contains the corresponding results for the model in (log) differences. To be consistent with the levels model, the differences model uses the change in the age cohort aged 50–54 as reference category. The results essentially replicate those obtained for the levels model qualitatively and quantitatively. In particular, the age pattern and the importance of heterogeneity in the effect of changes in the age composition of the workforce on output growth is very similar. The results also show that the estimated coefficient for lagged output per worker is positive and smaller than one in the levels model, and negative in the model in differences, providing evidence for the usual conditional

Table 2.1: Effects of Aging and Education on Economic Performance: Levels Model

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share < 20	-3.84***		-3.06**	-3.05**	-3.64***	-2.09*	-2.53**
	(1.22)		(1.21)	(1.40)	(1.21)	(1.20)	(1.18)
Share 20–24	-2.37**		-1.84*	-1.82	-3.23**	-1.16	-2.58*
	(1.10)		(1.10)	(1.38)	(1.37)	(1.11)	(1.36)
Share 25–29	-3.56**		-3.12**	-3.18*	-2.77*	-2.58*	-2.30
	(1.42)		(1.39)	(1.89)	(1.47)	(1.32)	(1.41)
Share 30–34	-3.06**		-2.74**	-2.74*	-3.96***	-2.33*	-3.61**
	(1.27)		(1.25)	(1.54)	(1.43)	(1.21)	(1.41)
Share 35–39	-4.01***		-3.71**	-3.63**	-2.97*	-3.33**	-2.49*
	(1.44)		(1.43)	(1.74)	(1.57)	(1.39)	(1.51)
Share 40–44	-1.53		-1.35	-1.37	-1.87	-1.12	-1.81
	(1.32)		(1.29)	(1.62)	(1.41)	(1.26)	(1.39)
Share 45–49	-3.19**		-3.07**	-3.31*	-3.98**	-2.92**	-3.89**
	(1.43)		(1.41)	(1.95)	(1.55)	(1.37)	(1.51)
Share 55–59	-4.66**		-4.37**	-4.76**	-4.40**	-4.00**	-4.16**
	(1.85)		(1.81)	(2.18)	(1.82)	(1.74)	(1.77)
Share 60–64	-5.48***		-5.50***	-5.96***	-6.06***	-5.53***	-6.33***
	(1.37)		(1.34)	(1.76)	(1.50)	(1.30)	(1.48)
Share 65+	-3.06*		-3.24**	-3.35*	-4.21***	-3.47**	-4.22**
	(1.58)		(1.55)	(1.80)	(1.62)	(1.56)	(1.68)
Share high-skill		0.97***	1.08***	0.85***	0.84**	2.45***	2.35***
		(0.34)	(0.40)	(0.32)	(0.41)	(0.76)	(0.77)
Output p.w. (t–1)	0.50***	0.48***	0.48***	0.59***	0.48***	0.46***	0.46***
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)
Capital p.w.	0.32***	0.33***	0.33***	0.29***	0.33***	0.35***	0.34***
	(0.05)	(0.04)	(0.05)	(0.02)	(0.05)	(0.05)	(0.05)
Cohort shares (p–value)	0.01		0.01	0.00	0.01	0.00	0.00
Skill share (p–value)		0.00	0.01	0.01	0.04	0.00	0.00
First-stage F–statistic					13.3	27.9	4.5
Hansen test (p–value)					—	0.25	0.28
Countries	120	120	120	120	120	120	120
Observations	1,098	1,098	1,098	1,053	1,098	1,098	1,098
R^2	0.86	0.86	0.87		0.87	0.86	0.86

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w., measured in logarithms, are included as controls in all specifications. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged shares of high skills of cohorts at the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

convergence patterns.[9] The estimated values for the capital income share α are 0.32 and 0.40 for the specifications in Column (1) of Table 2.1 and Table 2.2, respectively.

The results in the differences specification closely resemble the empirical specifications estimated by Feyrer (2007) and reproduce his results for a different data set.[10]

[9]The bias in coefficient estimates in specifications with lagged dependent variable and fixed effects should be moderate in a relatively long panel with 13 time periods, see Nickell (1981) and Judson and Owen (1999); see also the discussion in Section 2.4.4.

[10]In particular, Feyrer (2007) also finds point estimates that are negative relative to the relatively most productive age cohort. The point estimates are qualitatively very similar to the results for empirical specification proposed in this paper and quantitatively slightly smaller.

Table 2.2: Effects of Aging and Education on Economic Performance: Differences Model

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ Share < 20	-3.68***		-2.83**	-1.24	-4.23***	-3.35***	-5.22***
	(1.16)		(1.19)	(1.13)	(1.36)	(1.26)	(1.81)
Δ Share 20–24	-3.03***		-2.05**	-1.35	-4.03***	-2.54**	-4.91***
	(1.01)		(1.02)	(0.98)	(1.27)	(1.12)	(1.48)
Δ Share 25–29	-3.47***		-2.78**	-1.70	-4.34***	-3.12***	-4.91***
	(1.15)		(1.15)	(1.13)	(1.51)	(1.20)	(1.74)
Δ Share 30–34	-3.92***		-3.41***	-2.48**	-4.10**	-3.15***	-4.20**
	(1.13)		(1.14)	(1.15)	(1.76)	(1.11)	(1.89)
Δ Share 35–39	-4.97***		-4.58***	-3.75***	-4.44***	-4.89***	-4.29***
	(1.22)		(1.23)	(1.20)	(1.50)	(1.37)	(1.56)
Δ Share 40–44	-2.56**		-2.33**	-1.12	-2.70**	-2.36**	-2.60**
	(1.10)		(1.06)	(0.94)	(1.16)	(1.11)	(1.21)
Δ Share 45–49	-3.08***		-2.93***	-1.60*	-4.09***	-2.92***	-4.21***
	(1.12)		(1.09)	(0.94)	(1.19)	(1.09)	(1.24)
Δ Share 55–59	-2.35**		-2.17**	-1.01	-2.65***	-2.28**	-3.13***
	(0.96)		(0.95)	(0.97)	(1.01)	(1.00)	(1.12)
Δ Share 60–64	-5.26***		-5.14***	-3.08***	-5.63***	-5.29***	-6.25***
	(1.20)		(1.20)	(1.06)	(1.41)	(1.29)	(1.57)
Δ Share 65+	-6.61***		-6.22***	-1.71	-6.32***	-6.49***	-6.93***
	(1.67)		(1.64)	(1.43)	(1.80)	(1.70)	(1.99)
Δ Share high-skill		2.68**	3.30***	1.21	1.87	0.64	-3.09
		(1.07)	(1.15)	(0.87)	(1.16)	(4.23)	(4.75)
Share high-skill (t–1)		0.67***	0.55**	0.42***	0.55*	0.71*	0.83**
		(0.24)	(0.25)	(0.15)	(0.33)	(0.37)	(0.41)
Output p.w. (t–1)	-0.21***	-0.24***	-0.23***	-0.02***	-0.24***	-0.24***	-0.23***
	(0.03)	(0.03)	(0.03)	(0.01)	(0.03)	(0.03)	(0.03)
Δ Capital p.w.	0.40***	0.39***	0.41***	0.43***	0.40***	0.41***	0.40***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)
Cohort shares (p–value)	0.01		0.01	0.01	0.00	0.00	0.00
Skills shares (p–value)		0.00	0.00	0.00	0.05	0.01	0.11
First-stage F–statistic					8.5	52.1	5.9
AR(2) test (p–value)				0.65			
Hansen test (p–value)				0.33	—	0.45	0.50
Countries	120	120	120	120	120	120	120
Observations	1,098	1,098	1,098	978	1,053	1,053	1,053
R^2	0.42	0.39	0.43		0.42	0.43	0.40

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The dependent variable is the log difference in output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w., measured in logarithms, are included as controls in all specifications. Column (4) corrects for the dynamic-panel bias using the system GMM estimator by Arellano and Bover (1995) and Blundell and Bond (1998). The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged shares of high skills of cohorts at the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. For system GMM, also the p–values of the AR(2) test are reported. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

2.4.2 Human Capital and Distribution of Skills

As a next step, the analysis focuses on the role of the effect of human capital and the distribution of skills on economic performance. Column (2) of Table 2.1 presents the corresponding estimates for a specification that only includes the share of high-skilled in addition to lagged output and capital per worker. The point estimate of the share of individuals with high-skilled human capital is positive and highly significant. A one-percentage point increase in the share of high skilled of in an economy is accompanied by an increase in output of 0.97 percent.

Column (2) of Table 2.2 presents the corresponding results for a specification in differences. This specification also accounts for the possibility that, conceptually, human capital influences output (growth) through two channels. First, changes in the share of skills account for composition effects of productions factors, which can be accrued to the complementarity of human and physical capital in standard growth models (Solow, 1956; Lucas, 1988). Second, the accumulation of human capital may alleviate the diffusion and adoption of already existing technologies (Nelson and Phelps, 1966) or spur innovation as in the endogenous growth literature (Romer, 1990; Aghion and Howitt, 1992). Not accounting for both channels might lead to a potential bias in the estimates due to the omission of one relevant channel from the estimation, as indicated by the results of Sunde and Vischer (2015). The results provide evidence supporting the specification with both levels and changes of human capital, as suggested by the work of Sunde and Vischer (2015). Both point estimates of levels and changes in the share of high-skilled human capital are positive and individually and jointly significant.[11]

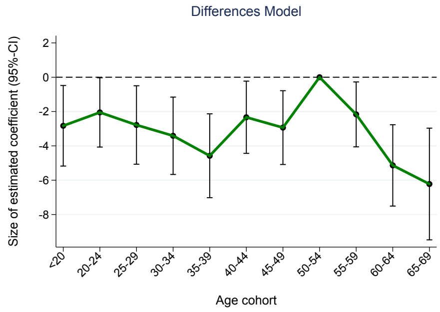
Quantitatively, the results imply that a one-percentage point larger share of skilled workers in the economy is accompanied by an 0.67-percent increase in growth of output per worker over a five-year period. In light of the literature, this effect works through innovation as well as diffusion and adoption of new technologies. Growth of one-percentage point in the share of high-skilled implies an increase in the growth rate of 2.68 percent over five years. The coefficient for lagged output per worker takes negative values for the differences model, indicating conditional convergence, and the coefficient of the capital income share is similar to the earlier results.

2.4.3 Considering Demographics and Skills in Combination

While the results so far have successfully reproduced the findings in the existing literature, the specifications have considered the demographic structure and the influence of human capital in isolation. However, in view of the possibility that the age structure and the human capital composition of the population are correlated and both influence macroeconomic performance, the estimates might suffer from omitted variable bias. In order to investigate this possibility and potential interactions, we now proceed to estimate more comprehensive models that accounts for both the demographic structure of the workforce and the distribution of skills in the population.

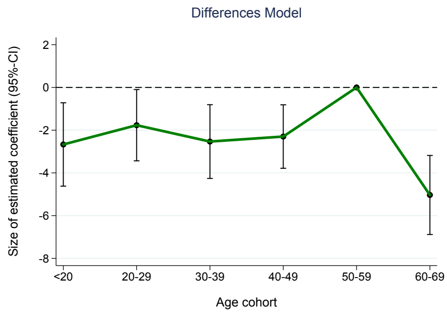
[11] The specification of the model in levels, which follows from equation (2.10), does not contain a term involving the change in the share of high-skilled in the population, because this term emerges from the dynamics of TFP. In unreported estimations, we nevertheless included changes in the share of high-skilled individuals in the specification of the empirical model of the levels estimation to estimate a symmetric empirical specification in both levels and differences and obtain directly comparable estimates of the coefficients of interest. Moreover, this specification provides a natural specification test, because the coefficient of the change in the skill share is hypothesized to be zero in light of the theoretical model (2.10). The findings suggest that the coefficient is indeed not significantly different from zero in an extended version of the specification in Column (2) of Table 2.1, which indirectly supports the empirical model. The estimation results are qualitatively and quantitatively almost identical when estimating a specification of the levels model that does additionally include the change in the share high-skilled, see Table A.2 in the Appendix for details.

(a) Five-year cohorts (Table 2.1, Column 3)



(b) Five-year cohorts (Table 2.2, Column 3)

(c) Ten-year cohorts (Table A.7, Column 3)



(d) Ten-year cohorts (Table A.8, Column 3)

Figure 2.3: Macro Productivity Profiles

Columns (3) of Tables 2.1 and 2.2 present the estimation results for such an extended specification. The coefficient estimates for the age structure and human capital are qualitatively similar, indicating that both dimensions affect macroeconomic performance. In the levels model, the skill pattern appears to exhibit slightly smaller coefficient estimates than in the specification without human capital, while the human capital effect is slightly larger than in the specification without controlling for the age structure. The same is true for the differences model, with the exception of the effect of the share of high-skilled individuals in levels, whose coefficient is also slightly smaller than in Column (2).

Figure 2.3 provides a graphical representation of the estimates of the coefficients for the different age shares obtained with (a) the levels model and (b) the model in differences. Panels (c) and (d) reproduce the respective productivity profiles for an alternative specification using ten-year instead of five-year age cohorts. The graphs illustrate that the age-related coefficients are somewhat smaller in absolute terms for the differences model but otherwise very comparable. Hence, an increase in the share of a specific age cohort relative to the 50- to 54-year-olds leads to a reduction in output. Moreover, the skill distribution positively affects macroeconomic performance through both, the innovation and adoption of technology channel and the composition of production factors. Most importantly, demographic structure of the workforce and human capital both jointly affect output. Therefore, both channels are conceptually relevant for themselves even if they interact substantially.

2.4.4 Alternative Estimation and Identification Strategies

The estimation results presented so far, in particular in the levels specification, were based on a two-way fixed effects estimator with a lagged dependent variable (log output per worker) as additional regressor. Consequently, the coefficient estimates, in particular for the lagged dependent variable might be biased, see Nickell (1981). To investigate whether this might be an issue for the coefficients of interest, Columns (4) of Tables 2.1 and 2.2 present the corresponding results obtained with an estimator that corrects for the potential bias in dynamic panels.[12] The estimation results are qualitatively and quantitatively very similar for the levels model, whereas they reflect some quantitative differences in the differences specification. Alternative estimation results obtained with a levels model without lagged dependent variable also deliver qualitatively similar results.[13] The simulation results shown below will be based mainly on the levels model.

Another potential concern is identification and endogeneity. The identification of the coefficients for aging and human capital so far was based on the implicit assumption that the current workforce (in terms of age structure and skill composition) is the result of fertility and education decisions in the past. Controlling for past income, capital, and country-specific intercepts related to productivity and other time-invariant factors account for country-specific differences in economic performance that might influence, or correlate with, the age and skill composition. Additionally, the results implicitly correspond to an intention-to-treat interpretation, where population shares reflect the potential size of the workforce of each age cohort, instead of accounting for the actual workforce, which might be affected by endogenous labor supply decisions at the extensive or intensive margin, and thus give rise to endogeneity concerns. The estimates thus implicitly assume changes in the workforce be arguably exogenous given the lagged dependent variable, country-fixed effects, and period-fixed effects. The finding of similar results in the levels and differences models is reassuring in this regard, because similar estimates for the respective coefficients are obtained despite the use of alternative variation for identification.

Nonetheless, in order to probe further into potential identification problems caused by unobserved variables that correlate with the factors of production and thus lead to problems of endogeneity bias, we present the results from three alternative identification approaches based on instrumental variables (IV).

[12] In the levels model the bias correction is implemented via the bias-corrected fixed effects estimator of Bruno (2005). To be internally consistent with the theoretical model outlined in Section 2.2, the bias correction in the differences model is performed using the system GMM estimator of Arellano and Bover (1995) and Blundell and Bond (1998).

[13] See Table A.3 in the Appendix. Due to the omission of the lagged dependent variable, the human capital variable picks up some of the persistence, with the consequence of slightly larger coefficient estimates for human capital obtained with this specification. In the following, we restrict attention to the model with lagged dependent variable, which delivers results for the role of human capital that are more conservative in this respect.

A first alternative way to obtain identification is to exploit the fact that the demographic structure of the working-age population follows very stable and predictable dynamics. Concretely, a cohort of individuals aged 40 years at a particular point in time will be of age 50 years ten years later. The same is approximately true for the age composition, because the relative sizes of cohorts of different ages are unlikely to change over time and thus provide valuable predictors over long periods of time. Simultaneously, however, lagged age shares satisfy the exclusion restriction of an instrument, which stipulates that they must be unrelated to unobserved factors driving macroeconomic performance some decades into the future. Hence, demographic dynamics lend themselves naturally to an instrumental variables approach in the present setting of panel data.

On the basis of these considerations, the IV strategy exploits the fact that the relative size of particular cohorts at some point in time predicts the size of these cohorts in the future. At the same time, it is unaffected by economic performance in the future and, thus, exogenous for the purpose of the estimation framework applied here. For a given country-period-cohort cell, this is plausibly the case, in particular, once conditioning on the lagged dependent variable and country-fixed effects. The share of the working-age population of a particular cohort—for instance, the share of the 25- to 29-year-olds in 1990—is instrumented using the respective share of this cohort in the previous period—that is, using the share of the cohort of 20- to 24-year-olds in 1985. Additional identification is obtained by exploiting that the shares of the youngest cohorts of working-age adults are instrumented by shares of cohorts that were not even in the labor force, and by using cohorts that have already left the labor force when the outcome variables are realized. Panels (a) and (b) of Figure A.3 in the Appendix illustrate this identification strategy.[14]

Columns (5) of Tables 2.1 and 2.2 present the second stage results from two-stage least squares (2SLS) estimations using this instrumentation strategy for the age shares in the corresponding specifications (in levels or differences). The first stage is sufficiently strong, as indicated by the respective F-statistics. The coefficients in the outcome equation closely resemble those obtained from standard panel estimation techniques, however. In particular, the overall patterns remain unchanged.

A similar identification strategy can be applied to the share of high-skilled individuals (or its change). The logic here is that the formal (tertiary or vocational) education attained by a given cohort is unlikely to change over the range of five or ten years. At the same time, using the lagged shares provides variation that is less likely to be affected by (or correlated with) contemporaneous macroeconomic performance, conditional on the full set of controls. The variation in the share of high-skilled over the course of five or ten years primarily depends on the education of young individuals entering the labor force and of the old individuals leaving the labor force. We therefore use the lagged skill shares

[14]Moreover, the figure illustrates that this IV approach can be applied regardless of whether the data are coded in five-year or than ten-year cohorts.

in these respective age cohorts as instruments. The logic of this identification approach is illustrated in Panels (c) and (d) of Figure A.3.[15] Columns (6) of Tables 2.1 and 2.2 present the corresponding results for the second stage of the 2SLS framework applied to human capital. Again, the first stage is strong as expected. The estimates for the outcome equation are qualitatively identical and quantitatively somewhat larger for human capital than those obtained with the baseline estimation approach, whereas the coefficient estimates for the age shares are somewhat smaller.[16] Again, this suggests that endogeneity bias appears not to be a serious concern for the qualitative patterns and, if anything, biases the coefficients of the share high-skilled toward zero in the conventional estimates.

Columns (7) of Tables 2.1 and 2.2 present the corresponding results for the second stage of the 2SLS framework applied to both the age structure and human capital. Again, the overall pattern is very similar.

2.4.5 Robustness and Further Results

We conducted several additional checks to investigate the robustness of these results.

A first robustness check concerns the estimation of the same model using an alternative data set for the age composition and human capital endowment provided by Barro and Lee (2013). The corresponding results reveal very similar estimation results.[17]

A second robustness check concerns the possibility of overfitting and multicollinearity by using data at the level of five-year age cohorts. Estimation results obtained with data for ten-year age cohorts deliver qualitatively and quantitatively similar results for the effects of the age structure of the population, and even slightly larger coefficient estimates for human capital.[18] The same holds when restricting the specification to only four or three age cohorts.[19]

As third robustness check, we considered an alternative coding of the human capital composition by considering the average years of schooling, while allowing for a more flexible specification. The results confirm the earlier findings regarding both, the age profile as well as the relevance of the share of high-skilled individuals with at least secondary education.[20]

[15]Notice that this additional instrumentation strategy provides the possibility of conducting overidentification tests, as there are more instruments than instrumented variables.

[16]One potential caveat with the instrumentation approach of human capital could be that the education composition of the share of the young entering the labor force might reflect anticipated macroeconomic performance, and thus pose a potential problem of endogeneity, while this is unlikely for the cohort leaving the labor force. In additional robustness checks, we therefore applied an alternative identification that only uses the education composition of the cohort that exits the labor force, paralleling the approach popularized by Acemoglu and Johnson (2007) by implicitly assuming that all newly entering cohorts are high-skilled. The results are qualitatively similar, see Table A.4 in the Appendix.

[17]See Tables A.5 and A.6 in the Appendix for detailed estimation results and Figure A.4 for the corresponding estimated age-productivity profiles.

[18]Detailed results are reported in Tables A.7 and A.8 in the Appendix. Panels (c) and (d) of Figure 2.3 show the corresponding productivity profiles obtained with ten-year panel data.

[19]See Figure A.5 in the Appendix.

[20]Detailed results can be found in Table A.9 in the Appendix.

Another robustness check refers to the use of income (output) per capita instead of output per worker as variable of main interest. While output per worker captures the notion of productivity and macroeconomic performance from the perspective of the production process, income per capita might be seen as more relevant from a policy perspective. The results are essentially the same for income per capita.[21]

Results obtained with an extended specification that also accounts for the age-related change in skills by incorporating cohort-specific information on the share of skilled individuals are also similar. In particular, these estimates only provide weak evidence for differences in the effect of the share of high-skilled individuals across different age cohorts when controlling for cohort-specific skill shares. Moreover, the qualitative and quantitative results for the effects of the demographic age structure as such remain largely unaffected.[22]

To some extent, these findings shed new light on the results of Cuaresma, Lutz, and Sanderson (2014), who conclude, based on an analysis that uses a comparable data set, that the demographic dividend is mostly the byproduct of increases in education. Whereas Cuaresma, Lutz, and Sanderson (2014) do not specifically control for the cohort-based demographic structure, but instead for labor force participation and the relative size of the working-age population to total population (that is, the inverse of the dependency ratio), the findings here suggest that the age structure might have an independent effect from the human capital endowment.[23]

Instead of considering the age composition of the workforce, the previous literature has focused on the old-age dependency ratio. Additional results suggest that adding the dependency ratio as well as the size of the working-age population (in logs or in absolute numbers) as further control variables leaves the results essentially unaffected.[24] This suggests that the role of aging for macroeconomic performance does not predominantly work through population size or the share of elderly, but through the age composition of the workforce. Consequently, a main economic implication of low fertility in the aftermath of the demographic transition appears to be population aging rather than a shrinking (or reduced growth) of the population at large. This issue will be discussed in more detail in the simulations below.

[21] Notice that the estimation equation for income per capita can also be directly derived from the conceptual framework. This requires defining the sizes of the age groups as shares of the total population (rather than of the working-age population) and controlling for the (young- and old-age) dependency ratio. The respective results are displayed in Table A.10 in the Appendix. The corresponding productivity profiles are displayed in Figure A.6 in the Appendix.

[22] See the results in Tables A.11 and A.12 in the Appendix for details.

[23] Moreover, the analysis of Cuaresma, Lutz, and Sanderson (2014) accounts for variation in labor force participation rates, which might be driven partly by cyclical phenomena instead of long-run trends, thus imposing problems for identification. As indicated before, the effects of the demographic structure presented here correspond to intention-to-treat effects, which are likely to provide a lower bound of the actual effect under the assumption of relatively stable participation patterns.

[24] See Appendix Tables A.13, A.14, and A.15 for detailed results.

Controlling for life expectancy also leaves the main results unaffected.[25] Using average years of schooling instead of the population share with a high-skilled education delivers similar results for the role of the age structure, but no significant effect for human capital in terms of average years of schooling.[26] This result potentially reflects the fact that skill shares provide a more appropriate measure of the skill endowment than the use of average years of schooling.[27]

Note that controlling for the dependency ratio, the size of the working-age population, and life expectancy at birth accounts for variation in fertility, health, and longevity across countries and over time, respectively. Controlling for these variables might lead to endogeneity concerns, if economic development unfolds a feedback mechanism on either of these dimensions in the long run and, at the same time, the corresponding variables correlate with the age structure or with education. If selection on observables is informative for selection on unobservables, the stability of parameter estimates across specifications with additional controls suggests that the bias is limited.[28] Hence, it is reassuring that including further control variables does not considerably affect the quantitative and qualitative results compared to our baseline specification.

2.4.6 Education to Counteract the Effects of Aging?

Instead of relying on qualitative assessments of the implications of population aging and education dynamics, the estimation framework and the corresponding estimates also allow to go one step further in the quantification of the increase in education that is needed to offset the effects of population on economic performance. In particular, the framework provides the possibility to estimate an elasticity of substitution between changes in the age structure and changes in the human capital structure of the economy that is needed to keep output per worker constant. An upper bound for this skills-aging elasticity in the levels model is given by

$$\eta^j_{max} = \frac{(1-\widehat{\alpha})\widehat{\lambda}^j}{(1-\widehat{\alpha})\widehat{\lambda}^h} = \frac{\widehat{\lambda}^j}{\widehat{\lambda}^h} < 0. \tag{2.13}$$

In the differences model, the elasticity takes the form

$$\eta^j_{max} = \frac{(1-\widehat{\alpha})\widehat{\lambda}^j}{(1-\widehat{\alpha})\widehat{\lambda}^h + \widehat{\theta}} < 0. \tag{2.14}$$

The corresponding parameters are the structural estimates of the empirical model in (2.10) or (2.12). Because the elasticity depends on the level of schooling in the previous

[25] See Table A.16 in the Appendix for details.

[26] See Table A.17 in the Appendix for details.

[27] See, for example, Hanushek and Woessmann (2012).

[28] Moreover, the coefficient estimates across the different specifications are very similar, while the variation explained is fairly comparable, indicating that selection and endogeneity concerns should be limited, following arguments in the spirit of Altonji, Elder, and Taber (2005) and Oster (2017).

period and changes in the current skill distribution, an increase in the share of skills in the same period can only work through the composition channel (that is, the denominator is $(1-\widehat{\alpha})\widehat{\lambda}^h$ in this case). In the following period, the skills-aging elasticity is given by the expression in (2.13) corrected for additional changes in the distribution of skills, which are again weighted by $(1-\widehat{\alpha})\widehat{\lambda}^h$. Since the denominator is positive and the cohort effects of the demographic structure are negative as long as the most productive cohort is chosen as reference group, the elasticity will always exhibit a negative sign. Hence, the elasticity is largest, when the denominator is maximized. This is the case when the share of high-skilled workers in the population increases over at least two consecutive periods and no human capital is lost due to retirement or emigration in the working-age population. Consequently, η^j cannot be greater than the expression stated in (2.13). In fact, it is lower whenever the gains in human capital are lost to some extent. Thus, η^j_{max} represents an upper bound for the skills-aging elasticity. Moreover, this upper bound has a natural interpretation in that it is the most favorable scenario under which negative feedback from changes in the demographic structure on output can be compensated.

The elasticity can be computed for each age group. For example, suppose an aging society, where a large fraction of the workforce (the baby boomer cohorts) shift out of the most productive group of the 50- to 54-year-olds into the less productive group of the 60- to 64-year-olds. Columns (3) of Tables 2.1 and 2.2 then provide upper bounds for the skills-aging elasticity of $\eta^{60-64}_{max} = -\frac{5.50}{1.08} \approx -5.09$ and $\eta^{60-64}_{max} = -\frac{5.14}{3.30+0.55} \approx -1.34$. Assuming constant returns to schooling, the share of high-skilled workers would thus have to increase by 1.34 to 5.09 percentage points in order to offset a one-percentage point shift out of the cohort 50–54 into the cohort 60–64. Because schooling takes place mostly at a young age, it is, however, unrealistic to increase the human capital of older workers by more than a small extent. Changes in the skill distribution must therefore come mostly through young cohorts. This is particularly problematic if young cohorts are small in size relative to the cohorts approaching retirement, such as in the case of the baby boomer generation. Therefore, even in the presence of large human capital increases, the demographic structure unfolds a forceful effect on macroeconomic performance. This may also be one of the reasons why large-scale extensions of schooling in developing countries in the context of the demographic transition (and the decline in fertility) were not associated by a strong development boost.[29]

However, as discussed in Section 2.4.5, a richer specification that considers the education composition of different age cohorts of the working-age population (age 15 to 69) delivers little evidence for a strong and systematic role of the skill distribution across age groups for output. The corresponding estimates are insignificant in most cases. Also the Wald test for joint significance of the estimated parameter sets fails to reject that estimates

[29] Another reason may be that schooling quality is generally low. For more information see, for example, Hanushek and Woessmann (2008).

are jointly different from zero in many cases.[30] Hence, we find little evidence for the obsolescence of high-skilled education embodied in older generations.

2.5 Implications for Future Economic Performance

By and large, the estimation results reveal a relevant role of demographic dynamics in terms of aging as well as in terms of changes in the human capital embodied in the working population, for economic development. At the same time, the heterogeneity across subsamples indicates that aging might not affect all countries in the same way. In particular, countries with a relatively old population, and with an ongoing aging process, appear to be affected most by the adverse effects of aging.

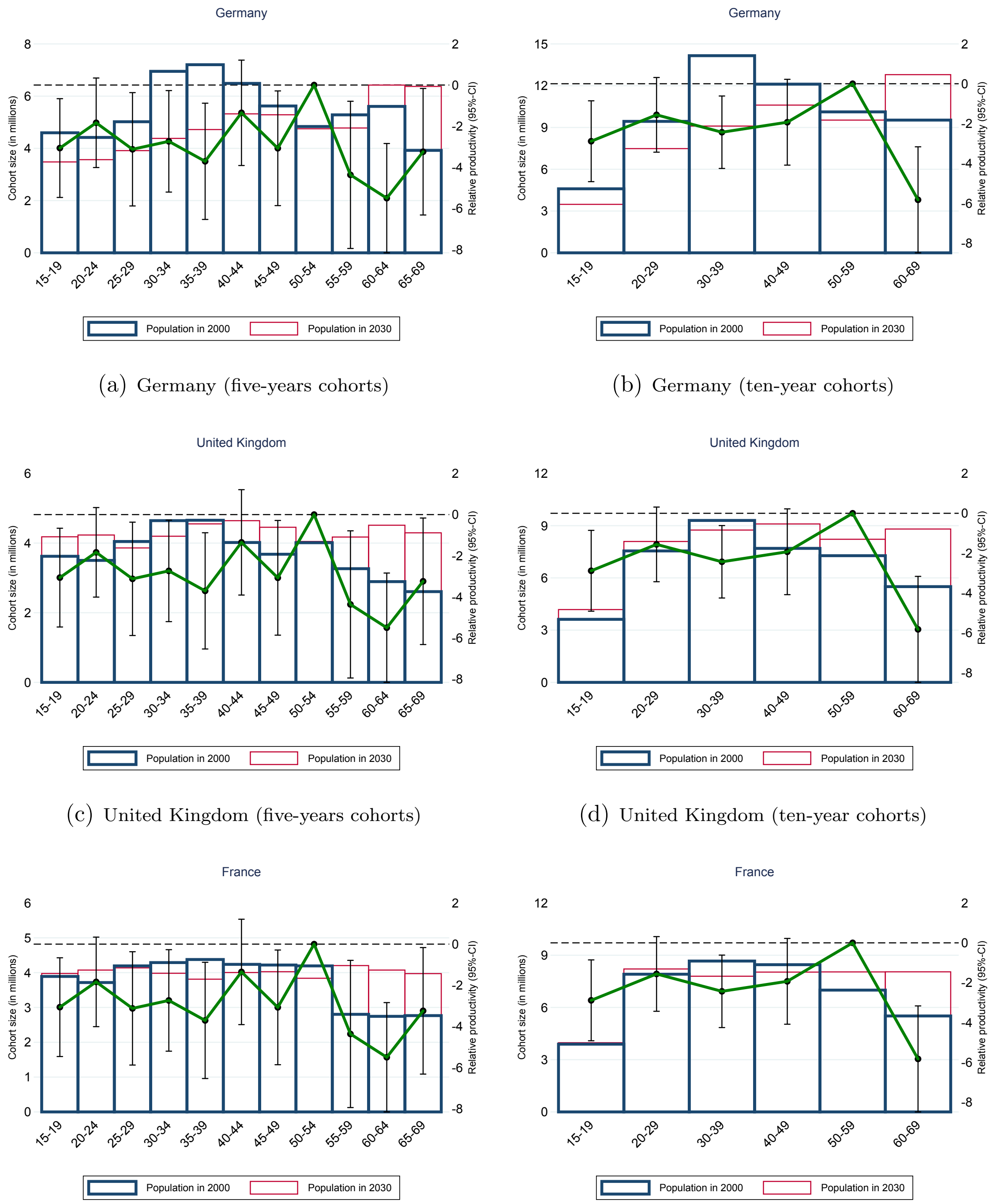
In order to obtain a more coherent picture of these patterns, and gain some understanding of the relative importance of human capital in offsetting the effects of population aging, this section presents the results of simulations of economic performance based on the baseline estimates of the previous section, and several alternative scenarios regarding aging and human capital dynamics.

2.5.1 Projecting the Effects of Aging and Education on Future Performance

While aging appears to be a process that is hard, if not impossible, to influence in the short and medium run, the skill composition of the population is a possible dimension through which policy might try to influence the economic prospects of a country. This raises the question about the relative importance of population aging and changes in the skill composition, and about the likely scenarios faced by countries with different age and skill compositions of their populations.

To illustrate the usefulness of the methodology developed in this paper for addressing these questions, consider Figure 2.4, which contrasts the coefficient estimates for the age structure obtained with specifications for five-year and ten-year age cohorts, with the projected change in the age structure for Germany, the UK, and France. Taking the age-profile of coefficients as a stable world average, the predicted economic performance will differ only because of heterogeneous aging patterns across countries. This is illustrated by the different age structures in Germany, the UK, and France. Similarly, one can use projections of the human capital composition for these countries to compute the predicted performance due to the changes in this dimension.

[30] In order to further test whether there might be an interplay between the demographic structure of the workforce and the distribution of skills, interacted models can be estimated. This allows to test the null hypothesis whether the effect of the demographic structure is stronger (or weaker) the larger the share of high-skilled workers in the population is. However, there is only weak evidence that this is the case (results not shown).



(a) Germany (five-years cohorts)

(b) Germany (ten-year cohorts)

(c) United Kingdom (five-years cohorts)

(d) United Kingdom (ten-year cohorts)

(e) France (five-years cohorts)

(f) France (ten-year cohorts)

Figure 2.4: Macro Productivity Profiles and Demographic Change for Selected Countries

In the following, we use the estimates to conduct counterfactual experiments to infer the relative importance of aging and changes in the skill composition of the population for future economic development. To this end, we use available projections of the prospective age and skill composition of the population. These projections can be conducted under several scenarios. The baseline scenario is to use the estimates to obtain an estimate of output per worker by inserting population projections in terms of age and skill structure and compute output as in equation (2.10) over the period over which projections for age structure and skill composition are available. This scenario uses all information about the evolution of the economy and, thus, provides a best practice projection.

As a consistency check for this methodology, we use the estimates obtained from the sample 1950–2000 to project economic performance until 2010, based on the actual observations for the changes in the age structure and in the skill share between 2000 and 2010. The results suggest that the model is able to track development rather well, with the exception that unpredictable events such as the global financial crisis of 2007/2008 imply deviations of the model projections from the actual data.[31]

As a first alternative scenario, we use only the projection for the human capital structure, but keep the age composition of the population constant. In other words, this corresponds to a simulation that stops the aging process and keeps the population at its current status quo in terms of age composition. Conceptually, this corresponds to the (deliberately extreme and unrealistic) counterfactual assumption of a stable population ("constant demographic structure").

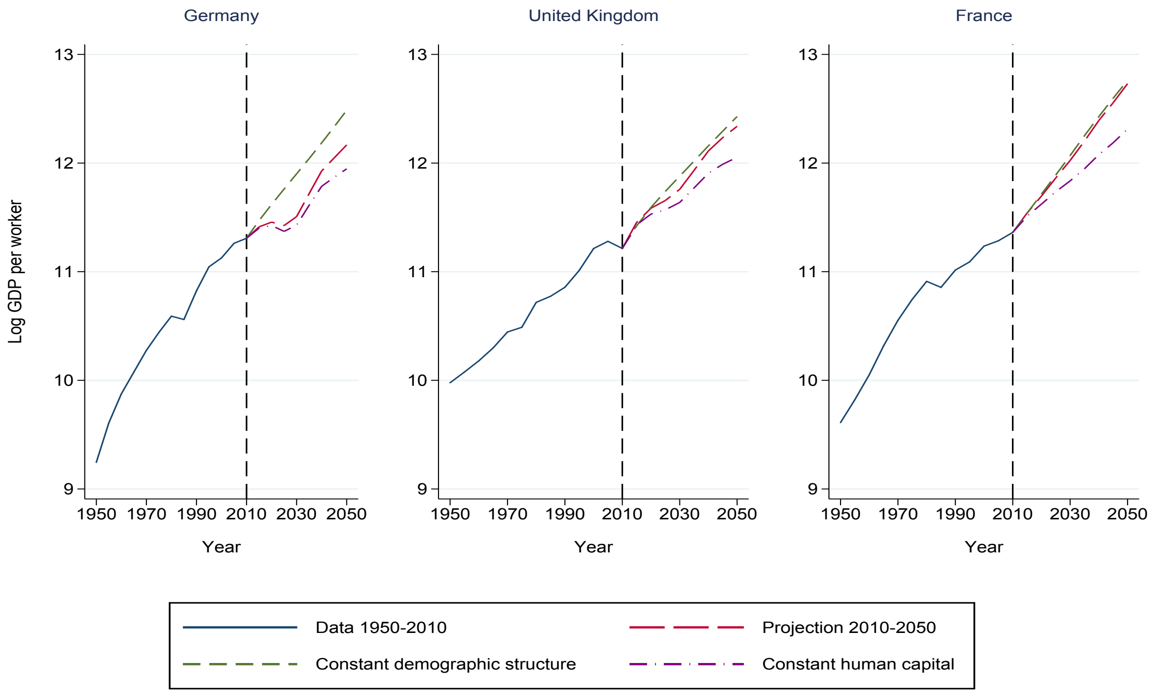
As a second alternative scenario, we simulate the model using the available population projections for the dynamics of the age structure, but keep the composition of human capital in the population constant at the present levels. This corresponds to a scenario that evaluates the consequences of aging in isolation while keeping human capital constant ("constant human capital"). The reference year for both counterfactual exercises is 2010.

In the following, we use Germany, the UK, and France as prime examples of developed economies that differ in terms of the speed of population aging. All three countries have comparable income levels and experienced roughly comparable patterns of economic development in the past. As Figure 2.1 shows, however, the three countries have different age structures of their population and correspondingly face different dynamics of population aging in the future. Moreover, the trajectories of educational attainment differ across these countries as shown in Figure 2.2.[32] Using the same set of coefficient estimates from the empirical analysis for these countries, we can therefore provide comparable simulations that allow us to identify the implications of the projected aging and human capital dynamics for future development.

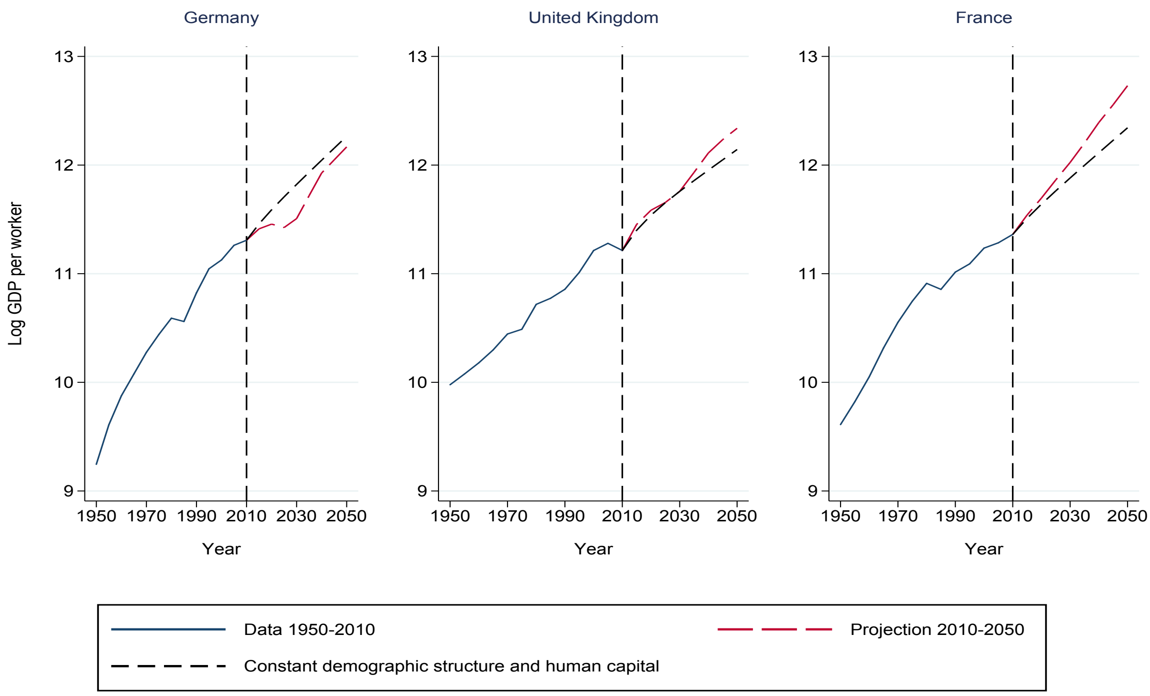
Figure 2.5 presents the corresponding projection results for the different scenarios for Germany, the UK, and France. For all three countries, the simulated performance using the available projections for aging and human capital suggest a dampened economic performance in the decades to come. The predicted slow-down is more pronounced for Germany than for the UK and France. Obviously, these projections are based on strong assumptions and should not be confused with forecasts of output growth, because important

[31] See Figure A.7 in the Appendix, where the economic performance of non-OECD countries is matched well, whereas the projection of economic performance is too benevolent for OECD countries, which have been affected more by the global financial crisis up to 2010.

[32] The same applies for education projections, see Figure A.8 in the Appendix.



(a) Selected Countries: Germany, United Kingdom, and France



(b) Aging vs. Human Capital Accumulation

Figure 2.5: Projections Under Different Scenarios

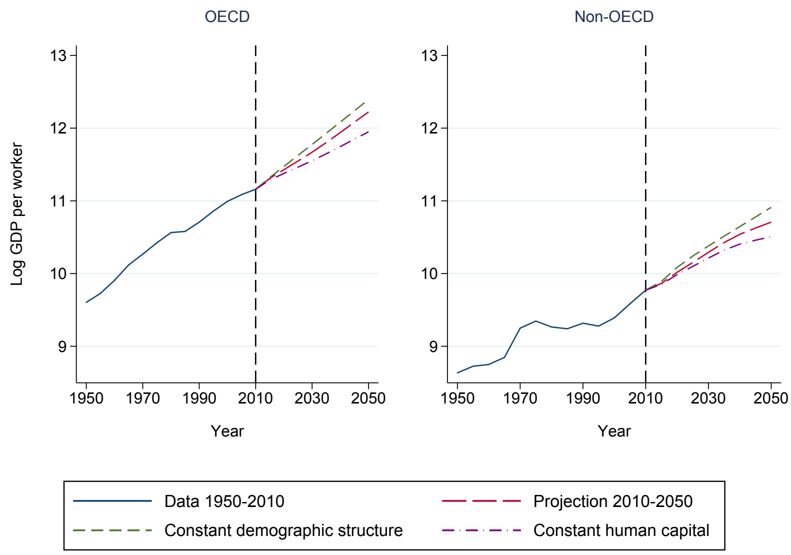
components like capital accumulation, depreciation, etc., are not adequately modeled in these simulations but held constant at their 2010 levels.[33] Nevertheless, they are useful as a benchmark for comparing the projections to the counterfactual simulations that freeze the demographic structure or the human capital distribution at their respective current shapes. When considering a constant age structure ("constant demographic structure"), the projection of the economic performance in the three countries is more positive than in the baseline projection, implying a negative effect of population aging. However, the difference between the best projection and this counterfactual projection is substantially more different in the case of Germany, which faces more pronounced population aging than the UK and France. Alternatively, keeping the human capital structure unchanged ("constant human capital") implies a moderately dampened economic development in Germany, whereas the development in the UK and France is affected more negatively by this scenario. While in Germany, freezing the skill share at its current level has relatively minor implications, in particular in the near future, a continued upskilling of the population in the UK and France seems to be a major factor for future development. Taken together, the results predict that in Germany population aging is a powerful dampening force for economic performance, which is likely to unfold its effects in the future, whereas the effect of changes in the education composition have rather limited power, because the population is already very skilled and young cohorts are small in size. In contrast, in the UK, and even more so in France, aging poses less of a problem, whereas a failure to keep pace with the projected education attainment might impose substantial negative effects on development. Panel (b) of Figure 2.5 illustrates this relative difference between aging and education: The figure contrasts the predicted dynamics of output per worker using the projections for demography and human capital to the counterfactual with both distributions frozen at their current shapes: Germany is predicted to exhibit a lower performance than under the counterfactual status quo, whereas the UK and France will develop faster. Part of this is due to the greater leverage for human capital implied by the different demographic structure. Greater aging pressure also limits the scope for human capital, which is embodied in the small young cohorts, to compensate direct effects of aging. Similarly heterogeneous results obtain for other developed economies, such as the USA and Japan.[34]

One noteworthy aspect in this context is that specifications without a control for the lagged development deliver somewhat larger effects for human capital. Correspondingly, projections obtained with these specifications deliver a somewhat more benign scenario, with greater scope for human capital in compensating the effects of population aging for countries that face substantial population aging but have a rather skilled population.[35]

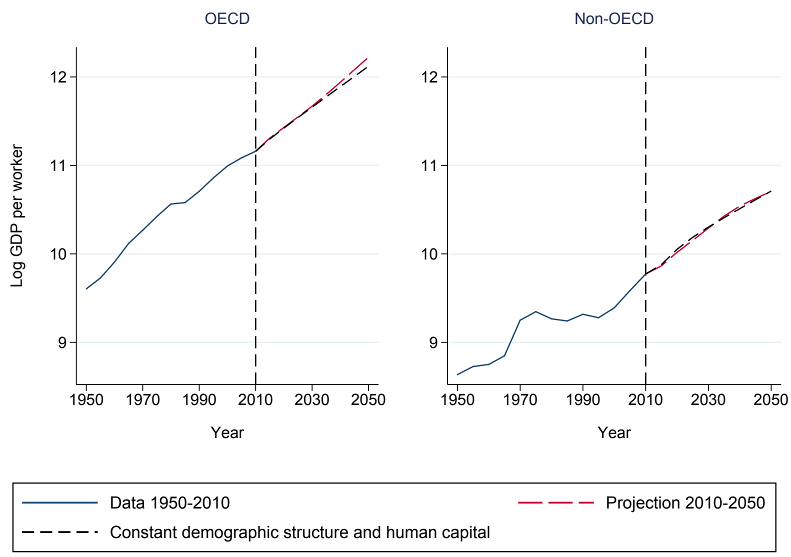
[33]Below, we turn to alternative simulation scenarios that also incorporate capital projections and alternative scenarios for education attainment.

[34]See Figure A.9 in the Appendix.

[35]See Figure A.10 in the Appendix for details.



(a) Developed vs. Developing Economies



(b) Aging vs. Human Capital Accumulation

Figure 2.6: Projections Under Different Scenarios

Figure 2.6 presents the corresponding results for OECD and non-OECD countries. Again, the benchmark projections deliver a rather pessimistic outlook about economic performance in both samples. Freezing the age structure at its current level implies faster development in the OECD countries, suggesting that aging will be a major impediment for economic performance in the future. A potentially more surprising result is that aging appears to have a similar negative effect on economic development in non-OECD countries, as evidenced by the simulation that keeps the age structure constant ("constant demographic structure"). The positive trajectory is mainly due to improvements in the skill composition of the population. Conducting the alternative scenario with constant skill structure but incorporating the demographic aging process reveals a worse performance for OECD and non-OECD countries compared to the baseline scenario.[36] This result highlights the importance of increasing human capital in the process of the demographic transition and the corresponding aging of the population in developing countries. The results are striking in showing the potential of human capital to counteract negative implications of aging, particularly when there is substantial scope for improvements in the education attainment of the population, as in less developed countries.

In countries with high fertility rates and a relatively young population, as it is the case in many African countries, population aging might even exert a positive effect: Because aging allows for the acquisition of more experience, it implies greater incentives to acquire more formal human capital. The key impediment for development in these countries appears to be a slow-down in the accumulation of human capital. Simulations illustrate this for Niger, Nigeria, Uganda and Mali.[37]

2.5.2 Sensitivity and Alternative Scenarios

The projection results are robust to changes in the specification of the estimation equation or the use of income per capita rather than income per worker.[38] Moreover, if based on the empirical results of the instrumental variables approach (Table 2.1, Column 7), the projection results remain qualitatively unchanged with education showing a slightly more promising quantitative role for future development.[39]

Obviously the quantitative dimension of the projections obtained with the different simulation scenarios are subject to a number of potentially restrictive and unrealistic assumptions. First, the alternative scenarios, in particular the assumption of a constant age structure as if the population were stable in its current form ("constant demographic structure"), were deliberately extreme. It should be clear that such a scenario is not only unrealistic but also inconsistent in terms of the implied demographic dynamics of non-stable populations. It should therefore be seen as an illustrative thought experiment, rather than a realistic (or even in some way implementable) possibility with normative character. A less extreme way to construct counterfactual scenarios in this context is to make alternative assumptions about vital rates. Such scenarios would fix fertility, or mortality, or both, at their 2010 levels—instead of fixing the age shares—and then

[36]Figure A.8 in the Appendix illustrates the education projections that are neglected in this scenario. Alternative scenarios comprise constant enrollment rates, see the discussion below.

[37]See Figure A.11 in the Appendix.

[38]Figures A.12, A.13, and A.14 in the Appendix present the corresponding projections for a coarser specification of age groups as well as for income per capita instead for selected countries and county groups. Moreover, Figures A.15 and A.16 show the corresponding projections when accounting for the size of the working-age population.

[39]Figure A.17 in the Appendix depicts the corresponding projections for Germany, France, and OECD and non-OECD countries.

compare the implied aging dynamics with the actually projected ones. The problem with these scenarios is that the resulting patterns of population aging differ only very mildly compared to the available projections.[40] This reflects the well-known difficulty for any policy aiming at vital rates to change the momentum of population aging. On the contrary, such scenarios camouflage the true extent of the consequences of aging for economic performance, as they imply differences in the age structure that are too minor to provide substantially new insights. We therefore view the alternative scenarios studied here as more illustrative.

Second, the simulations are based on the assumption of physical capital following the average growth trend over the period 1950–2010. This assumption is clearly counterfactual, but it allows us to focus on the demographic aspects while remaining agnostic about the implications for savings and capital accumulation. An alternative is to specify an auxiliary equation for the accumulation of physical capital per worker as a function of past output, past levels of physical and human capital, and the age structure of the population, along the line of the estimation framework for output. Such an auxiliary equation can be estimated and the coefficient estimates can be used to project physical capital under alternative scenarios, and in a second step output. The results from such a refined methodology leave the main results, in particular related to the relative importance of aging and human capital dynamics in different countries unaffected.[41]

Third, the simulations are obtained under the assumption of stable coefficients over time and across countries, as well as a constant growth trend. These assumptions allow for a comparable simulation across all countries, thereby providing the possibility to identify differences in economic performance that are due to projections in the demographic domain (the age structure) and projections in the domain of human capital. This delivers qualitative results that are internally consistent. To investigate the robustness of the findings with respect to less restrictive assumptions about parameter stability, we also conducted the same counterfactual projections based on estimation results obtained for the sample period 1990–2010 instead of the entire sample period 1950–2010. The results of these projections are qualitatively very similar.[42]

Overall, these projection results suggest that the implications of population aging on economic performance differ across countries with different age compositions and human capital projections. We therefore proceed by investigating this issue in more detail.

[40] This result is illustrated in Figure A.18 in the Appendix.

[41] Figure A.19 in the Appendix compares the projection scenarios for exogenous and endogenous capital. The simulation results suggest that capital accumulation is reduced in developed countries as consequence of aging, whereas the differences are less pronounced in low-income countries. This is consistent with the results from computable general equilibrium models that predict a negative effect of aging populations on capital formation, see, for example, Sánchez-Romero (2013). However, these estimates are based on the implicit and counterfactual assumption of closed economies without access to international capital markets. See Börsch-Supan, Ludwig, and Winter (2006) and Domeij and Floden (2006) for quantitative and empirical studies on the implications of population aging for international capital flows.

[42] See Figures A.20 and A.21 in the Appendix.

2.5.3 The Role of the Contemporaneous Age Composition

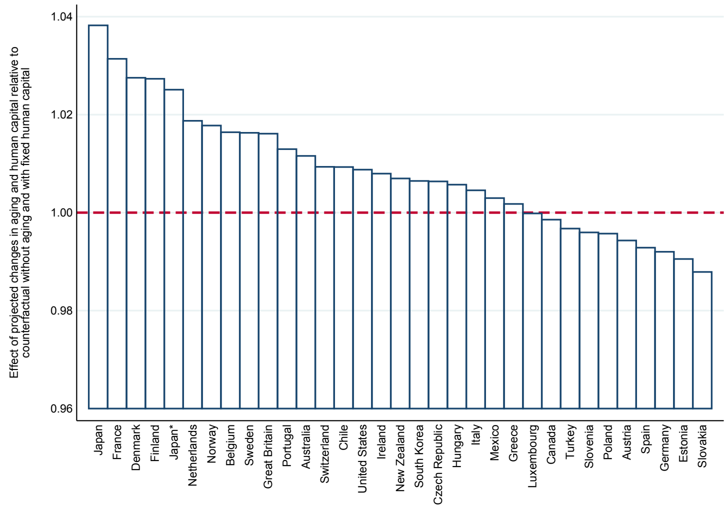
After having established how population aging and the contemporaneous changes in aggregate human capital affect macroeconomic performance and how this influences the prospects of future economic development, we return to the question whether investment in education can potentially offset the effects of population aging. Using the simulation methodology, we focus on which countries are predicted to suffer most from population aging, measured in terms of economic performance. We then contrast the positive impact of human capital acquisition to the predicted negative effect of aging, allowing us to quantify whether the former is large enough to offset the latter.

Using similar counterfactual experiments as in the previous section, one can simulate the effect of projected changes in the age and human capital structure of a country and compare it to a counterfactual scenario where both the age and human capital structures are fixed at their current (2010) levels. Figure 2.7 provides a plot of the predicted performance of OECD and non-OECD countries relative to the counterfactual status quo. Countries with a young population that are projected to further increase their share of skilled individuals, such as France, are predicted to exhibit a substantially better performance than they would absent population aging and continued skill acquisition. On the other end of the spectrum, countries like Germany but also Austria, Spain, Estonia, or Slovakia, that face substantial population aging with populations that do not have much scope for further upskilling, are likely to do worse than under the counterfactual status quo. Similar patterns can be observed for non-OECD countries.

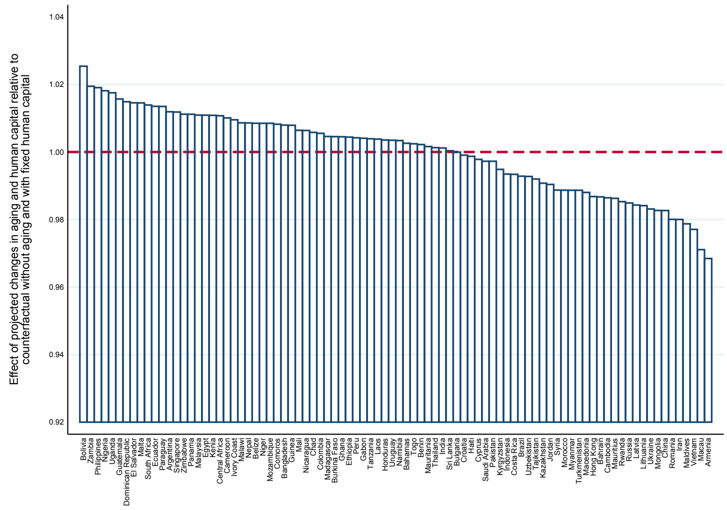
These predicted effects are mainly due to the differences in the scope for further acquisition of human capital at the aggregate level. This is illustrated by a decomposition, in which the contribution of the projected aging of the population to predicted economic performance is isolated from the contribution of the projected change in human capital.[43]

Returning to the estimation framework, additional results confirm these insights by revealing heterogeneity in the effects when splitting the sample into subsamples with different levels of economic or demographic development. In particular, the results reveal stronger effects of changes in the age structure on economic performance in non-OECD than in OECD countries.[44] This finding could indicate that the aging process is less pronounced once the demographic transition is completed. Alternatively, this finding could indicate that the adoption of technologies allows rich economies to insulate themselves from the negative effects of aging on average, as suggested by Acemoglu and Restrepo (2017). In contrast, poorer countries in which the process of population aging sets in with force might experience particularly adverse effects on their economic performance. This finding is also consistent with the finding that population aging has a more pronounced effect on societies with a large share of young people when considering a sample split.

[43]See Figures A.22 and A.23 in the Appendix for the respective graphs.
[44]See Table A.18 in the Appendix for details.



(a) OECD Countries



(b) Non-OECD Countries

Figure 2.7: Projected Performance Relative to the Counterfactual Status Quo

A more direct way of testing this conjecture is by investigating the stability of coefficients for samples split in terms of the observation period, or in terms of the relative age of the population. When considering estimates separately for the period before and after 1990, the results reveal that the importance of population aging appears to have increased in recent decades. Concretely, the effects are stronger when considering a subsample for the period after 1990.[45] When splitting the sample into countries with old populations and countries with young populations (relative to the median of the young-age dependency ratio), one obtains large and negative effects of the age structure in the "old" countries but large and positive effects of the age structure for the group of "young" countries.[46] This

[45]See Tables A.19 and A.20 in the Appendix for details.
[46]See Table A.21 in the Appendix.

is consistent with the patterns obtained from the simulation shown in Figure 2.7. Aging is predominantly a problem for economic development in countries with an unfavorable age composition (such as Germany, and many Eastern European and Asian countries), whereas countries with a young population are standing to gain from the projected aging and education patterns (such as France, or many countries in Africa).[47]

2.5.4 The Scope for Adjustment

We conclude the analysis by investigating several alternative margins besides additional investments in education along which the aging effect might be counteracted. A key dimension in this respect is the intensity with which human capital, and labor in general, is supplied. To explore whether the decline in the relative supply of labor by young cohorts due to population aging could be neutralized by an increased labor force participation, in particular by women, or by longer work weeks, we proceed in three steps. First, we provide an assessment of the scope of adjustment along this dimension by comparing the effective labor supply of the different age groups in 2010 to the projected levels in 2050. The effective labor supply in 2010 is computed by using the same information about the age structure of the workforce as in the estimates and combining it with data about the absolute size of the workforce and information about age-specific labor force participation at the extensive and intensive margin provided by the International Labour Organization (2011).[48] For 2050, we use the IIASA-VID projection for the working-age population and assume that age-specific labor force participation and hours worked by women attain the same level as for men in the 2010 ILO data. This allows us to construct several alternative scenarios compared to the (implicit) assumption of constant female labor force participation and hours worked in the results presented so far. In particular, we compare the baseline projection as before, which assumes constant labor force participation at the extensive and intensive margin, to three alternative scenarios: a projection that also accounts for changes in female labor force participation, for changes in hours, and for changes in both, until 2050. When comparing the magnitudes relative to the projected effects of aging for the size of the workforce of each age group, these increases appear rather moderate. In particular, these projections indicate that the overall patterns of aging will at most be moderated but not completely neutralized.[49]

[47]Interestingly, also countries with a rather old and well educated population, such as Japan, are projected to gain from the future dynamics. This is partly due to the fact that they already underwent substantial shifts in the age distribution in the past and stand to face a "pause" in the coming decades. For this reason, we report also a scenario Japan* where the demographic structure has been fixed in the year 2000 instead of 2010. In other countries, the development is projected to be positive due to considerable immigration that is projected to stabilize the age distribution, as, for example, in Switzerland.

[48]The data are available at `http://www.ilo.org/ilostat`.

[49]This is illustrated in Figures A.24 and A.25 in the Appendix, where we plot the labor supply in terms of the total weekly hours worked by the different age groups as of 2010 and the projected change under the different scenarios until 2050, for Germany and France, and OECD and non-OECD countries, respectively.

In a second step, we make use of novel age projections of labor force participation constructed by Fürnkranz-Prskawetz, Hammer, and Loichinger (2016) for 26 European countries to project the macroeconomic performance of these countries using the baseline estimates obtained from the age structure of the working-age population. In particular, using the 2010 data on labor force participation, we construct an index for the labor force participation that takes a value of 100 for each age group in 2010, and compute the relative change up to 2050 using the projection data by Fürnkranz-Prskawetz, Hammer, and Loichinger (2016). The projection results document that aging remains a substantial impediment for economic development for most countries.[50] A drawback of this analysis is that it relies on estimates that do not incorporate information about labor force participation but that only rely on the age structure of the working-age population. This reduced-form approach helps avoiding endogeneity problems related to labor force participation, but it introduces a slight methodological inconsistency.

In a third step, we therefore re-estimate the model using information on the effective labor force and, hence, effective human capital supply, by incorporating age-specific differences in the intensive margin, for the period 1980–2010 for which respective age-structured data are available from the ILO (International Labour Organization, 2011). Based on these estimates, we then project macroeconomic performance using the workforce projections by Lutz et al. (2007) and by Fürnkranz-Prskawetz, Hammer, and Loichinger (2016). The descriptive statistics indicate that the age structure of the working-age population and of the workforce are rather similar.[51] Consistently, the estimation results do not differ substantially from the baseline estimates and deliver the same qualitative results.[52] Correspondingly, the projections incorporating age-specific labor force participation patterns do not deliver substantially different results. In particular, aging remains to exert a major negative effect on macroeconomic performance, whereas human capital is only able to offset this effect partially.[53] Overall, the results from these three exercises suggest that the incorporation of age-specific projections for labor force participation and hours worked does not greatly affect the conclusions regarding the effects of aging and skills for future macroeconomic performance.

Another margin of adjustment is a potential shift in the age-profile of productivity. Whether in the future the productivity will peak at younger or older ages is an open question. Given the ongoing improvements of health status and labor force attachment of older cohorts in the workforce, and the observation of a stable experience premium that

Notice that the data and projections average over men and women.

[50] See Figure A.26 in the Appendix.

[51] In particular for age groups above 20 years of age, see Table A.1 in the Appendix for details.

[52] Table A.22 in the Appendix contains the corresponding estimation results.

[53] Figure A.27 in the Appendix compares projections obtained with the baseline methodology with projections based on estimates that incorporate variation in labor force participation for the 26 EU countries for which labor force projections are available. Figure A.28 plots the respective projections for Germany and France.

has led to the conjecture of experience-biased technical change (Caselli, 2015), one might expect that the most productive age range might shift from ages 50–54 to older ages.[54] To account for this possibility, we replicated the projections by shifting the estimated productivity profile by one age group (that is, considering ages 55–59 years as the most productive group instead of ages 50–54 years). The corresponding results reveal a modified projection of the consequences of population aging compared to the baseline in the sense that the negative effects of population aging in countries like Germany are delayed.[55]

Finally, the methodology allows us to address the question regarding the scope for technical change or productivity improvements to offset the effects of aging. Recent work has suggested that directed technical change might provide a countervailing force to the negative growth effects of population aging (Acemoglu and Restrepo, 2017). Likewise, improvements in the quality of human capital have been shown to affect macroeconomic performance across states (Hanushek, Ruhose, and Woessmann, 2017). To investigate this issue, we conduct another counterfactual exercise and compute the extent of skill-biased technical change or quality improvement, in the form of an increase in the relative productivity of high-skilled to low-skilled workers λ^h that is needed to offset the effect of population aging until 2050. The results of this exercise replicate the earlier findings about which countries are expected to suffer or gain from population aging, but this time the estimates provide a quantitative interpretation in terms of productivity. In particular, the relative productivity of high-skilled workers would have to increase by more than two-fold to counteract the effects of population aging in countries that are affected negatively by population aging, such as Germany.[56]

2.6 Conclusion

This study presents novel evidence regarding the role of the demographic structure of the workforce and the distribution of skills for aggregate economic performance. On the basis of an extended development accounting model, we derive a flexible empirical framework that can accommodate empirical models previously used in the literature. In particular, assuming that the quality of the labor force depends on the demographic structure allows incorporating workforce demographics into the production function and provides a coherent framework to evaluate the implications of population aging and education dynamics for

[54]For example, work by Kotschy (2018) described in Chapter 4 shows for the United States that over time life-cycle earnings profiles became slightly flatter at higher ages, thus potentially confining some of the negative effects of aging; see Chapter 4. Complementing this, research productivity as measured by scientific breakthroughs has shifted to older ages (Jones and Weinberg, 2011). However, there is also evidence that suggests that new technologies such as ICT might shift the productivity peak to younger years (Falck, Heimisch, and Wiederhold, 2016). In the present context, a shift to younger years would reinforce the effects of aging.

[55]See Figures A.29, A.30, and A.31 in the Appendix for details.

[56]See Figure A.32 in the Appendix for details.

future economic development.

The estimation results show that changes in the age structure of the working-age population have a strong effect on output, even when controlling for human capital. At the same time, the evidence suggests that the stock of human capital embodied in the population has a positive effect on economic performance, conditional on the age structure of the population. The effects of aging in terms of changing relative sizes of the different age cohorts mirror productivity profiles that have been found earlier in terms of hump-shaped productivity patterns over the age dimension. Consequently, the results show that population aging in old societies reduces the future growth potential. The estimates suggest that human capital can help to compensate for these aging pressures and deliver an upper bound for the elasticity between the age structure and the distribution of skills. This elasticity allows gauging the change in the distribution of skills that is required to offset the negative effects of aging of the workforce. The quantitative estimates of this elasticity predict that shifts out of the most productive age cohort into older and less productive age groups can be offset by higher investment in schooling. However, these offsetting effects might not be sufficient to fully compensate for aging, particularly in developed countries. Nevertheless, the results suggest that a continued expansion of education is crucial for future macroeconomic performance.

The results are also useful to infer the relative importance of aging and human capital accumulation for macroeconomic performance by ways of projections on the basis of different scenarios of population aging and human capital dynamics. Projections of future economic development predict that aging will play an important role by slowing down economic development in developed and less developed countries. Aging is, hence, not a problem of the developed world only. There is substantial heterogeneity in the projected macroeconomic performance as result of differential population aging patterns across countries. This heterogeneity emerges through the heterogeneous productivity but also as consequence of the implications for the scope of human capital in compensating this effect. The projections reveal a central role of human capital in ameliorating the negative consequences of aging. This is particularly the case in countries that are yet underdeveloped in terms of human capital endowments and that have considerable potential for an increase in the human capital endowment of the still largely low-skilled population. The scope of human capital improvements for compensating the consequences of population aging appears more limited in economies that age faster. However, the findings make clear that without further improvements in the skill composition of the workforce in these countries, the consequences of population aging will be much more dramatic. Additional projections suggest that increased female labor force participation or longer work hours will be unlikely to neutralize these effects or replace human capital. Moreover, skill-biased technical change will have to be substantial to counteract these developments.

Overall, the results are consistent with an important role of long-run demographic dynamics for future economic development, pointing toward the possibility of more stagnant development in the future. In this sense, the results complement recent findings by Cervellati, Sunde, and Zimmermann (2017).

Chapter 3

Health and Economic Growth: Reconciling the Micro and Macro Evidence

3.1 Introduction

Health is an essential component of human capital, supporting worker productivity by enhancing physical capacity and mental capabilities. Health improvements influence the pace of income growth through many pathways: Better health directly increases labor market participation and worker productivity (Strauss and Thomas, 1998; Bloom and Canning, 2000; Schultz, 2002; Bloom, Canning, and Graham, 2003); increasing life expectancy creates incentives to invest in education, innovation, and physical capital (Bloom, Canning, and Sevilla, 2003; Bloom et al., 2007; Bloom, Canning, and Moore, 2014; Cervellati and Sunde, 2013; Prettner, 2013); and better health, particularly of women, reduces fertility and spurs an economic transition from a state of stagnating incomes toward sustained income growth (Galor and Weil, 2000; Galor, 2005, 2011; Cervellati and Sunde, 2005, 2011; Bloom, Kuhn, and Prettner, 2015). Hence, health and development correlate positively at the macroeconomic level, as the unconditional correlation between the survival probability from age 15 to 60 and (log) gross domestic product (GDP) per worker illustrates in Figure 3.1.

In general, two prominent methods are used to assess the effect of health on economic growth. The first aggregates the results of Mincerian wage regressions of the return on individual health to derive the macroeconomic effects of population health. The second relies on the estimation of a generalized aggregate production function that decomposes human capital into its components, including population health. While the overwhelming majority of studies based on both methods indicate a positive effect of health on economic growth, the size of the effect remains subject to intense debate. The micro-based approach

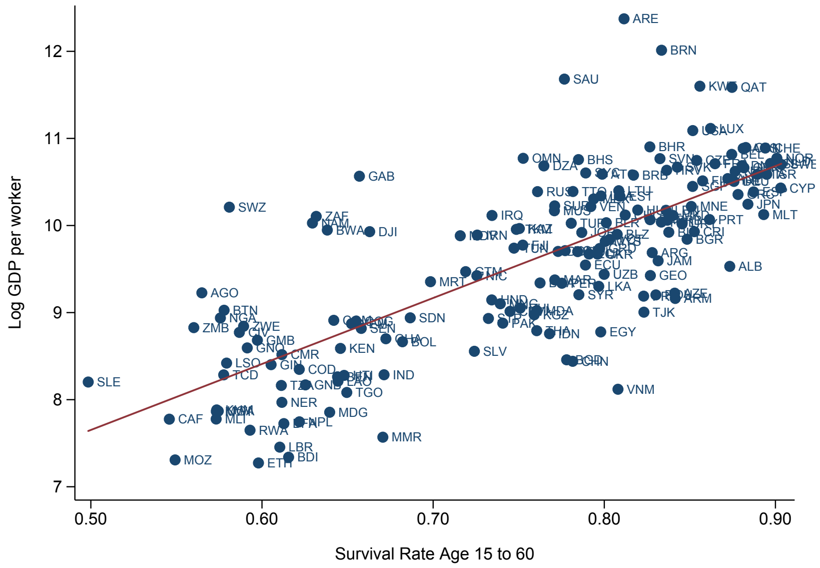


Figure 3.1: Unconditional Correlation: Health and Development

Data sources: Feenstra, Inklaar, and Timmer (2015) and United Nations (2017).

tends to find smaller effects than the macro-based approach, thus presenting a micro-macro puzzle of the economic return to health.

This paper aims to reconcile both approaches by showing that the estimates based on microeconomic results are compatible with the effects derived from a well-specified macroeconomic analysis. To this end, we develop a production function model of economic growth, keeping our specification as close as possible to a generalized Mincerian wage equation as in Weil (2007). This permits us to compare our macro-level estimates and his micro-level calibration directly. We account for reverse causality, omitted variable bias, and measurement errors in the explanatory variables using a large cross-country panel and exploiting the demographic structure for an instrumental variables approach (see the work by Kotschy and Sunde, 2018, described in Chapter 2, which uses this procedure to examine the implications of population aging and educational investment for macroeconomic performance). Our results show that the micro-based and macro-based estimates of the effects of health on economic development are consistent with each other. Thus, we provide a macro-based justification for using the micro-based approach to estimate the direct economic benefits of specific health interventions.

According to Weil (2007), a 10-percentage-point increase in adult survival rates translates into a 6.7-percent increase in labor productivity. Consequently, health differentials account for about 9.9 percent of the variation in output per worker across countries. Our analysis shows that a 10-percentage-point increase in adult survival rates is associated with a 9.1-percent increase in labor productivity. Weil's (2007) estimate falls well within the 95-percent confidence interval of our estimate, suggesting that the two models' results are compatible with each other. Accordingly, the micro-based and macro-based approaches to estimating the effect of health on income growth are reconcilable. Because we include physical capital and education in our empirical framework, the resulting estimate is a measure of the direct productivity benefits of health as in Weil (2007). Thus, the estimated effect excludes the role of better health in increasing the incentives for investment, saving, and education, and its role in reducing fertility and spurring a takeoff toward sustained growth. As such, the productivity benefits of health presented in this paper should be considered conservative.

Our results suggest that public health measures might be a lever for fostering economic development. These types of investments could include vaccination programs, antibiotic distribution programs, and iodine supplementation schemes, which lead to large improvements in health outcomes for relatively low expenditures (World Bank, 1993; Commission on Macroeconomics and Health, 2001; Field, Robles, and Torero, 2009; Luca et al., 2014).

The remainder of this paper is structured as follows. Section 3.2 reviews various approaches to measure the effect of health on economic performance. In Section 3.3, we derive the theoretical effect of health on output per worker from a human capital–augmented aggregate production function. In Section 3.4, we use these results to derive an empirical specification for estimating the influence of changes in the health stock of the population on output growth. Section 3.5 describes the data, while Section 3.6 presents the empirical results. Finally, Section 3.7 concludes.

3.2 Literature Review

A common early empirical approach to examining the effect of health on economic growth involves regressing income per capita growth against initial level of health for a cross-sectional sample of countries, controlling for initial income and other factors believed to influence steady-state income (see, for example, Barro, 1991, 1997; Durlauf, Johnson, and Temple, 2005). Nearly all studies investigating economic growth that use this approach find a positive, significant, and sizable influence of initial population health on the subsequent pace of economic growth (see, for example, Barro and Sala-i-Martin, 2004; Easterly and Levine, 1997; Sachs et al., 1995; Sachs and Warner, 1997; Bhargava et al., 2001; Bloom, Canning, and Sevilla, 2004). While the results of empirical growth equations for most other explanatory variables are not robust with respect to different specifications, Levine and Renelt (1992), Sala-i-Martin (1997), and Doppelhofer, Miller, and Sala-i-Martin (2004) find that initial population health (for example, as measured by life expectancy) is positively associated with subsequent growth in almost all permutations of explanatory variables they analyze. Hence, initial population health is one of the most robust predictors of subsequent economic growth.

More recent work analyzes the effects of health on economic growth via dynamic panel data regressions in the vein of Islam (1995), using the lagged dependent variable as one of the regressors to control for the convergence processes.[1] These studies typically employ an exogenous instrument for health to isolate the causal channel running from better health to income growth (see, for example, Lorentzen, McMillan, and Wacziarg, 2008; Aghion, Howitt, and Murtin, 2011; Cervellati and Sunde, 2011; Bloom, Canning, and Fink, 2014). Acemoglu and Johnson (2007) is one of the few studies finding no evidence for a causal positive effect of health improvements on economic growth. The authors argue that increasing life expectancy raises population growth, which, in turn, increases capital dilution in the neoclassical growth model and, therefore, reduces income growth during the convergence process. They support this theory empirically using the global epidemiological revolution as an instrument for life expectancy. Aghion, Howitt, and Murtin (2011) and Bloom, Canning, and Fink (2014), however, show that this result fails to hold when initial life expectancy is included in the regression. In addition, Cervellati and Sunde (2011) argue that Acemoglu and Johnson's (2007) results only hold for less-developed countries that have not yet undergone the demographic transition. In these countries, increasing life expectancy indeed raises population growth and reduces income growth. For post-demographic transition countries, however, fertility declines with mortality. As a result, health improvements do not lead to an increase in population growth and capital dilution does not intensify. Splitting the sample into pre- and post-demographic transition countries, Cervellati and Sunde (2011) find that the effect of life expectancy on growth is positive for post-demographic transition countries and negative but insignificant for pre-demographic transition countries (see Hansen and Lønstrup, 2015, for a recent discussion).

Another way to assess the size of the macroeconomic effect is by aggregating the microeconomic effects of health to infer the implications for aggregate income. For example, Fogel (1994, 1997) argues that much of British economic growth during 1780–1980 (about 0.33 percent per year) was due to increases in effective labor inputs that resulted from workers' better nutrition and improved health. More recently, the seminal works of Shastry and Weil (2003) and Weil (2007) employ an aggregate production function, in which the effects of health on productivity are calibrated from microeconomic wage regressions. In the microeconomic regressions, these studies explain income by means of various measures for health such as anemia, height, age at menarche, and the adult survival rate. The results of Shastry and Weil (2003) and Weil (2007) suggest that health is an important form of human capital, but that its effect on growth is smaller than that derived from macroeconomic cross-country regressions.

[1] By construction, the fixed effects in such regressions correlate with the error term. Thus, generalized method of moments (GMM) techniques are usually employed for estimation (see, for example, Arellano and Bond, 1991; Blundell and Bond, 1998; Judson and Owen, 1999).

3.3 Theoretical Framework

Assume that time $t = 1, 2 \ldots$ evolves discretely, and consider an aggregate production function of the form

$$Y_t = A_t K_t^{\alpha} H_t^{1-\alpha}, \tag{3.1}$$

where Y_t denotes aggregate output, A_t represents total factor productivity (TFP), K_t is the physical capital stock, H_t describes the aggregate human capital stock, and α constitutes the elasticity of final output with respect to physical capital. The aggregate human capital stock is the sum over the individual human capital levels $\nu_{j,t}$ of workers $j \in \{1, 2, \ldots, \mathcal{J}\}$ in the economy; that is, $H_t = \sum_j^J \nu_{j,t}$. Expressing output in per worker units yields

$$y_t = A_t k_t^{\alpha} \nu_t^{1-\alpha} \tag{3.2}$$

with $y_t = Y_t/L_t$, $k_t = K_t/L_t$, and $\nu_t = H_t/L_t$. Alternatively, output can be expressed in per capita units as

$$\tilde{y}_t = \frac{Y_t}{N_t} = \frac{L_t}{N_t} A_t k_t^{\alpha} \nu_t^{1-\alpha}, \tag{3.3}$$

where N_t refers to the total population size.

In a competitive labor market, one unit of composite labor ν_t earns the wage w_t, which is equal to its marginal product:[2]

$$w_t = \frac{\partial y_t}{\partial \nu_t} = (1-\alpha)\frac{y_t}{\nu_t}. \tag{3.4}$$

Furthermore, we assume individual human capital follows a generalized Mincerian wage equation along the lines of Hall and Jones (1999), Bils and Klenow (2000), and Weil (2007). Hence, we assume individual human capital $\nu_{j,t}$ follows the exponential function

$$\nu_{j,t} = \exp\left(\phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2\right), \tag{3.5}$$

where $h_{j,t}$ denotes the state of health, $s_{j,t}$ refers to educational attainment, $a_{j,t}$ describes worker's experience, ϕ_h is the semi-elasticity of human capital with respect to health, ϕ_s is the semi-elasticity of human capital with respect to educational attainment, and $\phi_{a,1}$ and $\phi_{a,2}$ refer to the semi-elasticities of human capital with respect to experience and experience squared. We include the latter to capture the diminishing marginal contribution of experience to productivity.[3]

[2] This holds under the assumption that a marginal change of individual human capital does not change the distribution of wages such that the marginal product of individual human capital and that of average human capital coincide.

[3] Conceptually, $h_{j,t}$, $s_{j,t}$, and $a_{j,t}$ need not represent all aspects of health, educational attainment, and experience—only those that are relevant for the production of final output.

Accordingly, a worker j with ν_j units of human capital earns a wage of

$$w_{j,t} = w_t \cdot \nu_{j,t} = w_t \cdot \exp\left(\phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2\right). \tag{3.6}$$

This notation normalizes the effective labor input of a hypothetical worker without any health capital, education, or experience to unity. Meanwhile, workers with better health, higher education, or more experience are equivalent in productivity terms to a larger number of such baseline workers. Logarithmic wages at the individual level thus take the well-known Mincerian form:

$$\ln(w_{j,t}) = \ln(w_t) + \ln(\nu_{j,t}) = \ln(w_t) + \phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2. \tag{3.7}$$

Hence, the aggregate production function in (3.1) with our measure for human capital in (3.5) is consistent with wage equations used in the microeconomic literature.

The Mincerian wage form implies that the aggregate human capital stock is given by

$$H_t = \sum_j^{\mathcal{J}} \nu_j = \sum_j^{\mathcal{J}} \exp\left(\phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2\right). \tag{3.8}$$

Accordingly, aggregating human capital requires raising individuals' educational attainment and health to the exponential power. This complication in the aggregation process vanishes, if human capital and, thus, wages follow a log-normal distribution. In this case, the log of the average wage corresponds to the log of the median wage plus one-half of the variance of log wages σ_t^2. As the log of median wages equals the average of log wages for a lognormal distribution, the aggregate human capital stock simplifies to

$$\begin{aligned}
\ln\left(\frac{H_t}{L_t}\right) = \ln\left(\frac{\sum_j^{\mathcal{J}} \nu_{j,t}}{L_t}\right) &= \frac{\left[\sum_j^{\mathcal{J}} \ln(\nu_{j,t})\right]}{L_t} + \frac{\sigma_t^2}{2} \\
&= \frac{\sum_j^{\mathcal{J}} \phi_h h_{j,t} + \phi_s s_{j,t} + \phi_{a,1} a_{j,t} + \phi_{a,2} a_{j,t}^2}{L_t} + \frac{\sigma_t^2}{2} \\
&= \phi_h h_t + \phi_s s_t + \phi_{a,1} a_t + \phi_{a,2} a_t^2 + \frac{\sigma_t^2}{2}.
\end{aligned} \tag{3.9}$$

Intuitively, a marginally better health status (for example, an increase in the adult survival rate by one percentage point) raises a worker's productivity and wages by $100 \cdot \phi_h$ percent. Analogously, additional marginal investment in education (for example, one year of schooling) raises a worker's productivity and wages by $100 \cdot \phi_s$ percent. This effect's absolute size is larger for highly educated high-wage earners than it is for poorly educated low-wage workers. Moreover, an extra year of education for a highly educated worker also represents a greater investment, because the worker must forgo a higher wage to undergo the extra schooling.

3.4 Empirical Framework

Suppose the production function in (3.2) applies to $i = 1, \ldots, \mathcal{I}$ different countries. Taking the logarithm of the production function and using the result from Equation (3.9), the log of production per worker is given by

$$\ln(y_{i,t}) = \ln(A_{i,t}) + \alpha \ln(k_{i,t}) + (1-\alpha)\left(\phi_h h_{i,t} + \phi_s s_{i,t} + \phi_{a,1} a_{i,t} + \phi_{a,2} a_{i,t}^2 + \frac{\sigma_{i,t}^2}{2}\right). \quad (3.10)$$

Using rates of return to calibrate the coefficient on education, ϕ_s, suggests a parameter value of 0.09 to 0.10 (Psacharopoulos, 1994; Bils and Klenow, 2000; Psacharopoulos and Patrinos, 2004). Regarding the elasticity of output with respect to capital, α, economists generally agree on values of around one-third (see, for example, Hall and Jones, 1999).

Heckman and Klenow (1997) and Krueger and Lindahl (2001) take a similar approach, deriving a formula that estimates the macroeconomic effects of schooling using an aggregated version of a Mincer wage equation. The major difference between these formulations is that education level's effect on output in their formulation is expressed as ϕ_s, whereas in our approach the effect of schooling is $(1-\alpha)\phi_s$. This difference arises, because they assume the cross-country differences and changes in the intercepts in (3.7) to be random and assign them to the error term in the regression. With our production function, increases in schooling increase the aggregate level of human capital and labor equivalent inputs in the economy and depress the wage paid per equivalent worker.

Equation (3.10) describes aggregate production as an identity that could be estimated directly, if all right-hand-side variables were available. In practice, however, the level of total factor productivity in country i at time t, $\ln(A_{i,t})$, is not observed. Several approaches can address this problem. We follow Bloom, Canning, and Sevilla (2004) and model total factor productivity as a diffusion process across countries, which allows for the possibility of long-run differences in TFP even after the diffusion is complete. Specifically, let the change in TFP be given by

$$\Delta \ln(A_{it}) = \lambda \left[\ln(A_{i,t}^*) - \ln(A_{i,t-1})\right] + \varepsilon_{i,t}, \quad (3.11)$$

where $\varepsilon_{i,t}$ constitutes an idiosyncratic shock. Each country has a period-specific upper bound, given by $\ln(A_{i,t}^*)$. A country's total factor productivity adjusts toward this bound at rate λ. We assume this upper bound depends on country characteristics $x_{i,t}$ and on the worldwide technology frontier μ_t. Moreover, schooling in previous periods may facilitate the diffusion and adoption of existing technologies (Nelson and Phelps, 1966) or spur novel innovation (Romer, 1990; Aghion and Howitt, 1992; Strulik, Prettner, and Prskawetz, 2013). Hence, lagged schooling $s_{i,t-1}$ constitutes another determinant of potential TFP (see also Cuaresma, Lutz, and Sanderson, 2014). Neglecting one of these channels might bias the empirical estimates, as Sunde and Vischer's (2015) results indicate. Because technological gaps are not directly observed, we follow Baumol (1986) and use lagged output per worker as a proxy (see also Fagerberg, 1994; Dowrick and Rogers, 2002). Hence, growth of total factor productivity reads

$$\Delta \ln(A_{it}) = \lambda \left[\mu_t + x'_{i,t}\Theta + \rho s_{i,t-1} - \ln(y_{i,t-1})\right] + \varepsilon_{i,t}. \tag{3.12}$$

Alternatively, a richer model derives lagged total factor productivity $\ln(A_{i,t-1})$ directly from the production function such that

$$\begin{aligned}\Delta \ln(A_{it}) = \lambda \left[\mu_t + x'_{i,t}\Theta + \rho s_{i,t-1} - \ln(y_{i,t-1}) + \alpha \ln(k_{i,t-1})\right] \\ +\lambda(1-\alpha)\left(\phi_h h_{i,t-1} + \phi_s s_{i,t-1} + \phi_{a,1} a_{i,t-1} + \phi_{a,2} a^2_{i,t-1} + \frac{\sigma^2_{i,t-1}}{2}\right) + \varepsilon_{i,t}.\end{aligned} \tag{3.13}$$

This slightly more comprehensive modeling approach, however, suffers from the disadvantage that including additional highly correlated explanatory variables inflates the estimated standard errors without providing additional insights into the parameters of interest. As such, we provide estimates for both models and show that they are qualitatively and quantitatively similar.

Related research suggests several variables $x_{i,t}$, which may affect the TFP level in the long run. For example, Hall and Jones (1999) argue that institutions and "social infrastructure" affect productivity, while Gallup, Sachs, and Mellinger (1999) emphasize the role of geography. Our empirical work experiments with several potential variables to control for these influences.

First-differencing (3.10) and inserting (3.12) provides the empirical estimation equation:

$$\begin{aligned}\Delta \ln(y_{i,t}) = \lambda \left[\mu_t + x'_{i,t}\Theta + \rho s_{i,t-1} - \ln(y_{i,t-1})\right] + \alpha \Delta \ln(k_{i,t}) \\ +(1-\alpha)\left(\phi_h \Delta h_{i,t} + \phi_s \Delta s_{i,t} + \phi_{a,1} \Delta a_{i,t} + \phi_{a,2} \Delta a^2_{i,t} + \frac{\Delta\sigma^2_{i,t}}{2}\right) + \varepsilon_{i,t}.\end{aligned} \tag{3.14}$$

De la Fuente and Domenech (2001) and Bloom, Canning, and Sevilla (2004) use this approach to model TFP diffusion in cross-country production function studies. It is formally equivalent to the autoregressive TFP model that Griliches and Mairesse (1998) and Blundell and Bond (2000) use in their studies of the production function based on firm-level data.

According to the specification in (3.14), output growth can be decomposed into three components. The first is growth of the input factors capital, health, schooling, and experience. The second is a catch-up term capturing the reduction of the technological gap to the leading countries in each time period such that the country converges to its TFP upper bound at the rate λ. The third component is an idiosyncratic shock to the country's total factor productivity $\varepsilon_{i,t}$.[4]

[4] We could allow the shock to grow over time to have a common component across countries, such

Equation (3.14) represents a model of conditional convergence, in which the speed of convergence λ describes the rate at which gaps in total factor productivity close. Therefore, this approach stands in contrast to models that take TFP differentials across countries to be fixed, such as those of Mankiw, Romer, and Weil (1992) and Islam (1995). The speed of convergence in these models depends on the time that capital stocks take to reach their steady-state levels given fixed investment rates. By including the growth rates of factor inputs directly in Equation (3.14), we can identify the catch-up term—that is, the effect of the gap between actual and steady-state output, given current input levels—as the impact of a TFP gap.

In the special case of no technological diffusion ($\lambda = 0$), the lagged level terms in (3.14) disappear. Hence, our approach encompasses the estimation of a production function in first differences, as advocated by Krueger and Lindahl (2001) and Pritchett (2001). Moreover, we can test if this restriction holds. Taking first differences nets out any fixed effects on TFP. Therefore, testing whether $\lambda = 0$ examines the plausibility that TFP differentials remain constant or, alternatively, narrow over time because of technological diffusion. Our model also encompasses the special case in which technological diffusion occurs, but the steady-state level of TFP is the same in every country. We can test this by examining whether the country-specific variables $x_{i,t}$ have zero coefficients.

When estimating Equation (3.14), we face the possibility that contemporaneous growth rates of factor inputs are endogenous and responsive to the current TFP shock $\varepsilon_{i,t}$. For health and education inputs, which are the objects of interest, we overcome this problem by exploiting the demographic structure to obtain plausibly exogenous instrumental variables (see Chapter 2 and Kotschy and Sunde, 2018). Specifically, inflows from young-age cohorts at the lower end and outflows from old-age cohorts at the upper end of the working-age population determine changes in overall health status and educational attainment of the working-age population. Hence, one can use the lagged level of health and the lagged level of educational attainment for the age cohorts that will enter or leave the working-age population in the next period as an instrumental variable for the contemporaneous growth rate of the corresponding factor input. This instrument is plausibly exogenous given the approximation of TFP growth rates, which controls for productivity gains that are due to past changes in input factors, past technology shocks, and convergence to the technological frontier. This approach is compatible with lagged TFP levels and expected TFP growth—the catch-up term in Equation (3.14)—affecting previous input decisions (for example, Bils and Klenow, 2000, suggest that schooling decisions depend on expected economic growth). The argument that lagged input levels are uncorrelated with future shocks to TFP is the rationale for estimating Equation (3.14) instead of the level relationship in Equation (3.10).

---

as worldwide oil or interest rate shocks. Such a shock, however, would be collinear to changes in the worldwide productivity frontier captured by the time effects and would thus not affect any of our results.

Finally, including fixed effects in a comprehensive specification also allows for the possibility of country-specific growth trends driven by unobserved heterogeneity with a persistent effect on TFP. We refrain, however, from using dynamic panel estimators based on generalized method of moments, which produce estimates with large standard errors and no statistical significance. In addition, we take the view that over the five-year intervals all the inputs potentially correlate with contemporaneous productivity shocks. Therefore, we would have to instrument all our regressors by lagged values, as opposed to firm-level studies, in which current inputs are treated as exogenous. Both of these factors imply a loss of precision in the estimates and make drawing inferences based on a fixed-effects approach difficult.

3.5 Data

We construct an unbalanced panel of 116 countries observed every five years from 1960 to 2010. Data on real output and physical capital, both in per worker units, are obtained from the Penn World Tables by Feenstra, Inklaar, and Timmer (2015).

Health inputs are proxied using adult survival rates derived from United Nations (2017). This variable measures the probability of surviving from age 15 to 60. Conceptually, this measure may relate more closely to adult health and worker productivity than life expectancy—a measure that is sensitive to infant mortality rates. Adult survival rates, however, act only as a proxy for the health of the workforce, because they measure mortality rates rather than morbidity. Our main reason for using adult survival rates is that it allows us to compare our results directly with those of Shastry and Weil (2003) and Weil (2007).

Following the Mincerian approach, educational input is proxied by years of schooling in the working-age population. To this end, we exploit measures on secondary and total schooling from Barro and Lee (2013) for the population above age 15. We combine age-specific years of schooling with population shares to construct average years of schooling for the working-age population, which we define from age 15 to 60 to match our measure of aggregate health.

We construct aggregate experience as the median age of the population obtained from United Nations (2017), net of an intercept of six years corresponding to early childhood. Moreover, we deduct compulsory schooling years, taken from UNESCO (1997) and UNESCO (2017), to account for differences in the age of workforce entry across countries. This correction is necessary, because countries with higher life expectancy and older populations tend to have later workforce entry due to longer schooling. As experience enters the regression framework in differences, this measure takes up variation from changes in median age and compulsory schooling following educational reforms.[5]

[5]For certain countries, the statistical yearbooks report values for specific regions. Moreover, the

To control for the effect of wage inequality, we use the disposable income Gini coefficient after taxation and transfers by Solt (2016a). These data provide standardized Gini coefficients that are comparable across countries and over time.

Finally, we include some country-specific variables that may affect long-run TFP levels. These include an indicator for the quality of economic institutions from Gwartney, Lawson, and Hall (2017), a measure for the value added by the agricultural sector from World Bank (2017) to control for structural change, the percentage of land area in the tropics by Gallup, Sachs, and Mellinger (1999) to control for geographical factors that may affect productivity and trading opportunities, and a set of regional dummies.[6]

3.6 Results

3.6.1 Baseline Results

Table 3.1 presents the main estimation results. We proxy education by average years of secondary schooling, which provides the most precise estimates for the return to education. Column (1) reports coefficient estimates of a parsimonious specification of our empirical model in Equation (3.14), including lagged educational attainment but omitting any additional controls. The point estimates show the sign expected from theory. Lagged per capita GDP is negative, implying conditional convergence as predicted by the neoclassical growth literature (Solow, 1956; Cass, 1965; Diamond, 1965) and as established empirically.[7] Capital accumulation positively relates to economic growth, which again conforms to the growth literature's results. Changes in the aggregate human capital stock positively affect productivity per worker: The coefficients for changes in average health and mean years of schooling both show a positive sign and differ significantly from zero at the one-percent level. Hence, health and education both constitute important dimensions of human capital. The opposing signs for $\widehat{\phi}_{a,1}$ and $\widehat{\phi}_{a,2}$ suggest a hump-shaped effect of average experience on growth of output per worker, which is consistent with the standard Mincerian framework; though only the coefficient of squared experience is (marginally)

educational systems of some countries allow for different categorizations such that alternative figures are conceivable. We correct for these fluctuations and code flatter, that is, less varying, values in the case of doubt. This procedure tends to render the measure for experience less informative and thus increases the corresponding standard errors. Table B.2 in the Appendix contains a complete list of coding decisions. Because we use only changes in experience over time, measurement error in compulsory schooling levels poses no threat to our identification.

[6]We also experimented with further indicators for landlocked countries by Gallup, Sachs, and Mellinger (1999) and controls for ethnic fractionalization and polarization by Alesina et al. (2003) and Reynal-Querol and Montalvo (2005). Given the set of other controls, however, these variables did not explain much of the remaining variation.

[7]See, for example, Barro (1991, 1997), Sala-i-Martin (1997), and Doppelhofer, Miller, and Sala-i-Martin (2004) in cross-section regressions, and Islam (1995), Caselli, Esquivel, and Lefort (1996), and Brückner (2013) in panel data settings. For interesting surveys and critical remarks on the literature, see Durlauf, Johnson, and Temple (2005) and Eberhardt and Teal (2011).

Table 3.1: Return of Health and Education to Productivity

	No Controls	Adding Controls	Fixed Effects	Adding Gini	Lagged Controls	IV $\Delta(h_{i,t})$	IV $\Delta(s_{i,t})$	IV $\Delta(h_{i,t})$, $\Delta(s_{i,t})$
Regressors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln(y_{i,t-1})$	-0.045***	-0.15***	-0.34***	-0.18***	-0.20***	-0.14***	-0.14***	-0.14***
	(0.0097)	(0.017)	(0.052)	(0.023)	(0.022)	(0.020)	(0.017)	(0.020)
$\Delta\ln(k_{i,t})$	0.52***	0.35***	0.29***	0.23***	0.48***	0.34***	0.34***	0.35***
	(0.054)	(0.063)	(0.10)	(0.067)	(0.075)	(0.061)	(0.064)	(0.064)
$\Delta(h_{i,t})$	1.12***	0.59**	0.64**	0.74**	0.68**	0.91	0.61**	0.80
	(0.30)	(0.29)	(0.27)	(0.35)	(0.28)	(0.94)	(0.28)	(0.90)
$\Delta(s_{i,t})$	0.099***	0.075**	0.091**	0.063**	0.063**	0.078***	0.049	0.046
	(0.032)	(0.031)	(0.046)	(0.027)	(0.029)	(0.030)	(0.096)	(0.093)
$\Delta(a_{i,t})$	0.011	0.0075	-0.0100	0.0089	0.0090	0.0070	0.0078	0.0073
	(0.0085)	(0.0088)	(0.0095)	(0.010)	(0.0085)	(0.0084)	(0.0085)	(0.0084)
$\Delta(a^2_{i,t})$	-0.00069*	-0.00057	0.00024	-0.00058	-0.00064*	-0.00056*	-0.00057	-0.00055
	(0.00036)	(0.00036)	(0.00036)	(0.00039)	(0.00033)	(0.00034)	(0.00035)	(0.00034)
$\Delta(\sigma^2_{i,t})$	—	—	—	-0.60	—	—	—	—
	—	—	—	(0.45)	—	—	—	—
R^2	0.29	0.38	0.38	—	0.41	0.38	0.38	0.37
First-stage F	—	—	—	—	—	26.7	30.2	17.8
Countries	116	116	116	109	116	116	116	116
Observations	613	613	613	461	613	613	613	613
Controls	—	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of log output per worker $\Delta\ln(y_t)$. All specifications include lagged schooling $s_{i,t-1}$ and a full set of time effects. Columns (2) to (8) add further controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, and a full set of regional dummies. First-stage F refers to the Kleibergen-Paap first-stage F-statistic. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

significant in this specification. Because experience varies strongly across individuals but very little across countries, however, obtaining a precise estimate of the effect of worker experience in macroeconomic models is difficult (Bloom, Canning, and Sevilla, 2004).

The specification used in Column (2) includes controls for the quality of economic institutions, the value added by the agricultural sector (as a proxy for structural change), the percentage of land area in the tropics, and the set of regional dummies. Adding these controls increases the model's explanatory power as reflected by an increase in R^2 and slightly improves the precision of the point estimates. Quantitatively, the computed parameters reduce in magnitude compared with the parsimonious specification, but the qualitative results remain unchanged. In particular, the estimates still indicate a positive and significant effect of changes in average health and education on output per worker. Interestingly, the reduction in magnitude of the return to health is almost entirely due to the inclusion of institutional quality. This confirms recent evidence by Weil (2014, 2017), who finds that institutional differences account for a considerable portion of the cross-country correlation between income and health. Nevertheless, our results also show that even after controlling for institutional differences, significant scope exists for a positive causal effect of health on output per worker.

In Column (3), we include country fixed effects to allow for country-specific growth trends. Again, the qualitative effects remain unchanged. Quantitatively, the estimated return to health does not change considerably, while the return to education slightly increases. Our result of a positive and significant return to health is therefore not driven by country-fixed unobservables. This specification substantially restricts the potential of omitted variables to bias our main outcome of interest and thus serves as a specification test for the model without fixed effects. In Column (4), we augment the specification in Column (2) by adding the Gini coefficient to approximate the variance of log wages. For reasons of data availability, the estimation sample shrinks to 461 observations.[8] The qualitative results again remain unchanged, while the estimated parameters do not change considerably in magnitude. The computed parameter for the disposable income Gini coefficient is negative and insignificant. Finally, Column (5) presents the results for the comprehensive model with lagged controls, which derives lagged TFP directly from the production function according to Equation (3.13). The results conform quantitatively and qualitatively to those in Columns (2) to (4).

Table 3.2 compares the results of the baseline specification in Column (2) with those of the literature. According to Weil (2007), an increase in adult survival rates of 0.1—or 10 percentage points—raises labor productivity by 6.7 percent. In comparison, our estimates indicate that an increase in adult survival rates of 0.1 translates into a 9.1-percent increase in labor productivity.[9] The slightly larger point estimate from the macro-based approach might be due to spillover effects that the micro-based approach omits by design. The 95-percent confidence interval of our estimate ranges from 0.47 to 17.7 percent and hence includes Weil's (2007) estimate. Consequently, our macro results are consistent with the micro results, reconciling the micro-based and macro-based approaches to estimating the effect of health on income growth.

Our coefficient estimate for changes in physical capital α is 0.35. This is in line with empirical estimates of output elasticity with respect to physical capital, which fall around 0.3 to 0.4 (Hall and Jones, 1999). Moreover, dividing our estimate of education by $(1-\alpha)$ yields a return to secondary schooling ϕ_s of 11.4 percent. This effect is consistent with the estimates obtained by Psacharopoulos and Patrinos (2004), who report a 13.1-percent return on secondary schooling. Finally, the signs of our coefficient estimates on the lagged dependent variable γ and our experience measures $\phi_{a,1}$ and $\phi_{a,2}$ accord with previous findings on conditional convergence and positive but diminishing returns to experience.

[8] To increase data availability, Solt (2016a) uses imputation procedures to reduce the number of missing values in the data set. As this procedure may understate the uncertainty in the data and thus lead to downward-biased standard errors, we conduct a standard error adjustment as suggested by Solt (2016b). In particular, we estimate the specification in Column (4) for 100 potential realizations of the Gini coefficient and compute the final estimates as the average over all individual results. For details, see Solt (2016b).

[9] To obtain the figures for ϕ_h and ϕ_s in Equation (3.14), divide the estimates in Table 3.1 by $(1-\alpha)$.

Table 3.2: Comparison Between Our Estimates and the Literature

Variable	Point Estimate	Confidence Interval	Target
γ	< 0		< 0
α	0.35	(0.23–0.47)	0.3–0.4
ϕ_h	9.1%	(0.47%–17.72%)	6.7%
ϕ_s	11.4%	(2.6%–20.26%)	13.1%
$\phi_{a,1}$	> 0		> 0
$\phi_{a,2}$	< 0		< 0

3.6.2 Instrumental Variables Approach

To address concerns about the endogeneity of health and education in our empirical model, we present instrumental variable regressions for the two human capital variables health and education in Columns (6) to (8) of Table 3.1. To this end, changes in health and education are explained by in- and outflows of young- and old-age cohorts at the lower and upper ends of the working-age population (see Chapter 2 and Kotschy and Sunde, 2018). Hence, we can use lagged levels for the youngest and oldest age cohorts to predict contemporaneous changes in aggregate health and education of the working-age population. Given the strong persistence in demographic patterns, these in- and outflows are plausibly exogenous in an empirical specification that controls for past levels of per capita output and time-specific fixed effects. Moreover, migration is less of a concern for these age groups in contrast to prime-age workers. In Column (6), we apply this identification strategy to instrument changes in health. In particular, contemporaneous changes in health are instrumented by the health status of the cohort aged 10–14 in the last period, measured in terms of infant mortality at time of birth. Optimally, we would also like to capture the outflow of aggregate health by using lagged health of the cohort aged 55–59; however, a lack of data on child mortality rates from before the world wars prohibits us from doing so. In Column (7), contemporaneous changes in education are instrumented by average years of secondary schooling of the cohort aged 55–59 years in the last period. We refrain from using the inflow of schooling for young-age cohorts, because individuals might anticipate future economic growth and thus increase educational attainment. Finally, Column (8) reports results for a specification, in which both health and education are instrumented. While the sign of each coefficient remains stable compared with the other regressions, the instrumented variables become insignificant. However, this loss of significance is of less concern, because the size of the estimated coefficients does not change considerably. Weak instruments are no concern, as sufficiently high values of the first-stage F-statistic indicate.

3.6.3 Robustness Checks

As a robustness check, Table B.1 in the Appendix contains the estimation results of an empirical model that proxies education by average years of total schooling instead of average years of secondary schooling. Again, we observe that the expected signs of the coefficients remain stable throughout all specifications. In particular, the estimated effect of health on output per worker is quantitatively almost identical compared with the baseline results. Therefore, we confirm our main result that micro-based and macro-based estimates of the return to health are consistent with each other.

The estimated return to average years of schooling loses its significance; however, the variation in this regressor may be less informative than the variation in average years of secondary education, because primary education does not change appreciably over the time period examined in most countries. In any case, this discrepancy is not a major concern, because the return of health, which is the main parameter of interest, remains relatively stable throughout the specifications.

3.7 Conclusion

Much of the economic growth literature has been devoted to studying the impact of education on aggregate economic performance and comparing the results with the rate of return to education identified by the Mincer (1974) wage equation. We believe our study is the first to show that the macroeconomic estimates of the effect of health on output are compatible with the microeconomic estimates of the effect of health on wages. According to our estimates, an increase in adult survival rates of 0.1, or 10 percent, increases labor productivity by about 9.1 percent, which is somewhat higher than, but still consistent with, Weil's (2007) calibrated value of around 6.7 percent. The slightly larger point estimate from the macro-based approach might be due to spillover effects that the micro-based approach omits by design. Altogether, our results support Weil's (2007) conclusion that health plays a role in explaining cross-country differences in the level of income per worker. Given that we find no evidence of substantial externalities, this result, moreover, suggests that calibration based on microeconomic data can serve as a reasonable means to estimate the macroeconomic impact of health changes.

As far as policy implications are concerned, public health measures might be an important lever for fostering economic development. Potential policies along these lines include vaccination programs, antibiotic distribution programs, and iodine supplementation schemes, which lead to large improvements in health outcomes for relatively low expenditures (World Bank, 1993; Commission on Macroeconomics and Health, 2001; Field, Robles, and Torero, 2009; Luca et al., 2014).

Chapter 4

Life Expectancy and Life-Cycle Wages: Evidence from the Cardiovascular Revolution in U.S. States

4.1 Introduction

Medical advancement in the twentieth century has spurred a substantial increase in longevity in the United States. As a consequence, the number of older but also healthier workers increased substantially. This development raises several questions. Do improved health conditions as measured by adult life expectancy lead to more productive workers? Moreover, do health shocks affect the population homogeneously? And, finally, what are potential channels for a causal link?

In order to answer these questions, I exploit variation in the unexpected sharp decline in mortality rates from cardiovascular diseases among U.S. states beginning in the 1960s. This decline, also referred to as cardiovascular revolution (for example, Foege, 1987), is used as an instrument for adult life expectancy in a balanced ten-year panel from 1940 to 2000 for the 48 contiguous U.S. states. The identification strategy exploits initial differences in mortality from cardiovascular diseases across U.S. states in 1960, when there existed little treatment possibilities for these diseases. Between 1960 and 1970, a number of path-breaking innovations in the treatment of cardiovascular diseases were introduced and behavioral risk factors identified. The availability of these treatments as well as follow-up inventions and public education about risks helped reduce mortality from cardiovascular diseases by roughly 50 percent between 1970 and 2000 (CDC, 1999b; National Heart, Lung, and Blood Institute, 2012a). The decline in cardiovascular mortality entailed a substantial increase in adult life expectancy, which varied across states, depending on the

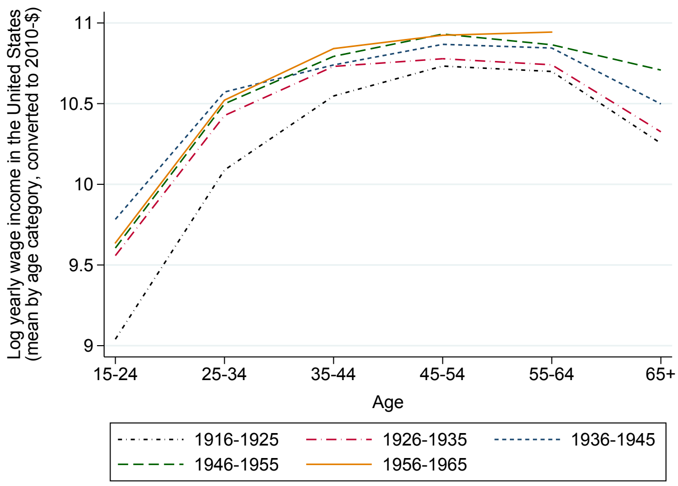


Figure 4.1: Life-Cycle Wage Profiles by Birth Cohort

Data source: IPUMS (Ruggles et al., 2015).

initial prevalence of cardiovascular diseases. Therefore, this quasi-experimental source of variation allows the estimation of a differences-in-differences model, where all states are treated though with varying treatment intensities. State-year observations for 1940–1960 constitute the pre-treatment and for 1970–2000 the post-treatment period.

The paper contributes to the literature in several ways. First, the empirical results establish a positive causal link between adult health, as measured by adult life expectancy, and age-specific wages per worker. The decline of cardiovascular mortality in the U.S. from 1968 onward led to an increase of life expectancy at 50 of approximately 3.16 years, or roughly two thirds of the increase between 1960 and 2000. According to the baseline estimation, this rise in life expectancy caused an increase of average gross wages for the group of the 45- to 54-year-olds of roughly 9,762$, or 31 percent of initial wages in 1960. This wage hike corresponded to 47 percent of the wage change observed in the same time window. Furthermore, the results reveal that wage gains accrued to workers in the prime-age group between 25 and 54 as well as to old-age workers above 65. Compared to earlier generations, the life-cycle earnings profile of an average worker thus increases more steeply at younger ages, while it flattens out more slowly at higher ages. Figure 4.1 illustrates this shift for wages of U.S. whites born between 1916 and 1955 and grouped in ten-year cohorts. Overall, this pattern is consistent with a workforce that over time becomes healthier at any given age, and at higher ages in particular.

Another contribution is the focus on the role of measurement of health conditions in the context of age-specific outcomes. In many studies, life expectancy, a one-dimensional summary measure of the survival experience of the population, serves as a proxy for the average health status of the population of interest, for example, the total workforce. In such a case, one implicitly assigns all individuals the same health status (or change thereof). This assumption may produce severe systematic measurement error, if the chosen proxy does not closely reflect the health conditions of the population of interest. For example, consider the third stage of the epidemiological transition during which life expectancy at birth substantially increased thanks to reduced infant mortality following the invention of vaccines and antibiotics (WHO, 2008). As important as this health shock was, it may grossly overstate the health improvement for the median American who is around age 30 at this time.[1] This type of mismeasurement introduces a systematic correlation between the proxy of the health shock and the error term, thus leading to biased estimates. In particular, mismeasurement leads to downward-biased estimates, if the change in average population health is overstated. Therefore, age-specific heterogeneity in the effect of health shocks and mismeasurement might be a reason for null results of life expectancy at birth on GDP per capita found by Acemoglu and Johnson (2007, 2014), Hansen (2014) and Bloom, Canning, and Fink (2014). Even though these papers use the mortality rate from infectious diseases as instrumental variable for life expectancy at birth, they cannot mend the measurement problem, because the first stage again overstates the health improvement for the median person. Hence, the published estimates can be considered a lower bound for the causal effect of health improvements on economic growth.

Lastly, this paper can make progress in analyzing potential channels through which adult health affects average wages by using individual data on health outcomes and economic variables, as well as by combining data on longevity from vital statistics with census data on wages, educational attainment, and labor supply. In particular, U.S. states provide a favorable setting, because the institutional environment for the labor market is homogeneous in contrast to cross-country studies. In addition, there is no binding statutory retirement age in the U.S., which offers a clearer picture of productivity and labor supply for old-age workers above 65 compared to other developed countries.[2] The timing of wage hikes suggests that potential channels are health improvements, in particular in the short-run, and higher educational attainment and changes in individual behavior toward a more healthy lifestyle in the long run. In contrast, adjustments in labor supply cannot explain the wage increase, because labor force participation rates as well as usual working hours and weeks either declined or remained unchanged during the treatment period. Moreover, heterogeneity in age group estimates preclude the possibility of unilateral indirect wage effects through out-selection or increased bargaining power. Thus, thanks to higher adult life expectancy, workers earn more, invest more in educational attainment, but work slightly less. This evidence confirms theoretical predictions and results from simulation exercises by Cervellati and Sunde (2013) and Strulik and Werner (2016). They show that individuals may invest more in schooling and, at the same time, reduce lifetime working hours, if leisure time while at work and consumption over the

[1]The median age of the U.S. population was 29.0 in 1940 and 30.2 in 1950 (Hobbs and Stoops, 2002).

[2]Nonetheless, certain age thresholds may still affect the timing of retirement. In particular, Americans become eligible for Medicare at age 65; full Social Security benefits can be claimed around age 66 depending on birth cohort; and there exist no further monetary incentives for delaying retirement beyond age 70.

life-cycle increase.[3] Therefore, higher lifetime labor supply is not a necessary condition for increased educational attainment, as was claimed by Hazan (2009).

The paper's main result of a positive association between adult health and average wages per worker also holds for long-differences models and specifications that either use a shorter pre- or post-treatment window, or both. Furthermore, the econometric model accounts for initial state-level differences in income, education, and the rural-urban gradient, as well as state-fixed effects and differential time trends across census regions. Moreover, robustness tests show that sub-state level heterogeneity in the prevalence of cardiovascular diseases or interstate migratory patterns are unlikely to produce a spurious correlation between adult life expectancy and average wages on the state level. Finally, the analysis reveals heterogeneity in the beneficial effects of health improvements on average wages between rural and metropolitan areas as well as different occupational groups.

This paper relates most closely to work by Hansen and Strulik (2017), who investigate the link between adult health and college enrollment of 18- to 24-year-old Americans by also exploiting variation from the cardiovascular revolution across U.S. states. Instead, this study examines specifically how adult health affects average wages of different age groups by using cohort-specific variation over time. In particular, this approach uncovers age-specific heterogeneity with respect to the causal effect of health gains on wages, which, otherwise, could not be detected: Gains in adult health exert a positive effect on wages for workers aged 25–54 but not for workers aged 55–64. This finding is also consistent with a side result of Hansen and Strulik (2017), who find no causal effect of adult life expectancy on wages pooling variation for workers aged 30–65. Furthermore, both papers complement each other: According to the results presented in this study, education constitutes one potential channel through which health improvements increase worker wages in the long-run; however, education cannot explain hikes in wages immediately following the treatment.[4] Based on micro data, this study additionally finds that health innovations have marginalized negative effects of cardiovascular diseases on individual income over relatively short time. This result indicates significant positive health effects on average wages per worker. Other closely related work is from Bleakley (2007) and Bhalotra and Venkataramani (2015), who exploit similar empirical strategies to identify positive long-run effects of health improvements during childhood on adult education, income, and labor supply. In contrast to their work, this study focuses on a health shock that predominantly affects adults and that, due to the nature of cardiovascular diseases, unfolds heterogeneous effects across age groups.

[3]This finding is also consistent with work by d'Albis, Lau, and Sánchez-Romero (2012), who demonstrate that gains in life expectancy may lead to earlier retirement given that mortality reductions occur at sufficiently young age to provide substantial increases in individual's expected lifetime human wealth.

[4]As an internal consistency check, I re-estimate the effect of adult life expectancy on college enrollment using the baseline specification of this paper. The resulting parameter estimates are quantitatively similar to those of Hansen and Strulik (2017).

Furthermore, this study relates to a large macro literature that has investigated the effect of aggregate health measures on economic outcomes. In particular, reductions in mortality and gains in longevity foster per capita income growth (Bloom, Canning, and Sevilla, 2004; Lorentzen, McMillan, and Wacziarg, 2008; Cervellati and Sunde, 2011; Strittmatter and Sunde, 2013; Hyclak, Skeels, and Taylor, 2016); are conducive to investment in educational attainment (Tamura, 2006; Jayachandran and Lleras-Muney, 2009; Hansen and Strulik, 2017); spur old-age savings (Bloom, Canning, and Graham, 2003; De Nardi, French, and Jones, 2009); and reduce fertility (Hansen, Jensen, and Lønstrup, 2018; Ager, Hansen, and Jensen, 2018). In addition, these channels potentially interrelate closely (Zhang and Zhang, 2005). This paper provides a cohort-based analysis of the effect of adult health on average wages, which is novel to this literature. By focusing on different age groups, the empirical analysis uncovers that health improvements benefit prime-age workers between 25 and 54 as well as old-age workers above 65. Therefore, the cohort analysis allows to track shifts in the life-cycle earnings profile of the average worker that follow from the standard theories of human capital by Mincer (1958) and Ben-Porath (1967). Related work reports mixed results on the change of life-cycle earnings profiles for specific occupational groups. For example, Jones and Weinberg (2011) find that creativity peaks of researchers as measured by scientific breakthroughs have shifted to higher ages. In contrast, evidence by Falck, Heimisch, and Wiederhold (2016) indicates that introduction of information and communication technologies might shift productivity peaks to younger ages. The evidence in this paper implies a steeper life-cycle profile at younger ages, which flattens out more slowly at higher ages. An optimistic interpretation of this result suggests that health gains for prime-age and old-age workers might boost aggregate productivity for aging societies and thus confine potentially adverse effects of demographic change. Therefore, this paper also connects to work that investigates the role of demographic change for past and future development, for example, Feyrer (2007), Sánchez-Romero (2013), Cuaresma, Lutz, and Sanderson (2014), and Kotschy and Sunde (2018).

The remainder of this paper is structured as follows. Section 4.2 presents background information on health improvements during the cardiovascular revolution. Section 4.3 introduces the data as well as the empirical framework and discusses key identifying assumptions. Section 4.4 presents the estimation results and examines potential channels through which adult health may affect wages. Finally, Section 4.5 concludes.

4.2 Background: The Cardiovascular Revolution

Over the course of the twentieth century, the United States experienced substantial improvements in public health leading to a marked increase in life expectancy. In particular, these improvements came down to two separate waves of medical breakthroughs: the epidemiological transition and the cardiovascular revolution. Figure 4.2 depicts the decline

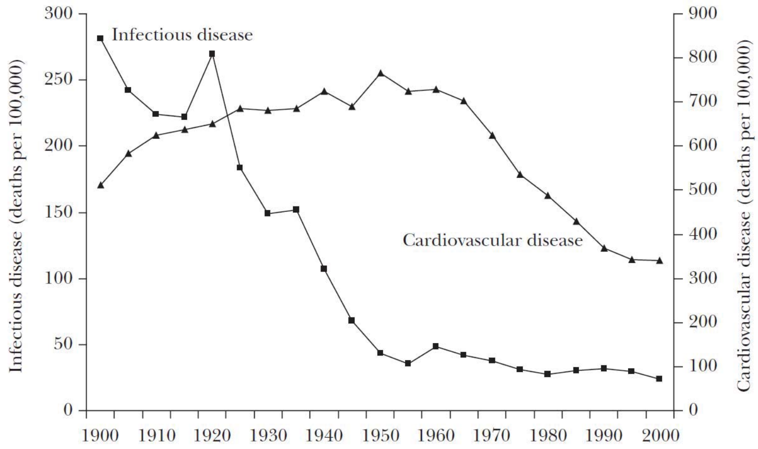


Figure 4.2: Mortality Rates from Infectious and Cardiovascular Diseases

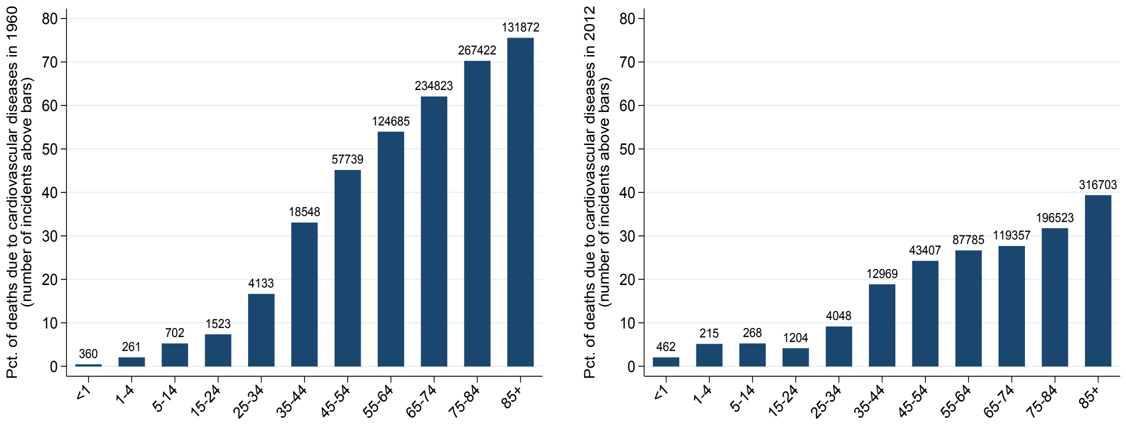
Source: Cutler, Deaton, and Lleras-Muney (2006, p. 104).

in mortality rates for infectious and cardiovascular diseases in the U.S., which resulted from these events.

The invention of antibiotics and vaccines in the first half of the twentieth century initiated a sharp reduction in mortality from communicable infectious diseases, which was especially pronounced for infants and children. The inventions during this period, which was termed the third stage of the epidemiological transition (Omran, 1971), caused an exceptional increase in life expectancy at birth, however, a significant but in comparison modest gain for higher ages.

The pointed increase in old-age life expectancy had to wait until the second wave of medical innovations around 1960, which was labeled cardiovascular revolution and identified as fourth stage of the epidemiological transition (Olshansky and Ault, 1986; Omran, 1998). The unexpected invention of new treatment possibilities for the non-communicable cardiovascular diseases boosted life expectancy predominantly through a decrease or delay in old-age mortality. Cardiovascular diseases become more likely as the tissues of the cardiovascular system age and lose some of their flexibility (Kirkwood, 2001). Therefore, mortality rates from cardiovascular diseases increase steadily with age, as exemplified by Figure 4.3.

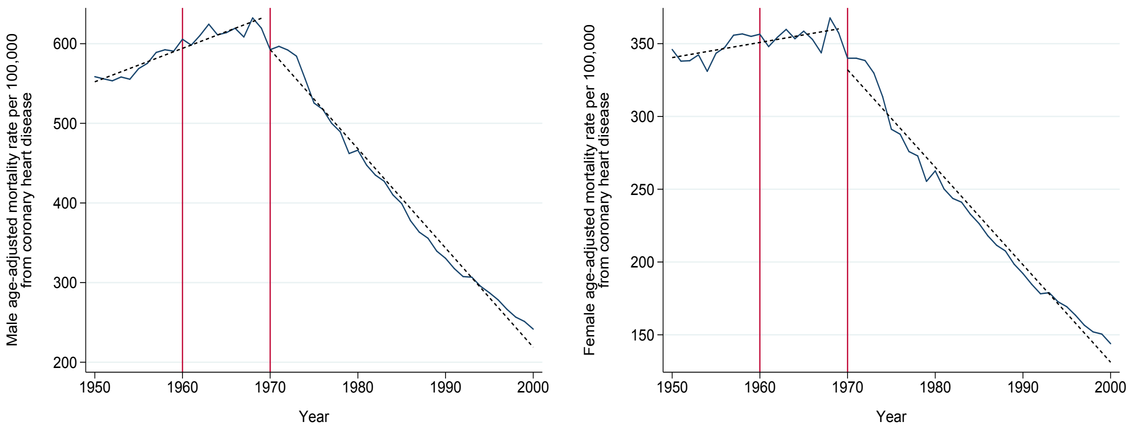
The cardiovascular revolution was successful in considerably reducing the mortality rates from a broad spectrum of cardiovascular illnesses; for example, coronary heart disease, which in 2000 still accounted for approximately twelve percent of total deaths in the U.S. (National Center for Health Statistics, 2017a), and which in 2004 still was the most common cause of death in high-income countries (WHO, 2008). Figure 4.4 showcases how powerful the decline in mortality from cardiovascular diseases was: Between the peak levels in 1968 and the year 2000, mortality from coronary heart disease fell by roughly



(a) 1960 (b) 2012

Figure 4.3: Percent of Deaths that are Attributable to Cardiovascular Diseases

Data sources: National Center for Health Statistics (1963) and National Heart, Lung, and Blood Institute (2012b).

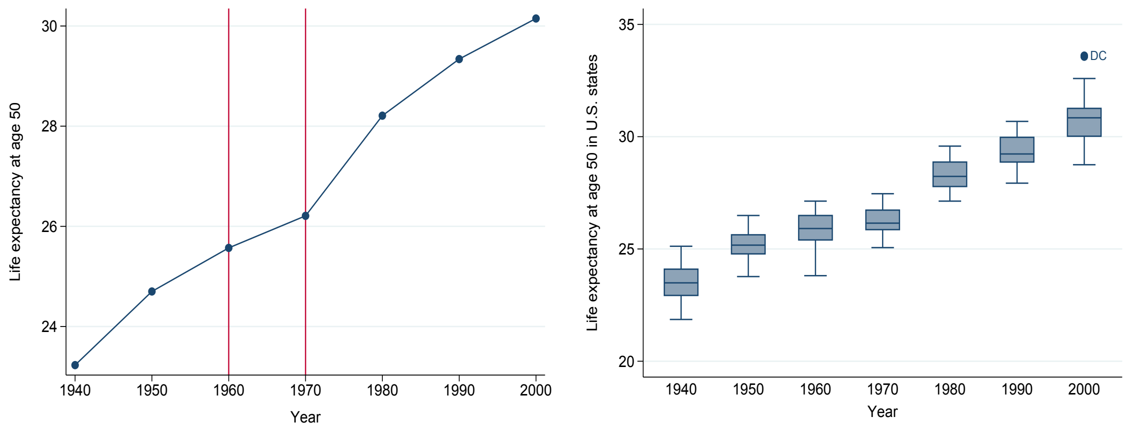


(a) Males (b) Females

Figure 4.4: Age-Adjusted Mortality from Cardiovascular Diseases

Data source: National Heart, Lung, and Blood Institute (2012a).

two thirds for both, men and women. As Figure 4.3 portrays, the number of incidents dropped for all age groups, except for infants and those above age 85, although the median age of the population had increased from 29.5 to 35.3 years during this period (Hobbs and Stoops, 2002). The decline was especially pronounced for individuals in the age range 35–84, thus especially boosting adult life expectancy as illustrated by Figure 4.5. In contrast, for the group above age 85, the number of incidents more than doubled during this period; however, the overall share of deaths that is attributable to cardiovascular diseases halved from almost 80 to slightly below 40 percent. One reason was that newly introduced drugs and treatment methods delayed the critical point at which the cardiovascular disease



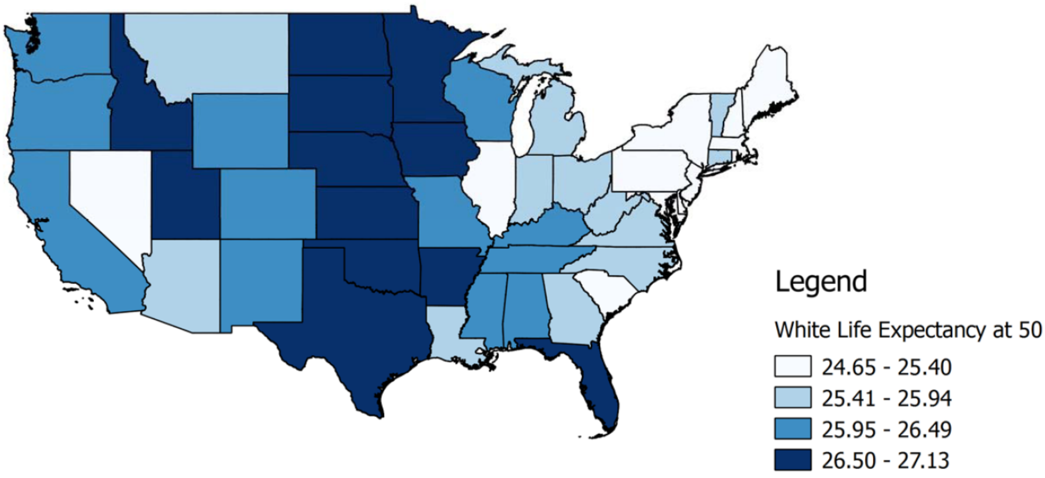
(a) United States (b) States

Figure 4.5: Life Expectancy at Age 50 in the United States

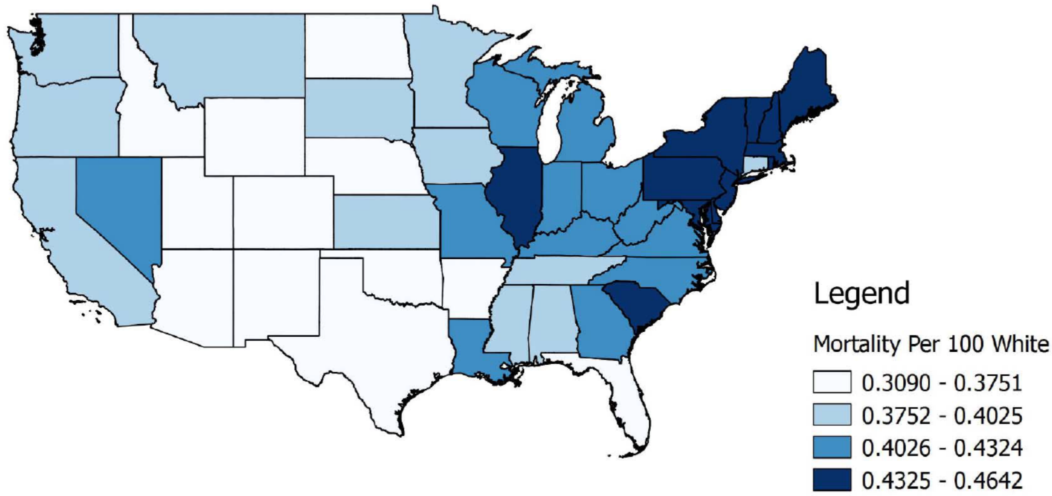
Data sources: United States Life Tables 1940–2000.

became lethal such that incidents occurred either at a higher age, or death originated from other sources such as cancer. Importantly, there have been striking geographic differences in the prevalence of cardiovascular diseases across U.S. states, which above all were rooted in social, cultural, and environmental factors (CDC, 1999b). The initial prevalence of cardiovascular diseases determined how beneficial the treatment was for states. Hence, the decline in mortality and, consequently, the increase in adult life expectancy varied across states. Figure 4.6 displays the geographic differences in life expectancy at 50 and mortality from cardiovascular diseases in the year 1960.

Reductions in mortality from cardiovascular diseases arrived through two channels. First, a number of medical innovations between the years 1960 and 1970 allowed to prevent certain diseases or to treat the symptoms. The most remarkable among these inventions were the artificial cardiac pacemaker, which was first implanted in 1958; the application of chest compression to restore blood circulation in a person that is in cardiac arrest beginning in 1960; the invention of the beta blocker in 1962, which is used to lower blood pressure and to treat cardiac arrhythmia; the invention of the portable defibrillator in 1959 and its application in the U.S. from 1966 onward; and the first adult human heart transplantation in the U.S. in 1968. Subsequent innovations include first thrombolytic therapies in 1986 to treat myocardial infarction, stroke, and pulmonary embolisms; the invention of cholesterol lowering statins, first marketed in 1987; and beginning in 1988, the implantation of intravascular stents to address acute closure of arteries and blood vessels. These new treatments improve health relatively quickly; for example, serum cholesterol reducing drugs achieve their full effect within five years (Law, Wald, and Thompson, 1994). These advances in the available technology were complemented by an increasing number of specialists and care centers for cardiovascular diseases (CDC, 1999b).



(a) Life Expectancy at 50



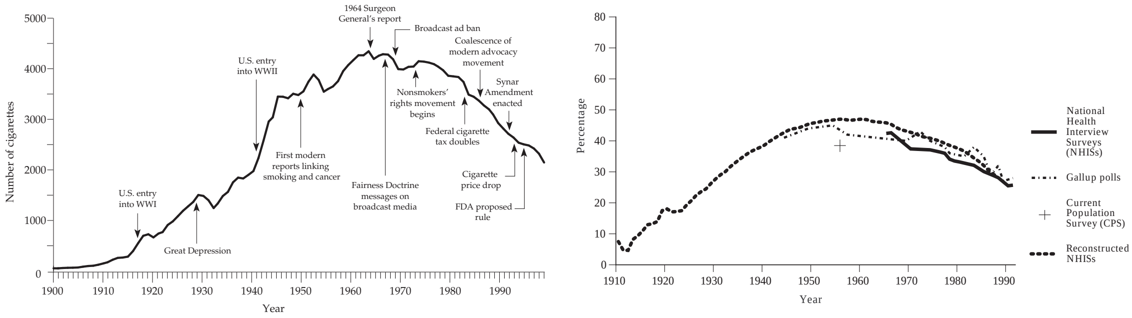
(b) Mortality from Cardiovascular Diseases

Figure 4.6: Mortality from Cardiovascular Diseases and Life Expectancy at 50

Data source: Grove and Hetzel (1968).

The second channel for the decline in mortality constituted increased awareness of major individual risk factors and changes in behavior. Research results by Keys et al. (1963), Keys (1980), and Dawber (1980) established, among others, high blood cholesterol, high blood pressure, physical inactivity, smoking, obesity, and unbalanced diet as major risk factors for cardiovascular diseases.[5] The federal government initiated national programs to educate specialists and the general public about risks of high blood pressure in 1972; of high blood cholesterol in 1985; and of the importance of cardiovascular health in 1989 (CDC, 1999a).

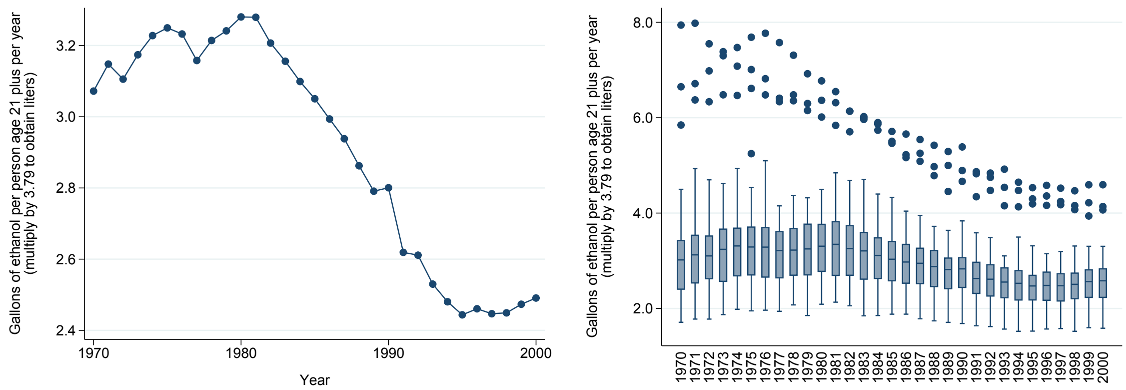
[5]According to Ezzati and Riboli (2012), high blood pressure and high blood cholesterol alone account for one half of the global incidence of coronary heart disease. Too high body weight and smoking are responsible for another 20 and 13 percent.



(a) Cigarette Consumption Per Capita (b) Share of Smokers in Population

Figure 4.7: Smoking in the United States

Sources: U.S. Department of Health and Human Services (1998, p. 123; 2000, p. 33).



(a) United States (b) States

Figure 4.8: Alcohol Consumption in the United States

Data source: Haughwout and Slater (2017)

This increased awareness helped raise the share of patients with too high blood pressure, who have their condition treated and under control (CDC, 1999b). Moreover, the report of the Surgeon General in 1964 (U.S. Department of Health, Education, and Welfare, 1964) highlighted the adverse effects of smoking on health, later followed by increased cigarette taxes in the 1980s (CDC, 1999a). Preventive measures against smoking were particularly successful, as Figure 4.7 shows. The share of smokers in the adult population was declining from 1960 onward, while the per capita cigarette consumption started to rapidly fall during the 1970s. Due to the cumulative damage from smoking, however, it takes about ten years after cessation until the risk of cardiovascular disease for former smokers reaches the same level as for non-smokers (Oza et al., 2011). Alcohol consumption, another risk factor, if enjoyed in excess (Marmot and Brunner, 1991; Murray et al., 2002), only started to decline after 1980 as Figure 4.8 reveals. Therefore, the positive effects of behavioral changes probably started only as early as the 1980s. Public health education, however, also had its limits. Even though the health risks were known, physical activity declined between 1970 and 2010, while the share of obese persons doubled (Flegal et al., 1998; CDC, 2001; Kohl and Cook, 2013).

The unexpected and concentrated surge of medical breakthroughs in the 1960s and the massive preventive efforts thereafter motivate a pre-treatment period until 1960 and post-treatment from 1970 onward in the estimation sample at hand. The next section discusses the empirical framework and the data.

4.3 Data and Empirical Framework

4.3.1 Data

The empirical analysis is based on a balanced ten-year panel of the 48 contiguous states of the U.S. for the period 1940–2000. Correspondingly, the estimation sample comprises 336 observations in total. Alaska and Hawaii are excluded because of missing data for early periods, the District of Columbia is omitted due to its special nature of a metropolitan region. Because life expectancy in 1940 is only available for whites, the entire sample is restricted to the white population.[6]

Data on gross wages, labor market outcomes, and educational attainment stem from individual data in decennial U.S. censuses (IPUMS) by Ruggles et al. (2015). Wages are adjusted for inflation and measured in logarithms. The variable comprises wages, salaries, commissions, cash bonuses, tips, and other money income received from the employer. Labor market outcomes cover individual labor force participation; usual hours worked per week; usual weeks worked per year; and usual hours worked per year, which are constructed by multiplying weekly hours with work weeks. Working weeks are not available as continuous measures in 1960 and 1970, while the series on usual working hours only starts in 1980. Intervalled hours and weeks, however, are available throughout all time periods. For this reason, I construct a continuous measure for weeks and hours from bivariate regressions of the continuous on the intervalled measure. For the cohort-specific analysis, hours and weeks are constructed based on age-specific regressions.[7] The share of college graduates is constructed from the number of individuals, who attended at least four years of college in their life relative to the entire number of individuals in the sample.[8]

[6] Table C.1 in the Appendix reports descriptive statistics for age-specific groups and the total workforce.

[7] This procedure will lead to downward-biased standard errors in the labor supply regressions, because the missing data points are replaced by fitted values from the corresponding regressions. The respective estimates, however, reveal no significant (positive) effect of adult life expectancy on labor supply such that this bias does not translate to inference.

[8] Results are qualitatively and quantitatively unchanged, if educational attainment contains all individuals who enrolled in college for at least one year.

For representativity, the data are collapsed to the state level using person-sample weights.

Data on life expectancy are obtained from the U.S. decennial life tables and vital statistics provided by the National Center for Health Statistics (2017b) of the Centers for Disease Control and Prevention (CDC). Adult life expectancy enters the estimation equation in logarithms.[9] In 1960, adult life expectancy differed considerably between U.S. states as shown by Panel (a) of Figure 4.6. Notably, white life expectancy at age 50 varied by 2.48 years between Florida, the state with the highest value, and Nevada, the state with the lowest value. Moreover, adult life expectancy was high in the West North Central and West South Central census regions, whereas it was comparatively low in New England and the Atlantic regions.

In order to capture the exogenous increase in adult life expectancy due to innovations in medical technology, the analysis exploits state differences in mortality from cardiovascular diseases prior to their introduction, that is, in 1960, as instrument for adult life expectancy. Age-adjusted cardiovascular mortality in 1960 is obtained from Grove and Hetzel (1968) and expressed in deaths per 100 whites.[10] Panel (b) of Figure 4.6 illustrates spatial differences in the prevalence of cardiovascular diseases as measured by mortality in 1960. The data reveal a strong negative unconditional correlation between adult life expectancy and mortality from cardiovascular diseases: Life expectancy at 50 was high in the census regions, where mortality rates were comparatively low, and vice versa. As shown in Section 4.2, age-adjusted mortality from coronary heart disease did not decrease until shortly before 1970—in fact, it even slightly increased between 1950 and 1968. Only from this point on, mortality from coronary heart disease decreased substantially.[11] For the baseline specification, innovations in medical technology are thus coded to occur from 1970 onward. This designates the time intervals 1940–1960 as pre-treatment and 1970–2000 as post-treatment periods ('differences-in-differences model'). In a more flexible specification, mortality from cardiovascular diseases in 1960 is interacted with a full set of year dummy variables ('flexible model').

Importantly, adult life expectancy and mortality rates provide a conservative view on the effect of health improvements on wages, because they cannot fully capture morbidity reductions following the cardiovascular revolution. In the absence of better health measures across U.S. states and time, they nonetheless represent the best option.

[9] Results are qualitatively and quantitatively similar, if, instead of a log specification, life expectancy enters the estimation framework directly.

[10] Age-adjustment allows to compare the mortality rates between states even if they have different age structures. Due to the adjustment, the mortality rates should not be interpreted as crude rates, unless a state exhibits the same age structure as the standard population. For this reason, not the absolute figures of mortality from cardiovascular diseases are of importance but the relative change over time.

[11] Declines in mortality rates and, thus, improvements in adult life expectancy slightly lag behind the actual development for the average person, because medical innovations might come too late for the very ill and the very old.

In order to ensure that initial mortality from cardiovascular diseases is as good as randomly assigned, further controls interacted with the treatment indicator are added. These controls comprise initial life expectancy by the CDC; initial income and the initial share of college graduates, both obtained from IPUMS by Ruggles et al. (2015); initial population density by Hobbs and Stoops (2002); and initial mortality from non-cardiovascular diseases by Grove and Hetzel (1968). Current values of these variables are not included, because they might themselves be affected by treatment and thus constitute bad controls.

Due to the collapsing process, wages, education, and labor supply are grouped on the state level. For this reason, I weight all regressions by the group size; that is, the initial white population of a specific age cohort or of the total population.[12]

Finally, I further exploit the Health and Retirement Study (HRS, 2017) to investigate whether improvements in adult health contributed to wage increases. This data set provides representative, individual longitudinal data on income and health status for more than 20,000 people over age 50 in the United States. Data on income comprise wages, salaries, bonuses, overtime pay, commissions, tips, second jobs, military reserve earnings, professional practice, or trade income, and refer to the previous year. Values are adjusted for inflation using the annual urban Consumer Price Index by the Bureau of Labor Statistics and measured in logarithms. Individual health is proxied by binary indicators, which indicate whether study participants have ever been diagnosed with heart problems or high blood pressure.

4.3.2 Empirical Framework

This section introduces the empirical framework to study the effect of changes in adult life expectancy on wages per worker and labor market outcomes. The structural model reads

$$y_{s,t} = \alpha x_{s,t} + w_s' \mathcal{I}_t^{1960} \beta + \gamma_s + \delta_t + \zeta_{r,t} + \varepsilon_{s,t}, \tag{4.1}$$

where $y_{s,t}$ denotes the outcome measure (for example, wages) for state s and time period t; $x_{s,t}$ represents log life expectancy of the age group under consideration; w_s is a vector of controls measured in 1960, interacted with the indicator matrix $\mathcal{I}_t^{1960}$, whose values take unity from 1970 onward and zero else; γ_s and δ_t denote state-fixed and time effects; $\zeta_{r,t}$ describes region-year interactions, which control for differential development trends across the nine U.S. census regions r; and $\varepsilon_{s,t}$ constitutes an idiosyncratic error term.

Due to omitted variables and reverse causality, log life expectancy is likely endogenous. In order to uncover the causal link between adult health and average wages, I exploit heterogeneity in the prevalence of cardiovascular diseases across U.S. states as exogenous

[12]Since the population equation of interest is the effect of improved health conditions on individual wages and labor supply, weighting the regression equation by the group size yields estimation results that are closer to the micro data than unweighted averages. See, for example, Angrist and Pischke (2009).

source of variation for instrumentation. The first-stage equation is given by

$$x_{s,t} = \eta z_s d_t^{1960} + w_s' \mathcal{I}_t^{1960} \theta + \kappa_s + \lambda_t + \mu_{r,t} + \xi_{s,t}, \tag{4.2}$$

where mortality from cardiovascular diseases in 1960, z_s, is interacted with the post-treatment indicator d_t^{1960}; κ_s and λ_t denote state-fixed and time effects; $\mu_{r,t}$ describes census-region-year effects; and $\xi_{s,t}$ constitutes the error term.

Conceptually, the first-stage equation compares differences in the increase of adult life expectancy to differences in the decline of mortality from cardiovascular diseases between the pre-treatment and post-treatment period across states. For this reason, it corresponds to a differences-in-differences approach, where all states are treated but with different treatment intensities. Moreover, the first stage has a natural interpretation in this context: A decline in the mortality rate from cardiovascular diseases initiates an increase in adult life expectancy, which, in turn, affects the economic outcomes of interest in the structural model.

For initial mortality from cardiovascular diseases to be a valid instrument, several conditions must be fulfilled. First, initial mortality must be as good as randomly assigned conditional on covariates. This assumption requires the instrument (initial cardiovascular mortality interacted with the treatment indicator) to be independent of potential outcomes and potential treatment assignments, given the complete set of covariates. To this end, the baseline specification contains controls for initial state levels of income, the share of college graduates, and population density. These controls take up state-level selection toward more health, which is attributable to disparities in income, educational attainment, or the rural-urban discrepancy between densely populated states at the coasts and spacious states in the middle of the country. Moreover, initial non-cardiovascular mortality and initial life expectancy control for the health environment prevailing before new medical treatment technologies for cardiovascular diseases were introduced. Finally, state-fixed, time, and census-region-year-fixed effects eliminate systematic state- and region-level variation due to differences in further social, cultural, or environmental factors that do not vary concomitantly over states and time, or that possess time-varying influences on mortality from cardiovascular diseases. In particular, these trends cancel out differentials between census regions.

Furthermore, mortality from cardiovascular diseases must affect outcomes only through the first stage; that is, through the channel of health and longevity. Accordingly, mortality from cardiovascular diseases is not part of the structural model. This exclusion restriction is fundamentally untestable. It would be violated, if changes in cardiovascular mortality rates were to affect the outcome of interest through a channel other than adult health as measured by adult life expectancy. Because the instrument is specific to the channel of health on the aggregate level, however, this assumption should plausibly be fulfilled in the context of this paper. Additionally, the empirical model accounts for initial differences in non-cardiovascular mortality to prevent the instrument from taking up beneficial effects attributable to medical advancement in the treatment of other diseases.

Finally, changes in mortality from cardiovascular diseases must be predictive of changes in adult life expectancy. Sufficiently high values of the first-stage F-statistics demonstrate that this assumption is fulfilled for the differences-in-differences model.

I also report results for a flexible model, in which the instrument is interacted with year dummies instead of the post-1960 treatment indicator. In this case, the first stage corresponds to

$$x_{s,t} = \sum_{\tau=1940}^{2000} \eta_\tau z_s d_t^\tau + w_s' \mathcal{I}_t^\tau \theta + \kappa_s + \lambda_t + \mu_{r,t} + \xi_{s,t}. \tag{4.3}$$

4.4 Results

This section presents the empirical results. First, I report evidence on the first-stage correlation between mortality from cardiovascular diseases and adult life expectancy as well as evidence on the reduced form effect of mortality on average wages. After having statistically established these relationships, I show results from two-stage least squares estimates for wages by age group and the total workforce.[13] Finally, this section concludes by investigating potential channels for a causal positive link between health improvements and wages.

4.4.1 First-Stage Evidence: Mortality and Adult Life Expectancy

First, I investigate the first-stage association between mortality from cardiovascular diseases and adult health conditions, proxied by log life expectancy at age 50. The analysis is based on the ten-year panel of the 48 contiguous U.S. states from 1940 to 2000, described in Section 4.3.1, with pre-treatment periods 1940–1960 and post-treatment periods 1970–2000. Table 4.1 reports least squares results for the differences-in-differences model from (4.2) in Panel (a) and for the flexible model from (4.3) in Panel (b).[14]

Column (1) shows results without covariates. In this case, initial mortality from cardiovascular diseases (interacted with the treatment indicator) and life expectancy at 50 correlate positively. This result is, however, driven by the omission of initial life expectancy. Based on how life tables are constructed, mortality rates and life expectancy must correlate negatively. Moreover, given better initial health conditions, there is less scope for future reductions in the mortality rates and, consequently, less potential for future improvements in life expectancy. Correspondingly, changes in life expectancy and initial life expectancy

[13] Table C.2 in the Appendix reports estimates from ordinary least squares (OLS).

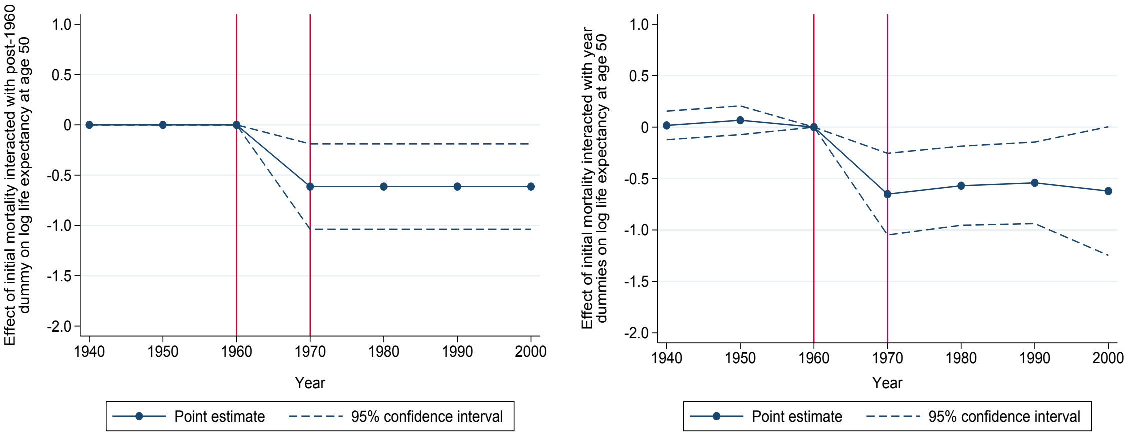
[14] Because life expectancy at age 50 on average provides the most accurate picture of health for workers around age 45–54, the sample is weighted by the initial white population of the 45- to 54-year-olds. The results are unaltered, if instead weighted by the entire white population.

Table 4.1: First Stage: Effect of Mortality on Adult Life Expectancy

	Dependent variable: log life expectancy at age 50				
	(1)	(2)	(3)	(4)	(5)
	(a) Differences-in-differences model				
Mortality × Post 1960	0.50***	-0.53***	-0.86***	-0.77***	-0.61***
	(0.11)	(0.15)	(0.17)	(0.20)	(0.21)
	(b) Flexible model				
Mortality × 1940	-0.13*	-0.13*	-0.15*	0.04	0.02
	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)
Mortality × 1950	-0.05	-0.05	-0.05	0.10	0.07
	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)
Mortality × 1970	0.21***	-0.82***	-1.16***	-0.78***	-0.65***
	(0.05)	(0.16)	(0.17)	(0.19)	(0.20)
Mortality × 1980	0.34***	-0.69***	-1.03***	-0.70***	-0.57***
	(0.07)	(0.15)	(0.17)	(0.19)	(0.19)
Mortality × 1990	0.50***	-0.54***	-0.88***	-0.64***	-0.54***
	(0.09)	(0.16)	(0.18)	(0.19)	(0.20)
Mortality × 2000	0.71***	-0.32*	-0.65***	-0.75***	-0.62*
	(0.15)	(0.18)	(0.21)	(0.25)	(0.31)
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Estimates are weighted by the initial white population of 45- to 54-year-olds. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

also correlate negatively. Therefore, the respective estimate is biased upward. Once, initial life expectancy is included as additional control for Columns (2) to (5), the sign turns negative. Accordingly, the larger the decline of cardiovascular mortality, the higher the gain in life expectancy. Column (3) adds initial mortality from non-cardiovascular diseases as control, which improves the fit of the first stage. Furthermore, this measure precludes that the instrument takes up health improvements that cannot be attributed to the cardiovascular revolution. This leads to a slightly more negative point estimate. Column (4) adds region-year-fixed effects that eliminate systematic trends reflecting economic, social, or cultural differences across U.S. census regions. Finally, the full specification in Column (5) adds additional controls for the initial share of college graduates, initial population



(a) DD Model (Table 4.1a, Column 5) (b) Flexible Model (Table 4.1b, Column 5)

Figure 4.9: Illustration: First Stage

density, and initial income. These variables cancel out any variation in life expectancy originating from initial differences in education and development. The corresponding point estimate takes a value of –0.61 and is significant at the one-percent level. Given the quasi-natural source of variation, the parameter has a quantitative interpretation: A reduction of mortality from cardiovascular diseases of one person per 1,000 whites leads to an increase in white life expectancy at 50 of approximately 6.1 percent. Taken at face value, the reduction in cardiovascular mortality by two persons per 1,000 whites (50 percent of its initial value) between 1960 and 2000 thus led to an increase in life expectancy at 50 of approximately 3.16 years, or two thirds of the overall increase over this time period.[15] This number conforms closely to the increase in life expectancy of 3.27 years, which, according to Cutler, Deaton, and Lleras-Muney (2006), can be attributed to medical advancement in the treatment of cardiovascular diseases between 1960 and 2000.

The flexible model in Panel (b) reports point estimates for the instrument interacted with year dummies for every period and the year 1960 as reference category. Mortality affects life expectancy negatively in the post-treatment period 1970 to 2000 for all specifications that control for initial longevity. In Column (5), the parameters are estimated to be of similar quantity and significantly different from zero. In contrast, there is no effect of mortality from cardiovascular diseases in 1960 on life expectancy for the pre-treatment periods 1940 to 1960.[16] Figure 4.9 plots the point estimates from the full specification in Column (5) with the corresponding 95-percent confidence interval for both models. The displayed coefficients of the flexible model in Panel (b) show a stable pattern for the pre-treatment and post-treatment period such that the assumption of a constant effect for each period in the differences-in-differences model appears appropriate.

[15] To arrive at these figures, compute $\Delta_x = \widehat{\eta} \cdot \Delta_z \cdot \bar{\mu}_x = (-0.61) \cdot (-0.20) \cdot 25.91 \approx 3.16$, where $\bar{\mu}_x$ is evaluated at the sample mean in 1960, and $\Delta_x/(x^{2000} - x^{1960}) = 3.16/4.76 \approx 0.66$.

[16] Single outliers do not drive these partial correlations as Figure C.1 in the Appendix shows.

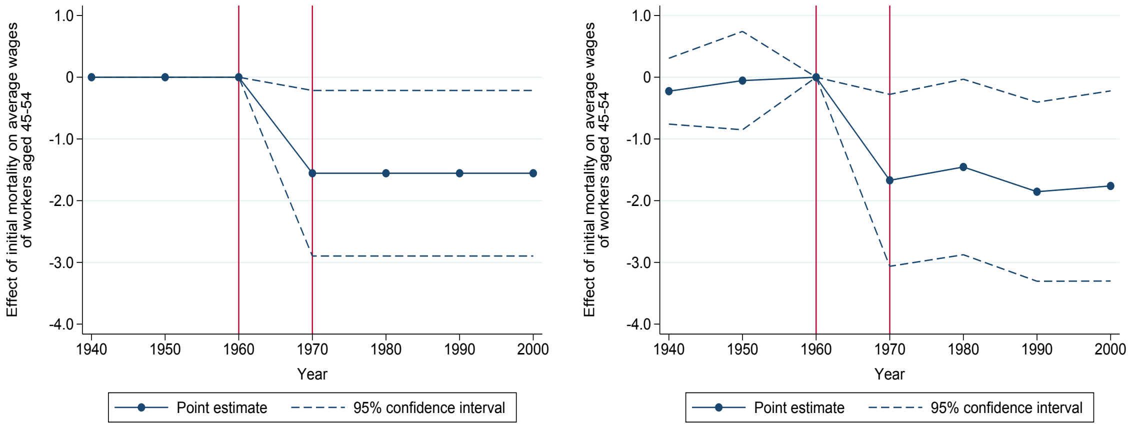
4.4.2 Reduced-Form Evidence: Mortality and Aggregate Wages

Table 4.2 reports reduced form estimates for the effect of mortality on wages for the 45- to 54-year-olds. This age group is of particular interest for two reasons. On the one hand, these workers are usually considered the most productive group of the workforce for their high participation rate and their considerable experience. Therefore, these workers are usually at the peak of their life-cycle earnings profile as illustrated in Figure 4.1. On the other hand, they become increasingly susceptible to cardiovascular diseases due to aging and behavioral risk factors, while still being young enough to profit quite considerably from new treatment possibilities and changed behavior. For these reasons, this is one of the age groups, which might profit from medical innovations in terms of both, health and

Table 4.2: Reduced Form: Effect of Mortality on Average Wages of Workers Aged 45–54

	Dependent variable: log wages of whites 45–54				
	(1)	(2)	(3)	(4)	(5)
	(a) Differences-in-differences model				
Mortality × Post 1960	-0.87***	-1.20**	-1.50***	-1.39***	-1.56**
	(0.21)	(0.45)	(0.53)	(0.49)	(0.67)
	(b) Flexible model				
Mortality × 1940	0.50*	0.50*	0.43	-0.43**	-0.23
	(0.27)	(0.27)	(0.27)	(0.20)	(0.27)
Mortality × 1950	0.38	0.38	0.34	0.23	-0.05
	(0.36)	(0.36)	(0.35)	(0.41)	(0.40)
Mortality × 1970	-0.42***	-0.76**	-1.06**	-1.39***	-1.67**
	(0.13)	(0.38)	(0.42)	(0.47)	(0.69)
Mortality × 1980	-0.85***	-1.18***	-1.50***	-0.98**	-1.45**
	(0.23)	(0.41)	(0.44)	(0.48)	(0.71)
Mortality × 1990	-0.42	-0.75	-1.09*	-1.70***	-1.85**
	(0.33)	(0.55)	(0.63)	(0.58)	(0.72)
Mortality × 2000	-0.60*	-0.93*	-1.30**	-1.73***	-1.76**
	(0.35)	(0.54)	(0.62)	(0.62)	(0.77)
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Estimates are weighted by the initial white population of 45- to 54-year-olds. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.



(a) DD Model (Table 4.2a, Column 5) (b) Flexible Model (Table 4.2b, Column 5)

Figure 4.10: Illustration: Reduced Form

economic outcomes.[17]

Panel (a) shows the results for the differences-in-differences model. As for the first stage, additional controls are included for initial values of life expectancy, mortality from non-cardiovascular diseases, region-year-fixed effects, the share of college graduates, population density, and income. The parameter estimate in the full specification in Column (5) takes a value of –1.56 and is significant at the five-percent level.[18] Given that the instrument is as good as randomly assigned, the coefficient estimate reflects the intention-to-treat effect. It measures the effect of being offered the treatment. Since not all individuals decide to take up the treatment (for example, some patients do not take a beta blocker, although they belong to high-risk groups for stroke or cardiac arrest), the intention-to-treat effect is too low relative to the average treatment effect on the treated (Angrist and Pischke, 2009). According to the point estimate, a reduction of mortality from cardiovascular diseases by one person per 1,000 whites leads to a wage increase of 15.6 percent for the group of the 45- to 54-year-olds.

In the full specification of the flexible model in Panel (b), there is no significant effect of cardiovascular mortality on average wages in the pre-treatment period, though the effect is significantly negative in the post-treatment period.[19] Figure 4.10 depicts the reduced-form estimates for the differences-in-differences and the flexible model. The flexibly-estimated coefficients show again a stable pattern over time giving credibility to the simpler differences-in-differences model.[20]

[17]Table C.3 in the Appendix reports qualitatively similar results for the entire workforce.
[18]The sample is again weighted by the initial population of the 45- to 54-year-olds. The resulting coefficient is quantitatively similar, if weighted by the entire white population.
[19]Single outliers do not drive these partial correlations as Figure C.2 in the Appendix shows.
[20]Figure C.3 in the Appendix plots the corresponding reduced form parameters for the entire workforce.

4.4.3 Life Expectancy and Aggregate Wages

The last two sections have established the existence of a first-stage correlation between mortality from cardiovascular diseases and adult life expectancy, and a reduced-form effect of mortality on wages for the 45- to 54-year-olds. Now, I turn to the average treatment effect on the treated, which corresponds to the ratio of the intention-to-treat effect from the reduced form, and the first-stage estimand, which corresponds to the compliance rate. Using two-stage least squares (2SLS), this quantity can also be directly estimated by instrumenting log life expectancy at age 50 with mortality from cardiovascular diseases in 1960 interacted with the treatment indicator. Table 4.3 reports the estimated effect of life expectancy on wages of the 45- to 54-year-olds.[21] Section 4.4.4 discusses the corresponding estimates for all age groups and the entire workforce.

The first column of Panel (a) shows estimates for the differences-in-differences model without any additional controls except state-fixed and time effects. In this case, an increase in life expectancy leads to a decline in wages of workers aged 45 to 54. As argued by Aghion, Howitt, and Murtin (2011) and Bloom, Canning, and Fink (2014), however, this specification is misspecified, because it omits initial life expectancy. In particular, initial life expectancy correlates with initial mortality from cardiovascular diseases and subsequent improvements in life expectancy. Furthermore, it concomitantly affects prospective wage gains. For a given reduction in mortality rates, the first stage, therefore, underestimates the corresponding improvement in life expectancy; in fact, the model suggests smaller improvements in life expectancy for states with higher initial prevalence of cardiovascular diseases. In addition, the reduced form underestimates the associated wage gains following the health improvements of the cardiovascular revolution. In combination, the omission of initial life expectancy results in a downward bias of the estimates. Correspondingly, initial life expectancy is included in all remaining specifications. The third column adds initial mortality from non-cardiovascular diseases to improve the fit of the first stage, leading to an increase of the Kleibergen-Paap F-statistic from 26.6 to 44.6. Moreover, this control prevents the mortality instrument from taking up beneficial effects of medical innovations that work through health channels other than the cardiovascular revolution. The last two columns additionally contain region-year-fixed effects that take up differential trends in wages and life expectancy across U.S. census regions.

Finally, the full specification in Column (5) adds initial values of the share of college graduates, population density, and income—all interacted with the post-1960 treatment indicator. Because there is a strong link between education and health (Grossman and Kaestner, 1997; Lleras-Muney, 2005), the initial share of college graduates is added to take up variation of life expectancy that is attributable to initial disparities in state-level education. Population density is included to account for the rural-urban gradient in

[21] The sample is weighted by the initial white population of the 45- to 54-year-olds.

Table 4.3: Adult Life Expectancy and Average Wages of Workers Aged 45–54

	Dependent variable: log wages of whites 45–54				
	(1)	(2)	(3)	(4)	(5)
	(a) Differences-in-differences model (2SLS)				
Log life expectancy at 50	-1.73***	2.26***	1.73***	1.81***	2.54***
	(0.40)	(0.84)	(0.61)	(0.57)	(0.95)
First-stage F–statistic	62.5	26.6	44.6	36.6	14.2
	(b) Flexible model (2SLS)				
Log life expectancy at 50	-1.36***	0.69	0.75**	1.73***	2.35***
	(0.37)	(0.44)	(0.38)	(0.53)	(0.89)
First-stage F–statistic	15.3	13.2	19.9	7.4	3.5
Hansen test p–value	0.09	0.03	0.04	0.2	0.8
	(c) Flexible model (LIML)				
Log life expectancy at 50	-1.47***	0.92	0.91**	2.05***	2.57**
	(0.41)	(0.57)	(0.44)	(0.66)	(1.00)
First-stage F–statistic	15.3	13.2	19.9	7.4	3.5
Hansen test p–value	0.09	0.03	0.05	0.2	0.8
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Estimates are weighted by the initial white population of 45- to 54-year-olds. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

cardiovascular diseases with a particularly high prevalence of coronary heart disease and hypertension in non-metropolitan areas (Pickle and Gillum, 1999; Cooper et al., 2000). Finally, initial income is added to account for a potential feedback effect from income to health (Ettner, 1996; Frijters, Haisken-DeNew, and Shields, 2005; Lindahl, 2005; Chetty et al., 2016). The corresponding coefficient estimate takes a value of 2.54 and is significant at the one-percent level. Therefore, an increase in life expectancy at age 50 by one percent, ceteris paribus, causes a wage hike of 2.54 percent for the group of the 45- to 54-year-olds. Thus, the average treatment effect on the treated is approximately 60 percent larger than the intention-to-treat effect from the reduced form. Taken at face value, the increase in life expectancy at 50 between 1960 and 2000 led to a hike in average gross wages of the 45- to 54-year-olds of approximately 9,762$ per year, or around 47 percent of the overall increase over this time period.[22] The first-stage F-statistic shows a value of 14.2, which indicates a sufficiently strong instrument given the conventional cutoff level of 10.

Panels (b) and (c) present results for the flexible model, estimated with two-stage least squares and with a heteroskedasticity-robust version of limited information maximum likelihood (LIML) due to the small value of the F-statistic. The full specification in Column (5) reports point estimates that are quantitatively similar to those of the differences-in-differences model. Since the p-value of the Hansen test for overidentification takes a value of 0.8, the null hypothesis that all instruments provide the same information is maintained.

4.4.4 Heterogeneity Across Age Groups and Measurement

Before turning to the effect of adult life expectancy for other age groups and the overall workforce, consider measurement of health improvements in light of the demographic structure of the population; a point that has so far largely gone unnoticed by the literature. The convention is to use a specific measure to capture the effect of a positive health shock, for example, life expectancy at a given age. By construction, this measure encompasses the expected remaining lifetime including all following age groups. Thus, it provides a gross approximation of expected health over the remaining part of the life-cycle. Accordingly, one implicitly assigns the same health to all individuals in the population of interest. This assumption is overly restrictive and masks heterogeneous effects across age cohorts. Moreover, it may introduce a systematic measurement error into the empirical model that cannot be solved by the instrumentation strategy.

This deficit becomes clear in light of the cohort structure of the labor force: On average, young workers have a relatively good health and can still expect to live a high number of years, while health of older workers is lower due to aging such that their remaining expected years of life are considerably smaller than for young workers. A measure that predominantly captures the health of young workers would thus overstate the health of older workers, and vice versa. Hence, there would be systematic measurement error.

Suppose, for example, one is interested in estimating the effect of health changes on average wages of the total workforce. In order to capture improvements in the health of the workforce, one might use the (log-) change in life expectancy at age 30. By assumption, every worker is assigned the expected health improvement of a thirty-year-old. This measure may over- or understate the average health improvement, depending on how the

[22]To arrive at these figures, compute $\Delta_y = \widehat{\alpha} \cdot \Delta_x / x^{1960} \cdot \bar{\mu}_y \approx 2.54 \cdot 3.16/25.91 \cdot 31515.35 \approx 9762.82$ with $\Delta_y/(y^{2000} - y^{1960}) \approx \frac{9762.82}{20666.62} \approx 0.47$, and $\bar{\mu}_y$ evaluated at the mean in 1960. The corresponding estimates based on the reduced-form estimate are 5,996$, or around 29 percent of the overall wage hike. Table C.4 in the Appendix reports similar results for a model in which life expectancy enters linearly.

health shock under consideration affects the average health status of the different age groups. Hence, one does not use the exact measure $x_{s,t}$ in the empirical framework but

$$p_{s,t} = x_{s,t} + \nu_{s,t}, \tag{4.4}$$

where $p_{s,t}$ is the observed proxy in the sample and $\nu_{s,t}$ is a measurement error. Whether this measure correctly captures the average health improvement of the workforce depends on the demographic structure. For example, in the case of the cardiovascular revolution, health gains concentrated among older adults with gains in life expectancy increasing with age. If, without loss of generality, the number of young workers is small, the average worker is older than 30. Therefore, the measure assigns a too pessimistic figure of the health improvement to the workforce for all observational units s with the extent of the error depending on state-level variation in the demographic structure of the workforce. Hence, the health proxy $p_{s,t}$ and the measurement error $\nu_{s,t}$ correlate negatively. Plugging the expression from (4.4) into the regression model yields

$$y_{s,t} = \alpha p_{s,t} + w_s' \mathcal{I}_t^{1960} \beta + \gamma_s + \delta_t + \zeta_{r,t} + \varepsilon_{s,t} - \alpha \nu_{s,t}. \tag{4.5}$$

Define $e_{s,t} = \varepsilon_{s,t} - \alpha \nu_{s,t}$ as the composite error term and suppose that there is no correlation between the proxy $p_{s,t}$ and the idiosyncratic error $\varepsilon_{s,t}$. Given a non-negative α and the negative correlation between $p_{s,t}$ and $\nu_{s,t}$ due to systematic error, the health proxy $p_{s,t}$ must correlate positively with the composite error. Hence, the point estimate $\widehat{\alpha}$ for this model will be biased upward. The same logic applies to a too optimistic measure of average health with the only difference that $p_{s,t}$ and $\nu_{s,t}$ correlate positively in this case. Accordingly, systematic mismeasurement of the health status leads to downward-biased estimates of the population parameter, if the health proxy overstates gains in average health, while estimates are biased upward, if the health proxy understates the improvement in average health. The more proxy and true health diverge, the more severe this bias will be.

For the age-group analysis, this problem can be solved by using age-specific life expectancy as right-hand variable. Panel (a) of Table 4.4 reports the corresponding results for the differences-in-differences model.[23] Columns (1) to (5) show results for age cohorts from 15 to 64 in ten-year intervals. Column (6) reports the estimated effect of adult life expectancy on wages for old-age workers above the age of 65. Finally, Column (7) provides the parameter estimates for the entire workforce.[24] To this end, life expectancy for the total workforce is approximated by the arithmetic mean over life expectancy at birth and all following age cohorts, thus providing a health indicator closer to the average person of

[23] Table C.5 in the Appendix presents estimates for the flexible model.

[24] Regressions are weighted by the initial white population of each specific age cohort in 1960 in the first six columns and by the entire white population for the last column.

Table 4.4: Adult Life Expectancy and Average Wages by Age Cohorts

	Differences-in-differences model (2SLS)						
	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Life expectancy of specific age group						
Log life expectancy	3.66	4.18***	2.82***	2.54***	1.09	5.21***	3.62***
(of specific age group)	(3.10)	(1.61)	(0.88)	(0.95)	(0.95)	(1.98)	(1.09)
First-stage F–statistic	10.8	22.5	46.9	14.2	26.0	16.0	31.7
	(b) Naive model: life expectancy at age 30						
Log life expectancy at 30	1.19	4.18***	4.30***	4.46***	2.26	11.47***	3.67***
	(1.77)	(1.61)	(1.48)	(1.53)	(1.66)	(3.73)	(1.31)
First-stage F–statistic	24.0	22.5	21.4	21.9	22.5	24.3	23.3
	(c) Naive model: life expectancy at birth						
Log life expectancy at birth	0.72	3.79**	3.46**	4.15**	3.15	15.26**	4.00**
	(1.78)	(1.73)	(1.64)	(1.77)	(2.06)	(6.10)	(1.56)
First-stage F–statistic	11.6	11.4	10.4	10.0	9.6	9.3	11.0
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

the workforce.[25]

The effect of life expectancy is positive and significant at the five-percent level for the prime-age groups of the working-age population from 25 to 54 and old-age workers above 65. A one-percent increase of life expectancy, ceteris paribus, induces a wage hike of roughly 3.5 percent with coefficients ranging from 2.54 to 5.21. The values of the Kleibergen-Paap F-statistic show a strong first-stage correlation between adult life expectancy and mortality from cardiovascular diseases. For very young workers, the effect of life expectancy on wages is large and positive but insignificant due to high standard errors. A potential explanation for this finding is that positive effects of health improvements are counteracted by an out-selection of the most productive workers into college. Moreover, the explanatory power of the instrument is weakest among all age groups with a value of the F-statistic of 10.8. Likewise, for the group of the 55- to 64-year-olds, health improvements do not significantly affect wages. The effect, however, is considerably stronger for those workers

[25] Similar results obtain for life expectancy at age 30 or 40 as health measure for the total workforce.

who decide to work even after age 65. One reason for this finding might be selection into retirement starting around age 60. In order to save enough for retirement, workers who optimally would like to retire early due to bad health or outdated human capital continue to work, thus lowering productivity for the 55- to 64-year-olds. Above age 65, only the healthy, motivated, and productive workers remain in the workforce: The overall labor force participation rate drops to slightly below 20 percent for this group. These workers are also likely those who gain most from improvements in health innovations. Hence, life expectancy shows the largest effect on wages with a coefficient of 5.21. Taken together, this evidence indicates a sizable positive effect of adult life expectancy on wages for workers in their prime-age and those above 65. Graphically, this translates into a shift toward steeper life-cycle wage profiles consistent with the unconditional evidence presented in Figure 4.1.

Panel (b) reports estimation results for a naive version of the differences-in-differences model, where life expectancy at age 30 is used for every age cohort. As outlined above, using a mismeasured proxy for health conditions leads to downward-biased estimates for age groups whose health gain is overstated by the measure—that is, the 15- to 24-year-olds—and upward-biased estimates for age groups, whose health gains are understated—that is, all groups above age 35. Correspondingly, the resulting point estimates would suggest a too large effect of life expectancy on average wages for older workers. For the 25- to 34-year-olds, the model is identical to Panel (a). Panel (c) repeats this exercise for life expectancy at birth, which is the most common health indicator in cross-country studies. For the age groups above 35, where changes in life expectancy at birth understate the actual change, the point estimates are overestimated. In contrast, for the age group 25 to 34, the results show slightly smaller effects of health innovations on average wages compared to the results in Panel (a), although gains in life expectancy are quantitatively similar for this age group. Because the first stage of life expectancy at birth is considerably weaker compared to Panel (a), this finding might be due to a loss of precision in the estimation. Compared to the more adequate specifications in Panel (a), the naive model with life expectancy at birth suggests a too large effect of health gains on average wages, because the health gains are relatively larger at higher ages.

Importantly, the instrumentation strategy cannot eliminate this measurement error, because the correlation between adult life expectancy and mortality depends on age. If measured at higher ages, gains in life expectancy do not contain health improvements resulting from reduced mortality at younger ages. In contrast, if life expectancy is measured at a too young age, the instrument assigns individuals beneficial effects from health innovations that do not apply to them due to their age. Therefore, the first stage again systematically over- or understates the average health improvement, if the wrong age-specific proxy for health conditions is assigned.

This subtle point is of significant practical importance for a large number of published work. For example, a branch of the growth literature has investigated the effect of health and life expectancy on long-run growth of output per capita. The conventional measure employed in these studies is life expectancy at birth. Long-run changes in output are thus explained by improvements of health conditions of infants, which might overstate the improvements for the workforce. In such a case, the corresponding point estimates would be biased downward. How severe the bias from measurement error is, depends on how well changes in life expectancy at birth capture changes in the health status of the workforce. For example, if improvements in medical technology mostly help infants and young children, as it was the case for the epidemiological transition, the gains in life expectancy at birth and for higher ages will differ greatly. In this case, the bias is most pronounced. This might explain why Acemoglu and Johnson (2007), Acemoglu and Johnson (2014), Bloom, Canning, and Fink (2014), and Hansen (2014) have not found a substantial positive effect of health and life expectancy on growth for the reduction in mortality from infectious diseases. If, in contrast, increases in life expectancy at birth mostly reflect improvements in health at older ages, as it is case for the cardiovascular revolution, the bias should be comparatively small, because changes in life expectancy at birth provide still a reasonable approximation of improvements in adult health.

4.4.5 Accounting for Inter-State Migration

The analysis so far has investigated the causal relationship between adult life expectancy and average wages, while treating states as closed entities. A potential concern relates to workforce migration between states. About 1.5 percent of the total U.S. population move between states per year and one third of the citizens do not live in the state, where they were born (Molloy, Smith, and Wozniak, 2011). If, on average, high productivity workers migrate into states, where life expectancy is higher, parameter estimates might be biased upward. In order to address this problem, individuals in the census data are dropped, if they do not live in the state, where they were born.[26] Panel (a) of Table 4.5 presents the respective results of the differences-in-differences model for state-level regression.[27]

The corresponding point estimates indicate a strong positive effect of increased life expectancy on average wages, which is statistically significant at the five-percent level, except for the age group of the 15- to 24-year-olds. For all age groups above 25, the resulting coefficients are quantitatively larger than in the baseline model in Table 4.4. The evidence from the non-migrant sample therefore conflicts with an upward bias due to migrant workers, unless indirect effects owing to the complementarity between domestic and migrating workers distinctly outweigh the direct effects. Hence, it is unlikely that the considerable positive effect of adult life expectancy on average wages in the baseline model is driven solely by workforce migration, and it appears conservative given the evidence presented in Table 4.5. Moreover, this evidence tends to the concern whether health should

[26] Note, however, that the effect of migrants cannot be deducted from the life expectancy measure.
[27] Estimates for the flexible model are reported in Table C.6 in the Appendix.

Table 4.5: Robustness: Migration

	Differences-in-differences model (2SLS)						
	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) No inter-state migration of workers						
Log life expectancy	2.21	5.02***	4.65***	3.91***	2.56**	6.61**	4.78***
(of specific age group)	(2.76)	(1.51)	(1.05)	(1.24)	(1.01)	(2.73)	(1.18)
First-stage F–statistic	13.4	22.6	37.4	13.3	32.9	15.5	31.0
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓
	(b) No old-age migration						
Log life expectancy	8.22	5.01**	3.11***	2.43**	1.07	5.21***	3.67***
(of specific age group)	(7.48)	(2.40)	(1.12)	(1.01)	(0.83)	(1.94)	(1.19)
First-stage F–statistic	2.7	11.7	33.7	12.4	43.8	18.6	24.2
States	45	45	45	45	45	45	45
Observations	315	315	315	315	315	315	315
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Average wages in Panel (a) are confined to workers, who work in the same state they were born in. Panel (b) excludes Arizona, California, and Florida from the sample. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

be measured by state of birth or state of residence: The results establish a causal link between adult life expectancy and average wages in the non-migrant sample, for which state of birth and state of residence coincide.

Furthermore, a frequent phenomenon is migration after retirement. While in the 1950s, fewer than one million people above age 60 moved from one state to another, the corresponding number had increased to 1.6 million between 1975 and 1980; whereupon, Florida, California, and Arizona were the most popular destination targets (Rogers and Watkins, 1987). High values of life expectancy at higher ages in these states reflect this popularity. If it is predominantly rich pensioners, who move for retirement, old-age migration might act as a positive demand shock to the destination states. If, at the same time, these pensioners are healthier than the average retiree, migration of the elderly might bias upward the estimated effect of health gains on average wages. Panel (b) of Table 4.5 presents results from the state panel without Florida, California, and Arizona. The parameter estimates are quantitatively similar to the baseline results, thus revealing again a positive link between adult life expectancy and wages per worker.[28] Overall, the evidence indicates that old-age migration does not cause a spurious correlation between life expectancy and wages.

4.4.6 Heterogeneity Along Further Dimensions: Metropolitan Areas, Occupational Choice, and Educational Attainment

The analysis so far has investigated the causal relationship between adult life expectancy and average wages on the state level. As noted by Cooper et al. (2000), however, there might be disparities in the prevalence of cardiovascular diseases, in particular, between rural and metropolitan areas that might not be fully taken up by controlling for population density. Therefore, this section analyzes the causal link between life expectancy and wages only for metropolitan areas in the corresponding states. To this end, census data are collapsed on the metropolitan-area level, and each area is assigned the corresponding state-level value of life expectancy.[29] Because metropolitan areas changed over time, only those areas that are consistently defined throughout all time periods from 1940 to 2000 enter the estimation sample. This leaves 623 time-year observations for 89 metropolitan areas in 33 states. Table 4.6 reports results for age groups and the workforce for the differences-in-differences model including a full set of controls.[30]

Table 4.6: Adult Life Expectancy and Average Wages: Metropolitan Areas

	Differences-in-differences model (2SLS)						
	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log life expectancy	9.91**	6.23***	4.09***	5.40***	3.05**	6.89*	6.31***
(of specific age group)	(3.88)	(1.78)	(1.11)	(1.86)	(1.32)	(3.53)	(1.45)
First-stage F–statistic	12.2	37.3	65.1	12.6	25.1	6.3	44.2
States	33	33	33	33	33	33	33
Metropolitan Areas	89	89	89	89	89	89	89
Observations	623	623	623	623	623	623	623
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

[28] Results for the flexible model are reported in Table C.7 in the Appendix.

[29] The motivation for using state-level health measures in this context is twofold: First, data on adult life expectancy are not available below the state level. Second, more disaggregated measures for age-specific life expectancy and disease-specific mortality rates may suffer from low quality, as relatively rare chance events may generate spurious patterns in small populations.

[30] Table C.8 in the Appendix shows results for the flexible model.

Throughout all age groups, the estimated coefficients are larger compared to state-level estimates. According to the workforce estimate in Column (7), a one-percent gain in life expectancy leads to a 4.17-percent increase in average wages per worker. All coefficient estimates are statistically significant at the five-percent level, and the F-statistic indicates a strong first-stage correlation between life expectancy and mortality from cardiovascular diseases. This finding is in line with a rural-urban gap in health improvements related to cardiovascular diseases as, for example, found by Kulshreshtha et al. (2014). One explanation for this divide is that behavioral risk factors such as smoking, drinking, obesity, and physical inactivity are more common in rural areas (CDC, 2017). Another potential reason is that access to treatment for cardiovascular diseases is more readily available in urban areas due to returns to scale. Overall, this evidence suggests that the positive effect of adult life expectancy on average wages is not an artifact of comparing rural with urban states. Moreover, the gains from health innovations are larger in metropolitan areas compared to the state-level estimates.

Occupational choice constitutes another dimension of heterogeneity. Blue-collar workers, on the one hand, execute tasks that are physically demanding, whereas white-collar jobs, on the other hand, require minimal physical labor but usually more investment in educational attainment. With increasing age, blue-collar workers are thus more vulnerable to negative income shocks as a consequence of worsened health status compared to white-collar workers. Hence, innovations in medical technology should benefit blue-collar workers especially at advanced working age. In contrast, higher prospective health encourages potential white-collar workers to invest more in human capital, thus resulting in higher productivity particularly among younger age cohorts.

In order to test these hypotheses, workers in the decennial U.S. census are categorized as blue-collar or white-collar, based on the occupation coding guidelines for the 1970 U.S. census (U.S. Bureau of the Census, 1972, pp. 152–154). Specifically, this coding classifies workers as white-collar, if they belong to the group of professional, technical and kindred workers; managers and administrators except farm; sales workers; or clerical and kindred workers. In contrast, blue-collar occupations comprise craftsmen and kindred workers; operatives except transport; transport equipment operatives; and laborers except farm. The remaining workers belong to farm or service occupations and are exempt from the analysis. Since occupational status was first reported in the U.S. census in 1950, the number of state-year observations decreases to 288.

Table 4.7 presents results for a regression of average wages per worker on adult life expectancy for a subsample consisting of white-collar workers in Panel (a) and blue-collar workers in Panel (b).[31] Adult life expectancy affects average wages of white-collar workers in the age group 25 to 44 positively and significantly, whereas the effect vanishes for the more advanced workers in the age range from 45 to 64. This finding points to better

[31]Table C.9 in the Appendix reports results for the flexible model.

Table 4.7: Heterogeneity: White-Collar and Blue-Collar Workers

	Differences-in-differences model (2SLS)						
	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) White-collar workers						
Log life expectancy	3.83	8.99***	4.08***	1.36	0.05	5.83**	4.51***
(of specific age group)	(6.76)	(3.27)	(1.58)	(1.29)	(0.95)	(2.85)	(1.68)
First-stage F–statistic	2.8	9.8	28.5	8.4	14.6	14.2	18.1
	(b) Blue-collar workers						
Log life expectancy	8.10	3.45	3.17*	4.07*	3.62**	0.35	4.15**
(of specific age group)	(6.36)	(2.85)	(1.64)	(2.17)	(1.67)	(2.60)	(1.99)
First-stage F–statistic	4.5	9.3	24.4	5.7	13.0	11.4	16.8
States	48	48	48	48	48	48	48
Observations	288	288	288	288	288	288	288
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Average wages contain observations from white-collar workers in Panel (a) and from blue-collar workers in Panel (b). Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

educated young white-collar workers.[32] Conversely, wages of older white-collar workers did not increase significantly following the cardiovascular revolution. Only for the selected group above age 65, life expectancy and average wages are again positively and significantly associated. This result is consistent with reduced sorting-out of productive workers due to improved health.

The coefficient estimates for the blue-collar workers in Panel (b) show a mirror image of the results for white-collar workers. Adult life expectancy and average wages show no statistically significant correlation for the young age groups and old-age workers above 65. Health gains, however, caused a significant wage rise for more experienced workers in the age range from 35 to 64. Hence, this finding supports the hypothesis that especially blue-collar workers were to benefit from improved health conditions due to the demanding physical activities they execute.

Finally, educational attainment represents another potential dimension of heterogeneity. Typically, wages of college-educated workers grow faster with every additional year of work experience compared to workers without any college education (see, for example, Ashenfelter and Rouse, 1999). If, for example, health gains at higher ages prolong the

[32]The effect is slightly more pronounced, if adult life expectancy is lagged. Results are available upon request.

Table 4.8: Heterogeneity: College and Non-College Workers

	Differences-in-differences model (2SLS)						
	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Workers with some college education						
Log life expectancy	-0.34	3.08	2.27**	1.50	0.41	13.67**	2.22**
(of specific age group)	(4.41)	(2.01)	(1.13)	(1.03)	(1.40)	(5.68)	(1.10)
First-stage F–statistic	8.5	20.9	45.5	16.6	26.5	16.9	29.1
	(b) Workers without college education						
Log life expectancy	2.14	1.69	0.91	1.53*	0.82	2.66	1.35
(of specific age group)	(3.07)	(1.35)	(0.85)	(0.88)	(1.06)	(1.76)	(0.94)
First-stage F–statistic	11.1	22.8	47.1	13.8	26.1	16.1	31.7
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Average wages contain observations from workers with at least some college education in Panel (a) and from workers without any college education in Panel (b). Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

work life, wage gains would, ceteris paribus, be larger for college workers as a result of their relatively more favorable earnings trajectory. At the same time, however, wages might show little growth for young college-educated workers, because individuals must invest additional time on acquiring college education. Lastly, gains in adult life expectancy likely raise the share of college enrollment in the population. If, at the margin, individuals, who would have otherwise not chosen this option, select into college because of higher prospective health, average wages for college workers remain unchanged or even decline. Hence, the extent of heterogeneity regarding educational attainment is a priori unclear.

In order to test whether adult life expectancy affects average wages differently along educational attainment, the sample is split into workers with at least some college education and those with at most a high-school degree. Individuals who did not report on their educational attainment are excluded from the sample. Table 4.8 reports parameter estimates for college workers in Panel (a) and non-college workers in Panel (b).[33] Overall, the results reveal no statistically significant association between adult life expectancy and average wages within educational groups. Therefore, health improvements do not alter the within-educational-group earnings trajectories. This finding does, however, not preclude wage increases as a consequence of structural shifts toward a more highly educated

[33]Table C.10 in the Appendix presents estimates for the flexible model.

workforce. The next section highlights this point more closely in the context of potential mechanisms that explain the wage hikes observed in the baseline sample.

4.4.7 Channels

What are the channels through which innovations in understanding and treatment of cardiovascular diseases affect average wages? This section discusses four potential channels that may explain how the treatment translates into higher wages: labor supply, educational attainment, behavioral changes, and improved health.

First of all, consider the possibility that wage hikes may result from changes in labor supply. At the intensive margin, individuals, who know about their improved health prospects, might decide to work more hours per week or more weeks per year and thus earn higher wages. Alternatively, at the extensive margin, workers might feel healthier particularly at higher ages and thus decide to remain in the workforce. Because workers typically earn higher wages with increasing experience and age, increased labor force participation at advanced ages might keep productive workers in the workforce and thus push average wages up. Table 4.9 reports the estimated semi-elasticities for the labor force participation rate (measured as 0 to 100 percent) in Panel (a), usual hours worked per week in Panel (b), usual weeks worked per year in Panel (c), and usual hours worked per year (derived from individual hours and weeks) in Panel (d).[34]

Strikingly, labor supply increased neither at the extensive nor at the intensive margin. In fact, the estimates in Panel (a) show that labor force participation rates decreased by roughly 1 to 1.5 percentage points for most age groups and for the total workforce. Moreover, the decrease was strongest among the 25- to 34-year-olds with no significant effect for the 45- to 54-year-olds. Hence, higher wages for prime-age workers cannot be the result of increased labor force participation. Furthermore, the evidence also precludes the possibility that cohort wages increased unilaterally as a consequence of lower labor supply. If wages and labor force participation were negatively correlated, one should observe an increase of wages for the 55- to 64- but not the 45- to 54-year-olds; however, this is not the case as shown in Table 4.4. The specifications in Panels (b) to (d) reveal moderate negative effects of higher life expectancy on usual hours or weeks worked. For the entire population, a one-percent increase in life expectancy leads to a decline of usual working hours per week by 0.16, or 10 minutes per week. Likewise, usual working weeks shrunk by 0.17, or roughly one workday per year. Neither effect is statistically significant at the conventional significance levels. Again, there is no clear pattern, which explains wage hikes as a consequence of reduced labor supply. Therefore, adjustments along the intensive margin of workers' labor provision cannot explain higher wages either. Finally, the combination of raised wages and stable or reduced labor supply implies an increase of

[34] Tables C.11 and C.12 in the Appendix present estimates for the flexible model.

Table 4.9: Adult Life Expectancy and Labor Supply by Age Cohorts

	Differences-in-differences model (2SLS)						
	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Labor force participation (0 to 100 percent)						
Log life expectancy	-86.15	-223.01***	-92.35***	-21.20	-116.37**	-140.11**	-149.56***
(of specific age group)	(95.18)	(61.93)	(30.32)	(30.71)	(48.82)	(59.82)	(36.88)
	(b) Usual hours per week						
Log life expectancy	-129.42**	-64.50***	-28.72*	-1.77	-35.32*	-38.76*	-15.95
(of specific age group)	(54.81)	(24.25)	(15.24)	(13.08)	(19.86)	(20.92)	(14.57)
	(c) Usual weeks per year						
Log life expectancy	-91.90*	-86.82***	-33.48**	0.08	-43.91*	-55.42**	-17.21
(of specific age group)	(54.38)	(31.75)	(16.58)	(14.45)	(22.95)	(25.55)	(15.58)
	(d) Labor supply of those working (weeks × hours)						
Log life expectancy	-3138.79	-2584.40*	-1147.17	350.41	-1780.77*	-1741.36*	-387.51
(of specific age group)	(2183.64)	(1322.99)	(801.69)	(655.21)	(986.28)	(896.19)	(736.01)
First-stage F–statistic	10.8	22.5	46.9	14.2	26.0	16.0	31.7
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the labor force participation in Panel (a), usual hours worked per week in Panel (b), usual weeks worked per year in Panel (c), and hours worked per year of those working in Panel (d). All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

productivity, as measured by wages per workers, per working hours, or per working weeks. In sum, labor supply cannot account for the observed increase in average wages.

Education provides an alternative channel through which gains in adult life expectancy might affect average wages per worker. Between 1960 and 2000, the share of U.S. whites who enrolled into college at least once roughly doubled from 15 to 30 percent. At the same time, the share of graduates with at least four years of college education tripled from 5 to 15 percent. Based on a prototype Ben-Porath model of human capital and life-cycle earnings, Cervellati and Sunde (2013) show that an increase in survival rates during working ages may raise the benefits of education relative to its costs. Health gains that take place at sufficiently young ages may consequently increase individuals' educational investment. Hansen and Strulik (2017) find that college enrollment increases by roughly ten percentage points as a consequence of higher life expectancy following the cardiovascular revolution in U.S. states. Reproducing their specification with college enrollment of the 15- to 24-year-olds as dependent variable in the empirical framework presented in this paper yields a quantitatively similar estimate of roughly nine percentage points.[35] Therefore, educational attainment is one possible channel through which adult life expectancy affects average wages. Human capital is, however, tied to the person who acquired it. Consequently, upskilling of older worker groups takes time. For example, if individuals around age 20 enroll in college due to the treatment in 1970, the direct benefits of education for wages of 45- to 54-year-olds will only take full effect after 20 to 30 years, when these individuals enter this age group. Hence, education may play a key role in explaining wage hikes, though, only after sufficient time has elapsed.[36] This timing structure will help in confining direct health from gains through educational and behavioral changes.

In combination, the results for education and labor supply indicate that increased life-time labor supply is not a necessary condition for higher educational attainment, as was claimed by Hazan (2009). In contrast, the evidence confirms simulation results of Cervellati and Sunde (2013) and Strulik and Werner (2016). They argue that higher educational attainment and lower life-time labor supply are compatible, if the income effect of higher life expectancy is large enough to afford both, increased life-time consumption and leisure time while at work.

Changes in individual behavior constitute another potential channel through which health may affect wages. For example, preventive measures against smoking following the report of the Surgeon General in 1964 (U.S. Department of Health, Education, and Welfare, 1964) have reduced smoking among U.S. adults considerably after 1970, as exemplified by Figure 4.7. Due to the cumulative damage of smoking, however, cessation requires up to ten years to take the full beneficial health effects (Oza et al., 2011). Consequently, health improvements from reduced smoking should show the full positive effect on wages per worker only starting from the 1980s. Behavioral changes that have more immediate positive effects are increased physical activity, reduced alcohol consumption, and a more healthy diet. Physical activity, however, has gradually declined between 1970 and 2010, while the share of obese persons doubled (Flegal et al., 1998; CDC, 2001; Kohl

[35]Table C.13 presents the effect of a health shock on college enrollment for different measures of adult life expectancy. Following Hansen and Strulik (2017), the estimation equation is weighted by the population at risk; that is, the initial population of 15- to 24-year-olds. To arrive at the numbers, compute $\Delta_y = \frac{\widehat{\alpha}}{100} \cdot \Delta_x = \frac{0.77}{100} \cdot 12.20 \approx 0.09$ with $\Delta_x = \frac{3.16}{25.91} \cdot 100 \approx 12.20$ and $\widehat{\pi} \cdot \Delta_z \cdot \bar{\mu}_x = (-0.61) \cdot (-0.20) \cdot 25.91 \approx 3.16$ computed from the first-stage estimates in Table 4.1, where $\bar{\mu}_x$ is evaluated at the mean in 1960.

[36]Training constitutes another dimension of educational attainment, which might raise worker productivity and wages more immediately, because it is mostly directed at prime-age workers (Carnevale, Strohl, and Gulish, 2015). Public expenditures on training, however, are quantitatively small compared to spending on tertiary education and cover less than one-tenth of a percent of U.S. GDP in the year 2000 (OECD, 2018a; OECD, 2018b). Private expenditures on formal training appear quantitatively more sizable with two to three per mill of U.S. GDP between 2010 and 2015, though, they are again minor in comparison to spending on tertiary education (Training Magazine, 2015; OECD, 2018a).

and Cook, 2013). At the same time, the consumption of alcohol started to decline only after 1980, as shown in Figure 4.8. Hence, behavioral changes due to a more healthy lifestyle do not cause immediate improvements in health conditions among U.S. adults. Nevertheless, there should be positive long-run effects from reduced drinking and smoking.

Taken together, the presented evidence suggests that labor supply cannot explain the observed wage hikes. Meanwhile, higher college enrollment and more healthy behavior only unfold a positive effect on productivity starting in the 1980s. By eliminating these channels, short-run effects of the treatment thus likely reflect health improvements. For example, gains in individual health status arise through new drugs and treatment possibilities such as the beta blocker or the cardiac pacemaker, which allow patients to continue to work only with minor restrictions.

In order to understand how the cardiovascular revolution affected wages over time consider the following fully-flexible model

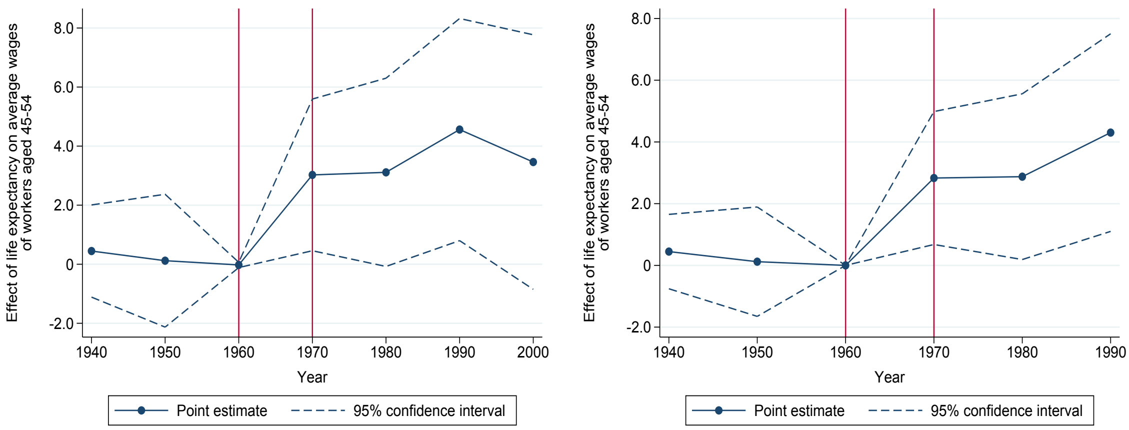
$$y_{s,t} = \sum_{\tau=1940}^{2000} \alpha_\tau x_{s,t} d_t^\tau + w_s' \mathcal{I}_t^\tau \beta + \gamma_s + \delta_t + \zeta_{r,t} + \varepsilon_{s,t}, \tag{4.6}$$

where log life expectancy is interacted with a full set of time dummies. This model allows to estimate the effect of life expectancy on wages for all six time periods relative to

Table 4.10: Effect of Treatment over Time

	45–54 Year Olds		Total Workforce	
	1940–2000	1940–1990	1940–2000	1940–1990
	(1)	(2)	(3)	(4)
Log life expectancy × 1940	0.45	0.45	0.21	0.21
	(0.80)	(0.61)	(1.81)	(1.41)
Log life expectancy × 1950	0.12	0.12	-0.20	-0.20
	(1.15)	(0.90)	(1.44)	(1.10)
Log life expectancy × 1970	3.03**	2.83***	5.26***	5.20**
	(1.31)	(1.10)	(1.97)	(2.22)
Log life expectancy × 1980	3.11*	2.87**	5.84**	5.76**
	(1.63)	(1.37)	(2.50)	(2.88)
Log life expectancy × 1990	4.56**	4.30***	7.73**	7.65**
	(1.92)	(1.63)	(3.32)	(3.68)
Log life expectancy × 2000	3.46	—	6.15**	—
	(2.20)	—	(2.39)	—
First-stage F–statistic	0.6	7.6	2.2	2.0
States	48	48	48	48
Observations	336	288	336	288
Full controls	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of 45- to 54-year-olds. Log life expectancy of the respective age group is interacted with time dummies. Initial log life expectancy is measured in 1960 and interacted with the post-1960 treatment dummy. All other control variables are measured in 1960 and interacted with time dummies. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.



(a) 1940–2000 (Table 4.10, Column 1) (b) 1940–1990 (Table 4.10, Column 2)

Figure 4.11: Illustration: Effect of Life Expectancy over Time (45- to 54-Year-Olds)

the reference year 1960. Using the flexible first stage from equation (4.3), the model is just-identified. The estimated coefficients for the 45- to 54-year-olds are reported in Table 4.10 and plotted in Figure 4.11 for the time periods 1940–2000 and 1940–1990. Due to the increased number of instruments, the value of the F-statistic falls to a level of below one for the model from 1940–2000. If observations in the year 2000 are dropped, the model is somewhat better identified with an F-statistic of approximately 7.6. The results are qualitatively consistent with moving-window and long-differences models, which add one additional year at a time. Therefore, it seems reasonable to provide a qualitative interpretation of the patterns shown in Figure 4.11.[37]

Panel (a) shows the results for the time period 1940–2000. In the pre-treatment periods, the effect of life expectancy is statistically insignificant and very close to zero. After the treatment in 1970, the estimated effect is positive, significant at the five-percent level, and takes a value of 3.03. The estimated parameter does not change much between 1970 and 1980 but becomes slightly less precise. Between 1980 and 1990, the effect increases by approximately 50 percent to 4.56 before it slightly declines thereafter. The model for the period 1940–1990 in Panel (b), which shows a higher value of the F-statistic, confirms these patterns. According to the channels outlined in this section, the immediate increase in 1970 and 1980 is likely due to health improvements. Finally, the gains from higher educational attainment and behavioral changes materialize in the data from 1990 onward, thus explaining an increase in the effect of life expectancy on wages of the 45- to 54-year-olds.

Longitudinal data from the Health and Retirement Study provide another piece of evidence that health effects contributed to wage increases following the cardiovascular revolution. For approximately 22,000 respondents, the data set contains up to twelve

[37]Parameter estimates for the moving-window and long-differences models are reported in Tables C.14 and C.15 in the Appendix.

Table 4.11: Effect of Individual Health on Wages

	Dependent variable: respondents' log wages					
	(1)	(2)	(3)	(4)	(5)	(6)
Heart Disease	-0.06**	-0.06**	-0.06**	-0.07**	-0.01	-0.06
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.06)
× born before 1910		-0.89***				
		(0.07)				
× born before 1920			-0.22			
			(0.23)			
× born before 1930				0.09		
				(0.10)		
× born before 1940					-0.10**	
					(0.05)	
× born before 1950						-0.00
						(0.06)
Individuals	22214	22214	22214	22214	22214	22214
Born before cutoff year	—	71	1062	5947	36112	63154
Observations with heart disease	10023	10023	10023	10023	10023	10023
Total observations	84041	84041	84041	84041	84041	84041

Notes: All regressions include individual-fixed, state-fixed, wave, and census-region-wave effects as well as a quartic age trend. Heart disease is a binary indicator that takes value one, if respondents reports to ever have heart problems diagnosed, and zero else. Heart disease is interacted with a dummy indicator that takes value one, if the individual has been born before a certain threshold level, for example, 1910, and zero else. Standard errors are clustered at the individual level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

observations from biannual interviews between 1992 and 2014. Table 4.11 shows results for the effect of individual health status on wages. Estimates are obtained from an OLS regression of log wage $y_{i,t}$ from individual i at interview wave t on a binary indicator for health status $h_{i,t}$; its interaction with a dummy, $h_{i,t}b_i^\tau$, which takes a value of one, if individual i has been born before a certain cutoff year τ; a quartic age trend $a_{i,t}^k$, $k \in \{1, 2, 3, 4\}$; and a set of fixed effects. Specifically,

$$y_{i,t} = \vartheta h_{i,t} + \rho h_{i,t} b_i^\tau + \phi_1 a_{i,t} + \phi_2 a_{i,t}^2 + \phi_3 a_{i,t}^3 + \phi_4 a_{i,t}^4 + \iota_i + \chi_t + \psi_{s,t} + \omega_{r,t} + \epsilon_{i,t} \quad (4.7)$$

where ι_i, χ_t, $\psi_{s,t}$, and $\omega_{r,t}$ denote individual-fixed, wave-fixed, state-fixed, and census-region-wave effects; and $\epsilon_{i,t}$ constitutes an idiosyncratic error term. In particular, the individual-fixed effect eliminates time-invariant heterogeneity in pivotal dimensions such as ability, educational attainment, and occupational choice. Moreover, state-fixed effects control for state-specific intercepts that pertain to individuals who migrate to another state. Finally, wave and census-region-wave effects address wage differentials that result from general wage trends over time.

The parameter ϑ describes the direct effect of having ever been diagnosed with a negative health status such as 'having heart problems'; in this case $h_{i,t}$ takes a value of one. The respective parameter estimate corresponds to roughly 0.06 throughout all specifications except Column (5). Taken at face value, workers who have been diagnosed with heart problems thus earn on average six percent lower wages compared to workers without heart problems. The causal effect of adverse health effects, however, is likely even more negative. Because the analysis examines variation in wages at the intensive margin, it cannot capture particularly severe cases of heart problems, which end lethally or in disability. Accordingly, cases in which the negative health effect corresponds to 100 percent are, by construction, omitted from the regression.

Furthermore, the parameter ρ captures heterogeneity with respect to the health effect for different birth cohorts. Individuals who have been born before the year 1910 were already around age 60, when new drugs and treatment procedures for cardiovascular diseases became available. Hence, the cardiovascular revolution came too late for them to affect most of their work life, or to provide significant incentives for further investment in educational attainment. The corresponding estimate shows a large and significant negative effect for this group compared to younger cohorts; however, a word of caution is needed. Due to the small number of only 71 observations before 1910, the resulting coefficient may be plagued by both, small sample properties and unobserved selection of individuals within this age group. Even though the results from Column (2) conform with the hypothesis of higher productivity and wages due to improved adult health, they should be seen as suggestive and not conclusive. Columns (3) and (4) illustrate that once the cutoff year is shifted toward younger cohorts for which new drugs and treatments for cardiovascular diseases were at least partly available, the interaction term becomes insignificant. Finally, the results in Columns (5) and (6) indicate that suffering from heart problems poses no negative effect on individual wages for birth cohorts that are young enough to fully harness beneficial effects of the cardiovascular revolution. Concretely, the direct effect of suffering from heart problems becomes statistically insignificant, once the interaction term splits the sample into individuals, who could not or only partly profit from the treatment, and individuals of the reference group who could fully avail of it. Overall, this evidence suggests that health innovations have marginalized negative effects of cardiovascular diseases on individual productivity and income over time.[38]

4.4.8 Discussion

The preceding sections argued that productivity gains from improved health and higher educational attainment explain the observed wage hikes. Here, I discuss to what extent the evidence is consistent with general equilibrium effects and alternative wage theories,

[38]Similar findings apply, if negative health status is measured by ‘high blood pressure’ as Table C.16 in the Appendix shows. For ‘stroke’ as proxy of negative health status, the results show qualitatively similar though statistically insignificant effects. This finding, however, is not surprising insofar that the sample contains considerably fewer observations for stroke and that selection out of the labor market is particularly strong for this group.

in particular: positive demand effects, agglomeration economics, efficiency wages, and compensating wage differentials.

First of all, newly available drugs and health services reduce the frequency of severe courses of disease and, thus, costs for patients and relatives. Therefore, individuals may reallocate income to commodities or additional, previously unaffordable health services, boosting overall demand and wages. Catlin and Cowan (2015) show that national health expenditures continuously increased over the period 1960–2000, and that annual growth rates of national health expenditures exceeded GDP growth rates in all but three years over this time period. As a result, the share of national health expenditures to GDP increased substantially. Hence, individuals devoted rather more than less resources to health services. This finding thus conflicts with a positive demand shock on commodity markets, but it conforms with a positive demand shock on the health sector.

Agglomeration economies and local multipliers constitute another potential source of prolonged income growth (see, for example, Moretti, 2010, and Kline and Moretti, 2013). Specifically, the demand for local goods and services increases with the equilibrium wage and the number of skilled workers in a city or economically-integrated area. The existence of such multiplier effects is consistent with the evidence presented above. Better health conditions and more training increase the number of skilled workers and raise average productivity and, therefore, wages per worker. Consequently, demand for local goods and services increases, thus providing further support for higher wages. Because skilled workers concentrate in metropolitan areas, agglomeration economies may explain heterogeneity in the size of health effects between rural and urban areas. Agglomeration economies, however, constitute a second-round effect that requires initial improvements in productivity or the distribution of skills within an area. Hence, they fail as the sole determinant of the observed positive effect of adult health on average wages.

The discussion so far implicitly assumed that workers be remunerated according to their marginal product on a competitive labor market. Efficiency wage theories depart from this assumption by allowing wages above the market rate, as this profits the firm (see, for example, Katz 1986, Stiglitz, 1986, and Krueger and Summers, 1988). In the context of cardiovascular diseases, for example, firms might find it profitable to pay healthy workers above their marginal product in order to reduce turn-over costs from replacing ill workers. This argument is, however, contradicted by the finding of a larger beneficial effect of health innovations in states with a high initial prevalence of cardiovascular diseases. Improved health conditions and better treatment possibilities, in particular through new drugs, lower firms' incentives to pay wages above market clearing. Therefore, efficiency wage arguments cannot explain the observed wage increase.

Finally, the theory of compensating wage differentials suggests that jobs with less favorable job characteristics must be remunerated with higher wages as "[t]he whole of the advantages and disadvantages of the different employments of labour and stock must, in the same neighbourhood, be either perfectly equal or continually tending to equality. If in the same neighbourhood, there was any employment evidently either more or less advantageous than the rest, so many people would crowd into it in the one case, and so many would desert it in the other, that its advantages would soon return to the level of other employments" (Smith, 1776, Book 1, Chapter 10). According to this prediction, workers would, ceteris paribus, demand higher wages for jobs and states, which pose more disadvantages due to higher risk of cardiovascular diseases. This reasoning, however, conflicts with a larger beneficial effect of health innovations in states with high initial prevalence of cardiovascular diseases. Following health innovations, the compensating wage differentials should collapse, thus implying lower, not higher wage growth in states with high initial prevalence of cardiovascular diseases, as shown by the baseline results. Hence, compensating wage differentials cannot explain the observed raise in average wages.

4.5 Conclusion

This paper establishes a positive causal link between adult health and average wages per worker by exploiting the sharp decline in mortality from cardiovascular diseases in U.S. states after the 1960s. This drop in mortality, also known as the cardiovascular revolution, provides a well-suited source of quasi-experimental variation for several reasons. First, because cardiovascular diseases become more likely with increasing age, they predominantly affect adult health conditions and, thus, adult life expectancy. Second, the decline in mortality rates was initiated by a number of unexpected, path-breaking medical innovations during the 1960s. Lastly, treatment intensities vary across states due to heterogeneity in the prevalence of cardiovascular diseases due to social, cultural, and environmental reasons. Hence, this variation allows to estimate a differences-in-differences model, where all states are treated but with varying treatment intensities. In order to account for endogeneity, adult life expectancy is instrumented by mortality from cardiovascular diseases prior and post the medical advancements in the 1960s.

The results suggest that the cardiovascular revolution was responsible for an increase of life expectancy at 50 of approximately 3.16 years, or roughly two thirds of the increase between 1960 and 2000. This rise in life expectancy can account for roughly 47 percent of the wage increase observed between 1960 and 2000 for workers aged 45 to 54. In particular, the results reveal that the gains concentrate on the prime-age workers between 25 and 54 as well as old-age workers above 65. Correspondingly, the life-cycle earnings profile for an average worker increases more steeply at younger ages, whereas it flattens out more slowly at higher ages. Overall, this pattern is consistent with a workforce that over time becomes healthier at any given age, and at higher ages in particular.

The paper's main finding of a positive causal link between adult life expectancy and average wages also maintains for empirical models that exploit metropolitan-area variation in wages or account for interstate migratory patterns. Adjustments in labor supply cannot explain the estimated wage increase, because labor force participation rates, working hours, and working weeks either declined or remained unchanged during the treatment period. Moreover, age group estimates preclude the possibility of unilateral indirect wage effects through out-selection or increased bargaining power. Furthermore, the analysis reveals that there exists heterogeneity in the beneficial effects of health improvements on average wages between rural and metropolitan areas as well as different occupational groups. The timing of the wage hikes suggests that potential channels are health improvements, in particular in the short-run, and higher educational attainment and potential adoption of a more healthy individual lifestyle in the long run. Evidence based on micro data further suggests that health innovations have marginalized negative effects of cardiovascular diseases on individual income over time. Overall, the evidence demonstrates that thanks to better adult health, workers earn more, work slightly less, and invest more in human capital.

Chapter 5

Population Aging and Income Inequality: Evidence from the Cardiovascular Revolution in U.S. States

5.1 Introduction

Medical breakthrough and demographic change led to substantial population aging in the United States during the twentieth century. This fundamental transformation of the demographic structure was accompanied by a substantial rise of investment in educational attainment, increased saving, better population health, a temporary preponderance of the working-age relative to the dependent population, and ultimately resulted in a (second) demographic dividend in terms of economic growth.[1] While the beneficial effects of the demographic dividend on growth and fertility have been analyzed in depth, little research focused on the quantitative link between population aging and economic inequality.[2]

In order to investigate the causal link between population aging and income inequality across cohorts, this study exploits variation in the unexpected sharp decline of mortality from cardiovascular diseases in U.S. states starting in the 1960s. This sharp reduction in mortality rates is used as an instrument for the change in adult life expectancy, which proxies for average population health and longevity. The identification strategy exploits

[1]A non-exhaustive list on the positive effects of health improvements and the demographic dividend on economic growth contains Bloom, Canning, and Sevilla (2003, 2004); Bloom, Canning, and Graham (2003); Zhang and Zhang (2005); Mason (2007); Lorentzen, McMillan, and Wacziarg (2008); Jayachandran and Lleras-Muney (2009); De Nardi, French, and Jones (2009); Cervellati and Sunde (2011); Hansen, Jensen, and Lønstrup (2018); and Ager, Hansen, and Jensen (2018).

[2]Exceptions are studies by Deaton and Paxson (1994, 1997, 1998a, 1998b), which descriptively show that income and to a lesser extent health inequality increase with age, and recent work by Guvenen et al. (2016), which descriptively tracks changes in lifetime incomes over cohorts. However, they do not assess the role of population aging across cohorts in a rigorous empirical framework.

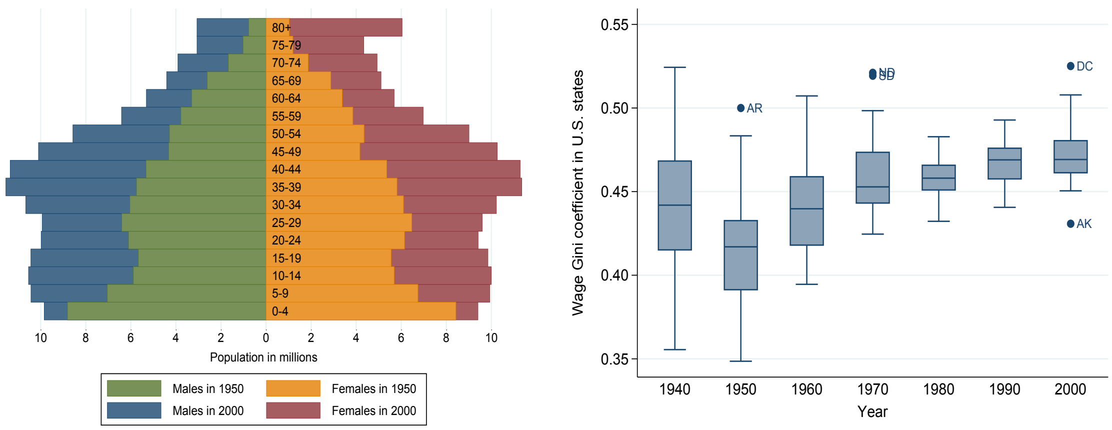
initial geographical differences in the prevalence of cardiovascular diseases across U.S. states in 1960, when medical services and treatment possibilities for patients were scarce. This changed drastically between 1960 and 1970, when a number of seminal innovations for identifying and treating cardiovascular diseases were introduced. New drugs, novel treatment procedures, follow-up inventions, and public education about behavioral risk factors contributed to halve mortality rates from cardiovascular diseases between 1970 and 2000 (CDC, 1999b; National Heart, Lung, and Blood Institute, 2012a). As a consequence, adult life expectancy rose substantially; with the degree of adjustment depending on the initial prevalence of cardiovascular diseases within each state.

This paper contributes to the literature by quantifying the contribution of population aging to income inequality in the United States between 1940 and 2000. From a theoretical viewpoint, this interrelation is mechanical, though the direction and size of its effect are a priori ambiguous. As individuals age, income inequality evolves within age cohorts: Life-cycle earnings profiles suggest that income inequality rises between ages 25 and 64, when the return to educational attainment unfolds, while the income gap contracts again during retirement. Moreover, within-cohort inequalities tend to accumulate over the life-cycle with inequality being most pronounced among the 55- to 64-year-olds (OECD, 2017).[3] Therefore, changes in the demographic structure cause a composition effect that may intensify or depress period income inequality, conditional on the relative size of age cohorts. For example, income inequality may first increase and later fall as large baby boomer cohorts work their way through the demographic structure (see Almås, Havnes and Mogstad, 2014, for the case of Norway). In addition, population aging may contribute to transformations of life-cycle earnings profiles, which become steeper for young ages and flatten out more slowly at higher ages, thus further reinforcing income dispersion over the life-cycle.

Panel (a) of Figure 5.1 illustrates shifts in the demographic structure of the United States between the years 1950 and 2000. Strikingly, the baby boomer generation born between 1946 and 1964 caused a substantial rise in population size which over time has slowly shifted through the demographic structure. In the year 2000, this generation constituted the sizable bulge for the age group of the 35- to 54-year-olds. Panel (b) depicts an increase in pre-tax wage inequality between 1950 and 2000. In combination with the dynamics of life-cycle earnings and within-cohort inequality, the demographic pattern in Panel (a) suggests an aging-driven increase in income inequality following the baby boom until 2020–2025, when the baby boomers start to retire, and a decrease thereafter. Conceptually, a non-linear effect of the demographic structure on income inequality is also consistent with the well-known Kuznets curve.[4] Over the course of the demographic

[3]See, for example, Mincer (1958), Becker (1962), Ben-Porath (1967), and Ashenfelter and Rouse (1999) for further information regarding human capital investment and life-cycle earnings.

[4]In his seminal work, Kuznets (1955) in fact acknowledges the role of demographic change on income inequality; however, he emphasizes differential rates of increase between the rich and the poor rather than dynamics of life-cycle earnings.



(a) Demographic Structure of the U.S. (b) Pre-Tax Income Inequality in U.S. States

Figure 5.1: Demographic Structure and Pre-Tax Income Inequality

Data sources: Panel (a): United Nations, Department of Economic and Social Affairs (2015). Panel (b): IPUMS (Ruggles et al., 2015).

transition, high population growth at first reduces income inequality as the share of young individuals in the working-age population increases relative to the share of older individuals. Once fertility declines, aging of the working-age population begins, thus raising life-cycle income inequality. Finally, aging-related income inequality declines again after fertility rates have stabilized sufficiently.

The empirical results in this paper confirm a positive causal link from population aging to income inequality for a balanced panel of the 48 contiguous U.S. states between 1940 and 2000. In particular, the baseline estimate indicates that, at the margin, a one-percent increase in adult life expectancy, measured at the age of 30, leads to an increase in inequality of 0.9 Gini points, measured on a scale from 0 (perfect equality) to 100 (perfect inequality). The first-stage estimate implies that the decline in mortality rates from cardiovascular diseases between 1960 and 2000 led to an increase in life expectancy at 30 of 3.17 years—that is, a plus of 7.2 percent compared to the initial value in 1960. Taken at face value, higher adult life expectancy thus raised pre-tax income inequality overall by 6.48 Gini points. Hence, population aging contributed considerably to the observed rise of earnings inequality in the United States. Moreover, an age-specific analysis reveals that the effect of improved adult health and longevity on income inequality reaches its maximum when measured for young and middle-aged groups between age 20 and 30. In contrast, the effect becomes small and even vanishes for higher ages. Correspondingly, increased income dispersion results from health improvements and higher prospective longevity during working ages rather than from population aging per se. In particular, this finding is consistent with increasing income dispersion over the life-cycle due to increased investment into educational attainment (Cervellati and Sunde, 2013), life-cycle earnings profiles that flatten out more slowly at higher ages (Chapter 4 and Kotschy, 2018), and wage polarization resulting from skill-biased technical change (Acemoglu and Autor, 2011; Autor and Dorn, 2013).

Furthermore, the empirical findings are robust to different measures of income, such as total income including capital returns or total family income allowing for within-family transfers. Moreover, the instrumentation strategy addresses concerns of reverse causality running from higher income inequality to public health by exploiting mortality reductions before the treatment period.[5] Additionally, the results show only weak signs of convergence in income inequality across states. Population aging therefore affects income inequality irrespective of the initial extent of within-state inequality. Finally, similar qualitative results obtain, if population aging is measured by alternative indicators such as median age or share of 45- to 64-years-olds in the total population. These findings also apply irrespective of initial differences in the demographic structure or inequality between states.

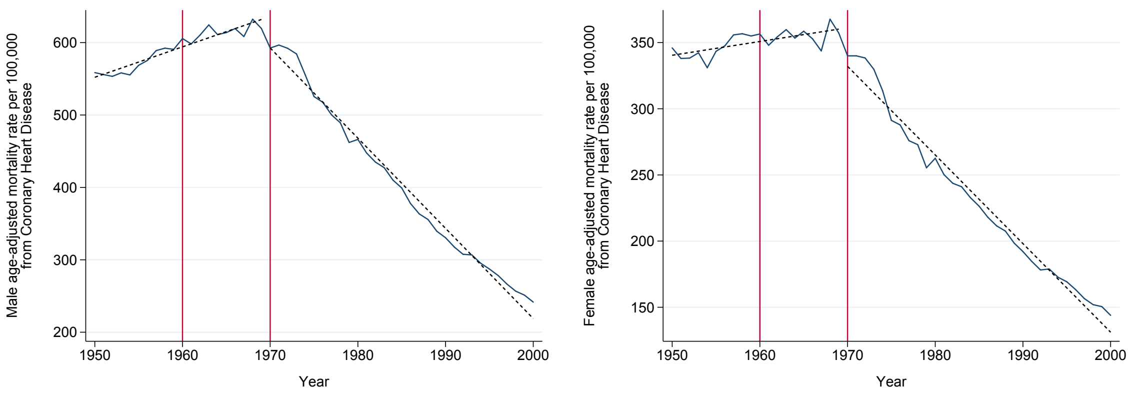
This paper is most closely related to work by Hansen and Strulik (2017) and Kotschy (2018). Both papers exploit variation from the cardiovascular revolution across U.S. states. Hansen and Strulik (2017) investigate the role of gains in longevity for college enrollment, whereas Kotschy (2018) examines the causal link between adult life expectancy and aggregate wages. In contrast, this study highlights the link between adult life expectancy and income inequality. Bhalotra and Venkataramani (2015) use a similar identification strategy to identify positive long-run effects of pneumonia during infancy on adult education, income, and labor supply.

The remainder of this paper is structured as follows. Section 5.2 presents background information on the cardiovascular revolution. Section 5.3 describes the data set and the econometric framework in detail. Section 5.4 presents the empirical results. Finally, Section 5.5 concludes. Because the empirical analysis uses the identification strategy from Kotschy (2018), described in Chapter 4, Sections 5.2 and 5.3 overlap considerably with the corresponding parts of Chapter 4. For clarity and completeness, the respective sections nonetheless repeat the background information and comprehensively describe the data set and the identification strategy.

5.2 Background: The Cardiovascular Revolution

During the second half of the twentieth century, the United States experienced substantial improvements in adult health leading to a marked increase in life expectancy. Specifically, these improvements came down to medical breakthroughs in the understanding, screening and treatment of cardiovascular diseases. This period is thus also referred to as cardiovascular revolution (Foege, 1987; Vallin and Meslé, 2009), and has been classified as fourth stage of the epidemiological transition (Olshansky and Ault, 1986; Omran, 1998). The

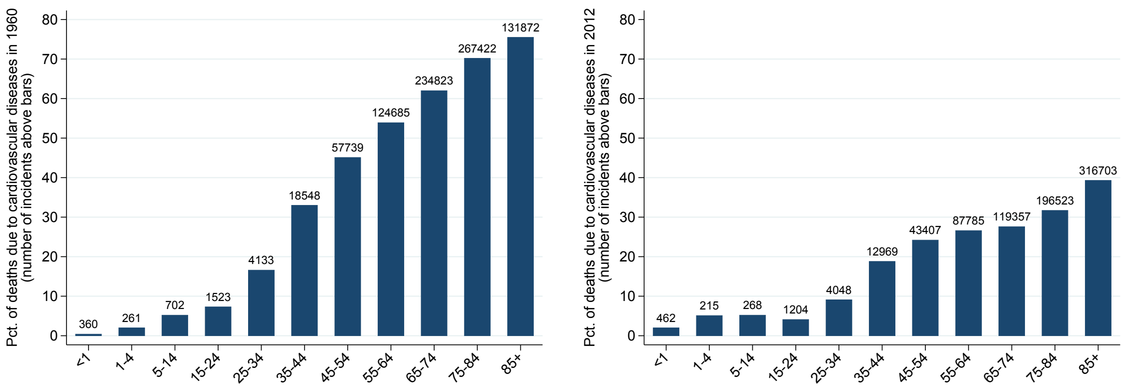
[5] For further information regarding a link from income inequality to population health, see, for example, Lynch et al. (2000) and Wilkinson and Pickett (2006).



(a) Males (b) Females

Figure 5.2: Age-Adjusted Mortality from Cardiovascular Diseases

Data source: National Heart, Lung, and Blood Institute (2012a).



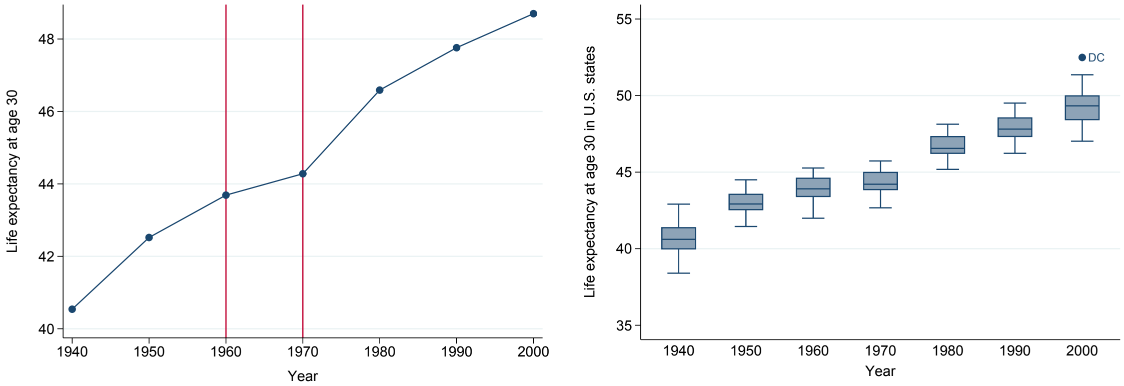
(a) 1960 (b) 2012

Figure 5.3: Percent of Deaths that are Attributable to Cardiovascular Diseases

Data sources: National Center for Health Statistics (1963) and
National Heart, Lung, and Blood Institute (2012b).

cardiovascular revolution considerably contributed to a reduction in mortality from a broad spectrum of cardiovascular illnesses to which a substantial number of deceases accrued. For example, coronary heart disease, the most prominent cardiovascular illness, still accounted for approximately twelve percent of total deaths in the U.S. in 2000 (National Center for Health Statistics, 2017a) and still constituted the most common cause of death in high-income countries in 2004 (WHO, 2008). Figure 5.2 showcases the substantial decline in mortality from coronary heart disease: Between the peak in 1968 and the year 2000, age-adjusted mortality rates per 100,000 shrunk by roughly two thirds for men and women.

In particular, the unexpected invention of new treatment possibilities for cardiovascular diseases boosted life expectancy predominantly through a decrease or delay in advanced-age mortality. Specifically, cardiovascular diseases become more likely as the tissues of the



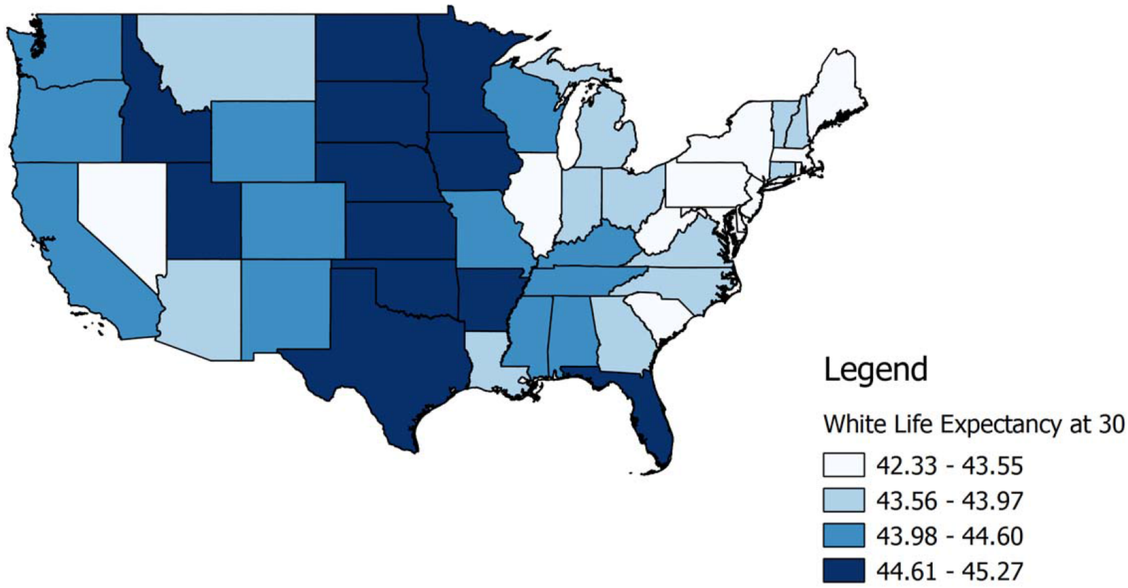
(a) United States (b) States

Figure 5.4: Life Expectancy at Age 30 in the United States

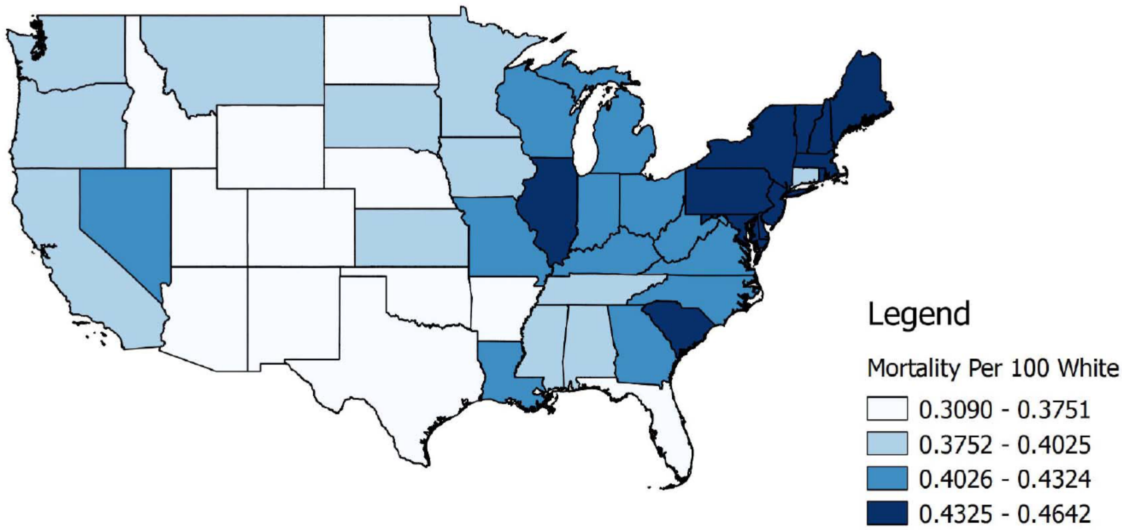
Data sources: United States Life Tables 1940–2000.

cardiovascular system age and lose some of their flexibility (Kirkwood, 2001). Therefore, mortality rates from cardiovascular diseases increase steadily with age, as exemplified by Figure 5.3. Notably, the number of registered deaths dropped considerably for almost all age groups, although the median age of the population increased from 29.5 to 35.3 years during that period (Hobbs and Stoops, 2002). The decline was especially pronounced for individuals in the age range 35–84, thus especially boosting adult life expectancy, measured at age 30, as illustrated Figure 5.4. In contrast, the number of incidents for the group above age 85 more than doubled during that period; however, the overall share of deaths attributable to cardiovascular diseases halved from almost 80 to slightly below 40 percent. A reason was that newly introduced drugs and treatment methods delayed the critical point at which the cardiovascular disease became lethal such that incidents occurred either at a higher age, or death originated from other sources as, for example, cancer. Importantly, there have been striking geographic differences in the prevalence of cardiovascular diseases across U.S. states, which, above all, were rooted in social, cultural, and environmental factors (CDC, 1999b). The initial prevalence of cardiovascular diseases determined how beneficial the treatment was for states. Hence, the decline in mortality and consequently the subsequent increase in adult life expectancy varied across states.

Panel (a) of Figure 5.5 displays the geographic differences in life expectancy at 30 and mortality from cardiovascular diseases in the year 1960. Notably, expected life time of whites at age 30 varied by 2.94 years between Kansas and Nevada—the states with the highest and lowest values. Moreover, adult life expectancy displayed high levels in the West North Central and West South Central census regions, whereas figures are comparatively low in New England and the Atlantic regions. Panel (b) illustrates spatial differences in the prevalence of cardiovascular diseases. In particular, Figure 5.5 reveals a negative strong unconditional state-level correlation between adult life expectancy and the prevalence of



(a) Life Expectancy at 30



(b) Mortality from Cardiovascular Diseases

Figure 5.5: Mortality from Cardiovascular Diseases and Life Expectancy at 30

Data source: Grove and Hetzel (1968).

cardiovascular diseases, as measured by mortality per 100 white people. Correspondingly, states with high mortality rates are more likely to possess low levels of adult life expectancy and vice versa.

Reductions in mortality from cardiovascular diseases arrived through two channels. First, a number of medical innovations between the years 1960 and 1970 allowed to prevent certain diseases or to treat the symptoms. The most remarkable inventions were the artificial cardiac pacemaker, which was first implanted in 1958; the application of chest compression to restore blood circulation in a person that is in cardiac arrest beginning in 1960; the invention of the beta blocker in 1962, which is used to lower blood pressure and to treat cardiac arrhythmia; the invention of the portable defibrillator in 1959 and its application in the U.S. from 1966 onward; and the first adult human heart transplantation in the U.S. in 1968. Subsequent innovations include first thrombolytic therapies in 1986 to treat myocardial infarction, stroke, and pulmonary embolisms; the invention of cholesterol lowering statins, first marketed in 1987; and beginning in 1988, the implantation of intravascular stents to address acute closure of arteries and blood vessels. The new treatment possibilities take effect relatively quickly, as for example, serum cholesterol reducing drugs achieve their full effect within five years (Law, Wald, and Thompson, 1994). These advances in the available technology were complemented by an increasing number of specialists and care centers for cardiovascular diseases (CDC, 1999b).

The second channel for the decline in mortality constitute increased awareness of major individual risk factors and changes in behavior. Research results by Keys et al. (1963), Keys (1980), and Dawber (1980) established, among others, high blood cholesterol, high blood pressure, physical inactivity, smoking, obesity, and unbalanced diet as major risk factors for cardiovascular diseases.[6] The federal government initiated national programs to educate specialists and the general public about risks of high blood pressure in 1972; of high blood cholesterol in 1985; and of the importance of cardiovascular health in 1989 (CDC, 1999a). This increase in awareness helped raise the share of patients with too high blood pressure who have their condition treated and under control (CDC, 1999b). Moreover, the report of the Surgeon General in 1964 (U.S. Department of Health, Education, and Welfare, 1964) highlighted the adverse effects of smoking on health, later followed by increased cigarette taxes in the 1980s (CDC, 1999a). As a result, the share of smokers in the adult population was declining from 1960 onward, while the per capita cigarette consumption started to rapidly fall during the 1970s (U.S. Department of Health and Human Services, 1998, 2000). Alcohol consumption, another risk factor if enjoyed in excess (Marmot and Brunner, 1991; Murray et al., 2002), started to decline after 1980 (Haughwout and Slater, 2017). However, public health education also had its limits. Even though the health risks were known, physical activity gradually declined between 1970 and 2010, while obesity doubled (Flegal et al., 1998; CDC, 2001; Kohl and Cook, 2013).

The unexpected and concentrated surge of medical breakthroughs in the 1960s and the massive preventive efforts thereafter motivate a pre-treatment period until 1960 and post-treatment from 1970 onward in the estimation sample at hand. The next section discusses the empirical framework and the data.

[6]According to Ezzati and Riboli (2012), high blood pressure and high blood cholesterol alone account for one half of the global incidence of coronary heart disease. Too high body weight and smoking are responsible for another 20 and 13 percent.

5.3 Data and Empirical Framework

5.3.1 Data

The empirical analysis is based on a balanced ten-year panel of the 48 contiguous states of the U.S. for the period 1940–2000. Correspondingly, the estimation sample comprises 336 observations in total. Alaska and Hawaii are excluded because of missing data for early periods, the District of Columbia is omitted due to its special nature of a metropolitan region. Because life expectancy in 1940 is only available for whites, the entire sample is restricted to the white population.[7]

Data on wages and incomes stem from individual data in decennial U.S. censuses (IPUMS) by Ruggles et al. (2015) and are adjusted for inflation using the CPI-U price index by the Bureau of Labor Statistics. In the baseline specification, income inequality is measured by the state-level wage Gini coefficient constructed from pre-tax data on wages and salaries with values ranging from zero (perfect equality) to unity (perfect inequality).

Data on life expectancy between 1940 and 2000 are gathered from the U.S. decennial life tables and vital statistics provided by the Centers for Disease Control and Prevention (CDC) of the U.S. Department of Health and Human Services. Adult life expectancy enters the estimation equation in logarithms.[8]

In order to capture the exogenous increase in adult life expectancy due to innovations in medical technology, the analysis exploits state differences in cardiovascular mortality prior to their introduction, that is, until 1960. Age-adjusted cardiovascular mortality in 1960 is obtained from Grove and Hetzel (1968) and expressed in deaths per 100 whites.[9]

Data on age-adjusted mortality rates from coronary heart disease over time are obtained from National Heart, Lung, and Blood Institute (2012a). As Figure 5.2 shows, mortality rates from coronary heart disease increased slightly until around 1968 and only declined substantially thereafter.[10] Hence, for the baseline specification innovations in medical technology are coded to occur from 1970 onward. This designates the time intervals 1940–1960 as pre-treatment and 1970–2000 as post-treatment periods ('differences-in-differences model'). In a more flexible specification, mortality from cardiovascular diseases in 1960 is interacted with a full set of year dummies ('flexible model').

In order to ensure that initial mortality from cardiovascular diseases is as good as

[7]Table D.1 in the Appendix reports descriptive statistics.

[8]Results are qualitatively unchanged, if, instead of a log specification, life expectancy enters the estimation framework directly. Results are available upon request.

[9]Age-adjustment allows to compare the mortality rates between states even if they possess different age structures. Due to the adjustment, mortality rates should not be interpreted as crude rates, unless a state exhibits the same age structure as the standard population. For this reason, not the absolute figures of mortality from cardiovascular diseases are of importance but the relative change over time.

[10]Declines in mortality rates and, thus, improvements in adult life expectancy probably slightly lag behind the actual development for the average white person as medical innovations might come too late for the very ill and the very old.

randomly assigned, the empirical analysis includes further controls interacted with the treatment indicator. These control variables comprise initial life expectancy by the CDC; initial income and initial share of college graduates, both obtained from IPUMS by Ruggles et al. (2015); initial population density by Hobbs and Stoops (2002); and initial mortality from non-cardiovascular diseases by Grove and Hetzel (1968). Current values of the corresponding variables are not included in the empirical framework, because they might themselves be affected by treatment and thus constitute bad controls.

Individual data are collapsed to the state level using person weights in order to ensure representativity of the sample. Due to the collapsing process, wages and educational attainment are grouped on the state level. For this reason, I weight all regressions by the group size, that is, the white population of a given state in 1960.[11]

5.3.2 Empirical Framework

In order to examine the causal link between adult life expectancy and income inequality, I estimate the following model:

$$y_{s,t} = \alpha x_{s,t} + w_s' \mathcal{I}_t^{1960} \beta + \gamma_s + \delta_t + \zeta_{r,t} + \varepsilon_{s,t}, \tag{5.1}$$

where $y_{s,t}$ denotes a measure of inequality for state s and time period t; $x_{s,t}$ represents log adult life expectancy; w_s is a vector of controls measured in 1960, interacted with the treatment matrix $\mathcal{I}_t^{1960}$, whose values take unity from 1970 onward, and zero else; γ_s and δ_t denote state-fixed and time effects; $\zeta_{r,t}$ describes region-year-fixed effects, which control for differential development trends across the nine U.S. census-regions r; and $\varepsilon_{s,t}$ constitutes an idiosyncratic error term.

Due to omitted variables and reverse causality, log life expectancy is likely endogenous. In order to uncover the causal link between life expectancy and income inequality, I exploit heterogeneity in the prevalence of cardiovascular diseases across U.S. states as exogenous source of variation for instrumentation. The first-stage equation is given by

$$x_{s,t} = \eta z_s d_t^{1960} + w_s' \mathcal{I}_t^{1960} \theta + \kappa_s + \lambda_t + \mu_{r,t} + \xi_{s,t}, \tag{5.2}$$

where mortality from cardiovascular diseases in 1960, z_s, is interacted with the post-treatment indicator d_t^{1960}; κ_s, λ_t, and $\mu_{r,t}$ denote state-fixed, time, and region-year-fixed effects; and $\xi_{s,t}$ constitutes the error term.

Conceptually, the first-stage equation compares differences in the increase of adult life expectancy to differences in the decline of mortality from cardiovascular diseases between the pre-treatment and post-treatment period across states. For this reason, it corresponds

[11]Since the equation of interest is the effect of adult life expectancy on income inequality, weighting the regression equation by the group size yields estimation results that are closer to the micro data than unweighted averages. See, for example, Angrist and Pischke (2009).

to a differences-in-differences approach, where all states are treated though with different treatment intensities. Moreover, the first stage possesses a natural interpretation in this context: A decline in the mortality from cardiovascular diseases initiates an increase in adult life expectancy, which, in turn, affects the income inequality in the structural model.

For initial mortality from cardiovascular diseases to be a valid instrument, several conditions must be fulfilled. First, initial mortality must be as good as randomly assigned conditional on control variables. To this end, the baseline specification contains controls for initial state levels of income, the share of college graduates, population density; state-fixed, time, and census-region-year-fixed effects; as well as mortality from non-cardiovascular diseases and life expectancy in 1960. Second, the instrument must affect income inequality only through the first stage. This exclusion restriction is fundamentally untestable. It is, however, plausibly fulfilled in the context of this paper, because the instrument is specific to the channel of health and life expectancy on the aggregate level. Moreover, the empirical model accounts for differences in initial health conditions that are due to factors other than cardiovascular diseases. This monotonicity condition is mechanically fulfilled by the construction principle of life tables. Finally, initial cardiovascular mortality must be correlated to adult life expectancy in the first-stage regression.

I also report results for a more flexible model, in which the treatment is interacted with a full set of year dummies instead of the post-1960 treatment indicator. In this case, the first stage corresponds to

$$x_{s,t} = \sum_{\tau=1940}^{2000} \eta_\tau z_s d_t^\tau + w_s' \mathcal{I}_t^\tau \theta + \kappa_s + \lambda_t + \mu_{r,t} + \xi_{s,t}. \tag{5.3}$$

5.4 Results

5.4.1 First-Stage Evidence: Mortality and Adult Life Expectancy

First, I start by presenting evidence on the first-stage correlation between life expectancy at age 30 and mortality from cardiovascular diseases. The analysis is based on a balanced ten-year panel for the 48 contiguous U.S. with 336 observations in total. Table 5.1 reports results for the differences-in-differences specifications in Panel (a) and the flexible specifications in Panel (b). All regressions include state-fixed and time effects.

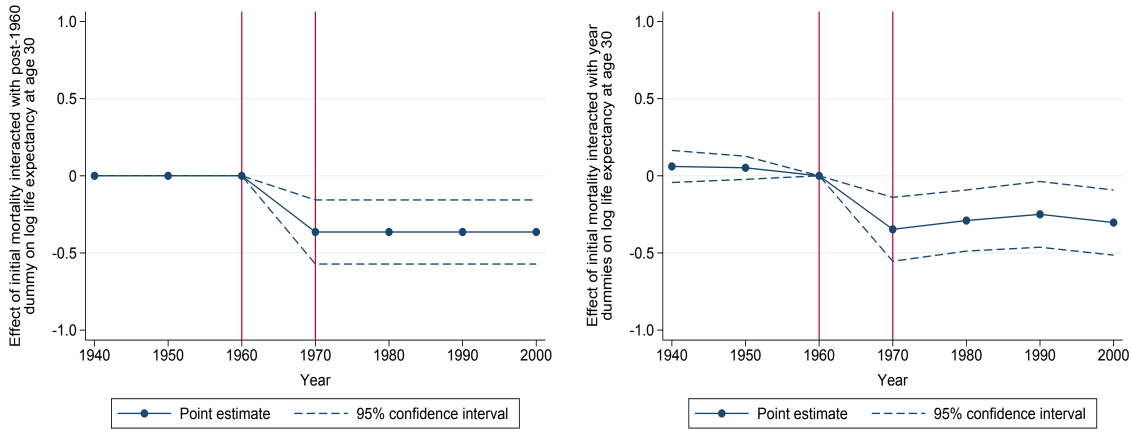
The first column reports parameter estimates for a parsimonious specification without any additional covariates. According to the resulting parameter, life expectancy at age 30 correlates positively with initial mortality from cardiovascular diseases (interacted with the post-1960 treatment indicator). This counter-intuitive result, however, follows directly from the omission of initial life expectancy in the empirical model. By the construction of life tables, age-adjusted mortality rates and life expectancy within a given year must correlate negatively. Moreover, given better initial health conditions, there is less scope for future

Table 5.1: First Stage: Effect of Mortality on Adult Life Expectancy

	Dependent variable: log life expectancy at age 30				
	(1)	(2)	(3)	(4)	(5)
	(a) Differences-in-differences model				
Mortality × Post 1960	0.26***	-0.17*	-0.41***	-0.45***	-0.36***
	(0.07)	(0.09)	(0.11)	(0.08)	(0.10)
	(b) Flexible model				
Mortality × 1940	-0.03	-0.03	-0.06	0.03	0.06
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Mortality × 1950	-0.01	-0.01	-0.02	0.05	0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Mortality × 1970	0.13***	-0.30***	-0.55***	-0.48***	-0.35***
	(0.04)	(0.07)	(0.10)	(0.09)	(0.10)
Mortality × 1980	0.20***	-0.23***	-0.47***	-0.42***	-0.29***
	(0.05)	(0.08)	(0.11)	(0.08)	(0.10)
Mortality × 1990	0.28***	-0.15*	-0.40***	-0.33***	-0.25**
	(0.05)	(0.08)	(0.11)	(0.10)	(0.11)
Mortality × 2000	0.35***	-0.08	-0.33***	-0.46***	-0.30***
	(0.09)	(0.10)	(0.12)	(0.08)	(0.10)
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

reductions in mortality rates and, consequently, also less scope for future improvements in adult life expectancy. Hence, the corresponding estimate is biased upwards. Once the empirical model accommodates for initial life expectancy in the remaining columns, the estimates turn negative, indicating the expected negative correlation between mortality from cardiovascular diseases and life expectancy. The third column adds initial mortality rates from non-cardiovascular diseases to the empirical model. Including this control helps increasing the precision of the estimation and avoiding that the instrument takes up variation from other health improvements that are not attributable to the cardiovascular revolution. As a consequence, the point estimate becomes quantitatively larger, that is, more negative. The fourth column adds region-year effects that account for differential trends in health gains across U.S. census regions. Such differential trends might arise from



(a) DD Model (Table 5.1a, Column 5) (b) Flexible Model (Table 5.1b, Column 5)

Figure 5.6: Illustration: First Stage

economic, social, or cultural factors, for example, disparities in wage trends or migration. Finally, the last column presents results for a full set of covariates comprising the initial share of college graduates, initial population density, and initial average state income. These variables control for variation in life expectancy at 30 that arises from disparities in initial development. The resulting parameter estimate in Panel (a) takes a value of –0.36 and is significant at the one-percent level. Given the quasi-natural source of variation, this parameter possesses a quantitative interpretation: Ceteris paribus, a reduction of mortality from cardiovascular diseases by one person per 1,000 whites leads to an increase in life expectancy at 30 of 3.6 percent. Therefore, the reduction in cardiovascular mortality by approximately two persons per 1,000 whites (50 percent of its initial value) between 1960 and 2000 led to an average increase in life expectancy at 30 by roughly 3.17 years.[12]

Panel (b) reports estimates for the flexible model, in which initial mortality from cardiovascular diseases is interacted with a full set of year dummies instead of the post-1960 treatment indicator; the year 1960 serves as reference category. In particular, differences in the initial prevalence of cardiovascular diseases exert a significant effect on life expectancy at 30 only in the post-treatment period between 1970 and 2000. In contrast, disparities in mortality from cardiovascular diseases cannot explain cross-state differences in life expectancy in the pre-treatment periods from 1940 to 1960. Hence, the evidence suggests common trends in life expectancy before and divergence in these trends after the onset of the cardiovascular revolution.

Figure 5.6 displays parameter estimates for the full model in Column (5) with the corresponding 95-percent confidence intervals. The point estimates are quantitatively similar for the differences-in-differences model in Panel (a) and the flexible model in Panel (b). In particular, the computed coefficients show a stable pattern over the post-treatment

[12] To obtain this figure, compute $\Delta_x = \widehat{\eta} \cdot \Delta_z \cdot \mu_x \approx (-0.36) \cdot (-0.20) \cdot 44.06 \approx 3.17$, where μ_x corresponds to the sample mean in 1960.

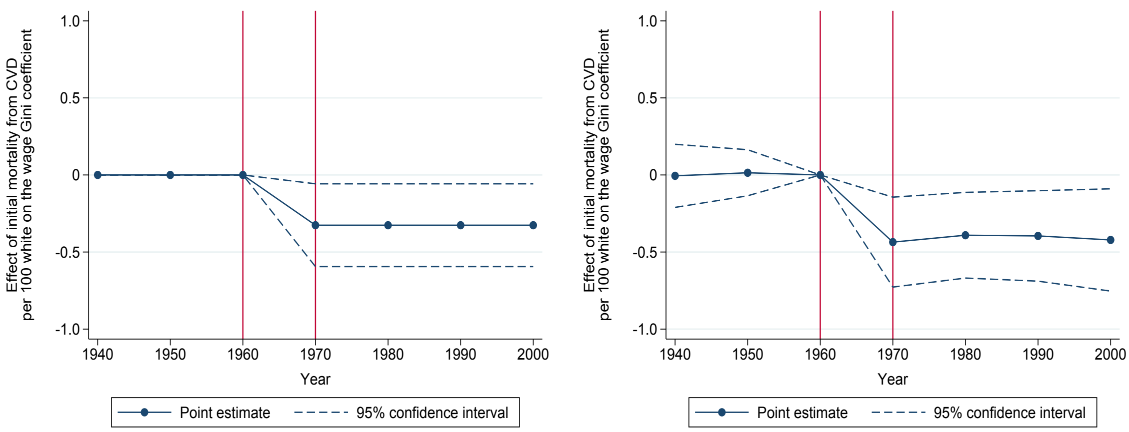
period between 1970 and 2000. Therefore, the differences-in-differences model with a constant effect of initial mortality from cardiovascular diseases on adult life expectancy within the pre- and post-treatment period appears appropriate.

5.4.2 Reduced-Form Evidence: Mortality and Wage Inequality

Table 5.2 presents the reduced-form estimates for a regression of wage inequality, as measured by the wage Gini coefficient, on initial mortality from cardiovascular diseases. All regressions include state-fixed and time effects. As before, the first column reports results for a parsimonious specification without any additional covariates. The subsequent columns add controls for initial levels of life expectancy; mortality from non-cardiovascular diseases; regional-year-fixed effects; and, finally, education, population density, and income.

Panel (a) shows results for the differences-in-differences specification. The parameter estimate for the full model in Column (5) takes a value of –0.33 which is significant at the five-percent level. A reduction in cardiovascular mortality by one person per 1,000 whites leads to an increase in wage inequality by 0.033 unit points of the normalized Gini coefficient, or 3.3 Gini points on a scale from 0 to 100. Given the conditional independence assumption, this figure corresponds to the intention-to-treat effect—that is, the offer of receiving treatment. However, not all individuals comply to the treatment (for example, some individuals do not take a beta blocker, although they belong to high-risk groups for cardiovascular diseases). Hence, the intention-to-treat effect understates the average treatment effect on the treated in absolute terms (Angrist and Pischke, 2009).

Panel (b) presents estimates for the flexible model. According to the full specification in Column (5), initial mortality from cardiovascular diseases affects wage inequality only significantly in the post-treatment period between 1970 and 2000 but not before. Therefore, initial differences in the prevalence of cardiovascular diseases had an impact



(a) DD Model (Table 5.2a, Column 5) (b) Flexible Model (Table 5.2b, Column 5)

Figure 5.7: Illustration: Reduced Form

Table 5.2: Reduced Form: Effect of Mortality on Wage Inequality

	Dependent variable: wage Gini				
	(1)	(2)	(3)	(4)	(5)
	(a) Differences-in-differences model				
Mortality × Post 1960	0.39***	-0.22	-0.61***	-0.81***	-0.33**
	(0.11)	(0.20)	(0.22)	(0.17)	(0.13)
	(b) Flexible model				
Mortality × 1940	-0.09	-0.09	-0.09	-0.00	-0.01
	(0.07)	(0.07)	(0.07)	(0.08)	(0.10)
Mortality × 1950	-0.15***	-0.15***	-0.14***	0.01	0.01
	(0.04)	(0.04)	(0.04)	(0.06)	(0.07)
Mortality × 1970	0.14**	-0.47**	-0.88***	-0.86***	-0.44***
	(0.06)	(0.18)	(0.20)	(0.17)	(0.15)
Mortality × 1980	0.31***	-0.30	-0.69***	-0.78***	-0.39***
	(0.08)	(0.19)	(0.21)	(0.17)	(0.14)
Mortality × 1990	0.36***	-0.24	-0.63***	-0.75***	-0.40***
	(0.09)	(0.18)	(0.20)	(0.17)	(0.15)
Mortality × 2000	0.44**	-0.17	-0.56**	-0.85***	-0.42**
	(0.16)	(0.23)	(0.25)	(0.20)	(0.16)
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

on the trajectories of state-level life expectancy only after the onset of the cardiovascular revolution. Figure 5.7 graphically illustrates this result in Panel (b). The estimates for the flexible model again reveal a stable pattern for the causal link between adult life expectancy and wage inequality within the pre- and post-treatment period, thus providing an argument for the simpler differences-in-differences model depicted in Panel (a).

5.4.3 Adult Life Expectancy and Wage Inequality

This section presents results for the instrumental variables model with initial mortality from cardiovascular diseases (interacted with the treatment indicator) as instrument for

Table 5.3: Adult Life Expectancy and Wage Inequality

	Dependent variable: wage Gini				
	(1)	(2)	(3)	(4)	(5)
	(a) Differences-in-differences model (2SLS)				
Log life expectancy at 30	1.53***	1.24**	1.49***	1.81***	0.90***
	(0.19)	(0.58)	(0.31)	(0.29)	(0.32)
First-stage F–statistic	43.9	8.4	29.9	50.9	23.3
	(b) Flexible model (2SLS)				
Log life expectancy at 30	1.50***	1.31***	1.44***	1.66***	1.10***
	(0.18)	(0.28)	(0.21)	(0.25)	(0.32)
First-stage F–statistic	11.3	6.1	9.1	12.4	5.4
Hansen test p–value	0.3	0.3	0.3	0.7	0.9
	(c) Flexible model (LIML)				
Log life expectancy at 30	1.53***	1.40***	1.50***	1.71***	1.14***
	(0.19)	(0.33)	(0.23)	(0.26)	(0.34)
First-stage F–statistic	11.3	6.1	9.1	12.4	5.4
Hansen test p–value	0.3	0.3	0.3	0.7	0.9
States	48	48	48	48	48
Observations	336	336	336	336	336
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed and time effects. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

life expectancy at age 30.[13] The respective estimate corresponds to the average treatment effect on the treated. This quantity equals the ratio of the intention-to-treat effect obtained from the reduced-form model to the compliance rate, which corresponds to the first-stage estimate. Table 5.3 reports results for the effect of life expectancy at 30 on the wage Gini coefficient. Panel (a) contains parameter estimates for the differences-in-differences model, whereas Panels (b) and (c) present results for the flexible model. Estimates are obtained from two-stage least squares (2SLS) and additionally from limited information maximum likelihood (LIML) for the overidentified flexible model.

[13]Table D.2 in the Appendix reports estimates from ordinary least squares (OLS).

The first column shows results for a parsimonious specification without any additional covariates. As Aghion, Howitt, and Murtin (2011) and Bloom, Canning, and Fink (2014), however, argue in the context of economic growth, this model is misspecified, because it does not control for initial life expectancy in order to capture convergence in health conditions over time. Specifically, initial life expectancy correlates with both, mortality rates of cardiovascular diseases in 1960 and subsequent improvements in life expectancy. At the same time, initial life expectancy affects the subsequent trajectory of wage inequality. Without controlling for initial life expectancy, the first stage thus underestimates the beneficial effect of the cardiovascular revolution on subsequent improvements in life expectancy. In addition, the reduced form underestimates the increase in wage inequality for a given reduction in mortality from cardiovascular diseases. The results in Tables 5.1 and 5.2 demonstrate how serious these biases are: Once initial life expectancy enters the empirical framework as additional control, the parameter estimates change their sign. In combination, these biases lead to a too large estimate for the effect of adult life expectancy on wage inequality. Hence, all remaining specifications control for initial life expectancy. In the third column, the empirical model includes initial mortality from non-cardiovascular diseases. This control precludes the instrument from taking up variation from health improvements that cannot be attributed to the cardiovascular revolution. Moreover, the inclusion of these variables improves the goodness of fit of the first-stage model as documented by an increase of the Kleibergen-Paap F-statistic from 8.4 to 29.9. The fourth column adds region-year-fixed effects that accommodate for differential trends in life expectancy and wage inequality across census regions. Finally, the full specification in the fifth column controls for initial cross-state disparities in educational attainment, population density, and average income. These additional explanatory variables account for the link from education to health (Grossman and Kaestner, 1997; Lleras-Muney, 2005), the rural-urban gradient in risk factors for cardiovascular diseases (Pickle and Gillum, 1999; Cooper et al., 2000), and potential feedback effects from income to health (Ettner, 1996; Frijters, Haisken-DeNew, and Shields, 2005; Lindahl, 2005; Chetty et al., 2016).

The parameter estimate for the full specification in Column (5) takes a value of 0.90. The estimated coefficient is significantly different from zero at the one-percent level. The corresponding F-statistic shows a value of 23.3 indicating a sufficiently strong first-stage correlation between initial mortality from cardiovascular diseases (interacted with the treatment indicator) and subsequent improvements in life expectancy at 30. Ceteris paribus, a one-percent gain in life expectancy at 30 thus leads to an increase of wage inequality by approximately 0.009 of the normalized Gini coefficient, or 0.9 full Gini points on a scale between 0 and 100. Taken at face value, this estimate suggests that the gain in life expectancy at 30 from the cardiovascular revolution between 1960 and 2000 led to an increase in wage inequality of 0.0648 unit points of the normalized Gini.[14] Albeit

[14]To obtain these figures, compute $\Delta_y = \widehat{\alpha} \cdot \Delta_x / x^{1960} \cdot \bar{\mu}_y \approx 0.90 \cdot 3.17/44.06 \approx 0.0648$, where μ_y

this figure is quantitatively large, it is in line with the overall state-level development of wage inequality as the raw data in Panel (b) of Figure 5.1 illustrate. In particular, states that started from low levels of wage inequality simultaneously experienced substantial gains in adult life expectancy and wage dispersion over time. For example, the states of Massachusetts, New Jersey, and New York improved their life expectancy at 30 by roughly six years between 1960 and 2000 according to the raw data. At the same time, wage inequality in these states increased between by up to eleven Gini points on a scale from 0 to 100. In contrast, states that experienced smaller improvements in life expectancy at 30 of roughly three to four years such as Arkansas, Nebraska, or Oregon also show considerably smaller increases in wage inequality (and even a decline in the case of Arkansas).

Panels (b) and (c) report results for the flexible model. All specifications are estimated with two-stage least squares and a heteroskedasticity-robust version of limited information maximum likelihood in order to account for the low first-stage correlation of the endogenous variable and the instruments, as expressed by small values of the F-statistic. Column (5) reports parameter estimates that are quantitatively somewhat larger compared to the differences-in-differences model though also slightly less precisely estimated. The Hansen test of overidentification shows high p-values such that the null hypothesis that all instruments provide the same information cannot be rejected.

5.4.4 Robustness of Results

This section briefly discusses the robustness of the results with respect to alternative measures of income equality and population aging as well as controlling for convergence in income inequality across states.

Different Measures of Income Inequality. In the baseline specification, income inequality across states was proxied by the wage Gini coefficient computed from individual census data. However, this measure includes only employees' wages and salaries, thus neglecting incomes of the self-employed as well as income derived from other sources such as savings, businesses, social security, or transfers from family members. In order to account for these additional income dimensions, I additionally construct two income Gini coefficients based on total personal income and total family income. Table D.3 in the Appendix confirms the positive link between adult life expectancy and income inequality for the alternative measures. Quantitatively, the corresponding estimates take slightly larger values, indicating the important role of health improvements and aging on income inequality during the period 1940 to 2000.

corresponds to the sample mean in 1960. Quantitatively similar results (not reported) obtain, if wage inequality is measured in logarithms. Results are available upon request.

Different Measures of Population Aging. The baseline model revealed a positive effect of population aging, as measured by adult life expectancy, on income inequality. Table D.4 in the Appendix reports results for specification that instead proxy population aging by the changes in median age or the share of 45- to 64-year-olds in the workforce. The findings again confirm the positive association between population aging and income inequality. Quantitatively, the estimated parameters for median age conform closely to the baseline results, whereas estimates for model with the workforce share of the 45- to 64-year-olds predict a somewhat more moderate increase in income inequality over time.

Convergence in Income Inequality. Adjustments in income inequality over time might depend on the initial extent of income inequality within each state. Correspondingly, the specifications in Tables D.3 and D.4 include initial inequality to control for convergence dynamics in the adjustment process. In particular, the results show only weak signs of convergence in income inequality across states. Hence, the evidence suggests a causal link from population aging to income inequality irrespective of the initial within-state dispersion of incomes and regardless of the employed inequality and aging measures.

5.4.5 Heterogeneity Along the Age Dimension: Does It Matter Who Profits from Health Improvements?

The empirical results in Section 5.4.3 show a positive causal link from population aging proxied by life expectancy at age 30 to wage inequality for the observation period from 1940 to 2000. However, individuals earn most of their work incomes during working age between 16 and 64. Hence, it might matter for individual life-cycle earnings at which age health improvements take place. Health gains at young ages should economically benefit individuals the most, whereas health improvements shortly before retirement likely exert a minuscule effect of individual income. In order to test this hypothesis, I vary the age at which life expectancy is measured. Table 5.4 presents results for the full specification with controls for initial levels of life expectancy, mortality from non-cardiovascular diseases, education, population density, and average income. All regressions include state-fixed, time, and region-year-fixed effects.

Panel (a) reports estimates for the differences-in-differences model. Values (equal to or) above the conventional threshold of 10 for the F-statistic suggest a sufficiently strong first-stage correlation between life expectancy and initial mortality from cardiovascular diseases (interacted with the treatment indicator) for all specifications. Column (1) reveals a significant positive effect of life expectancy at birth on wage inequality. The computed parameter takes a value of 0.64 corresponding to roughly 70 percent of the baseline estimate of 0.90. The effect of life expectancy on wage inequality reaches maximum values, if measured around ages 20 and 30 as documented in Columns (2) and (3). For health

Table 5.4: Adult Life Expectancy Measured at Different Ages

	Dependent variable: wage Gini					
	Life expectancy measured at					
	birth	20	30	40	50	60
	(1)	(2)	(3)	(4)	(5)	(6)
	(a) Differences-in-differences model (2SLS)					
Log life expectancy	0.64**	0.97*	0.90***	0.45**	0.26	0.24
	(0.32)	(0.51)	(0.32)	(0.20)	(0.20)	(0.19)
First-stage F–statistic	11.0	10.0	23.3	49.4	14.3	30.1
	(b) Flexible model (2SLS)					
Log life expectancy	0.69**	1.02**	1.10***	0.62***	0.34*	0.26
	(0.32)	(0.46)	(0.32)	(0.20)	(0.18)	(0.16)
First-stage F–statistic	3.4	3.2	5.4	8.7	3.6	6.4
Hansen test p–value	0.9	0.9	0.9	1.0	1.0	1.0
	(c) Flexible model (LIML)					
Log life expectancy	0.72**	1.07**	1.14***	0.62***	0.34*	0.26
	(0.35)	(0.49)	(0.34)	(0.20)	(0.18)	(0.16)
First-stage F–statistic	3.4	3.2	5.4	8.7	3.6	6.4
Hansen test p–value	0.9	0.9	0.9	1.0	1.0	1.0
States	48	48	48	48	48	48
Observations	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

improvements that mostly benefit individuals at age 40 or above, the estimated effect life expectancy on income inequality roughly halves to 0.45 as Column (4) shows. Finally, Columns (5) and (6) reveal that health gains accruing above age 50 exert no significant effect on income inequality. Hence, the evidence indicates that health gains during working ages rather than aging per se constitute the driving forces behind income inequality.

This result is consistent with theoretical predictions and simulation results from Cervellati and Sunde (2013). In particular, they show that mortality reductions must take place at sufficiently early ages to increase educational attainment and individual life-time wealth. In this case, the benefits of education increase relative to its cost such that individuals invest more in educational attainment implying a steeper trajectory for their life-time earnings, particularly at young ages. Consequently, there is more scope for income dispersion over the life-cycle. Evidence by Hansen and Strulik (2017) supports this view: College enrollment increased considerably as a result of the cardiovascular revolution. Moreover, empirical results for the same setting by Kotschy (2018)—described Chapter 4—indicate a transformation towards flatter life-cycle earnings profiles at older ages. Therefore, health gains may shift the old-age contraction of life-cycle inequality to higher ages. Finally, the evidence is consistent with wage polarization as a consequence of skill-biased technical change (Acemoglu and Autor, 2011; Autor and Dorn, 2013). Given higher educational investment following health gains from the cardiovascular revolution, a larger amount of high-skilled workers allows a structural shift to highly productive but skill-intensive jobs that cannot be replaced by machines. In contrast, routine jobs vanish as a consequence of automation, and workers reallocate to low-skill low-productivity service jobs. Hence, the polarization between an increasing number of high-skilled individuals with steep life-cycle earnings profiles and a sizable number of low-skill service workers with flat life-cycle earnings profiles exacerbates income inequality across cohorts.

Panels (b) and (c) report results for the flexible model. The corresponding parameters are quantitatively slightly larger than those obtained for the differences-in-differences model. Qualitatively, the estimates again indicate a large effect of life expectancy on income inequality during working ages, whereas the effect becomes small and marginally significant for higher ages. The null hypothesis that all instruments provide the same information is maintained given the p-values of the Hansen test for overidentification.

5.5 Conclusion

This paper documents a positive causal link from life expectancy during early working-ages on pre-tax income inequality for the United States between 1940 and 2000. To this end, the empirical identification strategy exploits quasi-natural variation in the decline of mortality rates from cardiovascular diseases in U.S. states after 1960. This variation allows the estimation of a differences-in-differences model, where all states are treated by health gains following the cardiovascular revolution though with different treatment intensities. In order to account for endogeneity concerns, adult life expectancy is instrumented by initial cross-state disparities in mortality from cardiovascular diseases (interacted with a treatment indicator).

The empirical results indicate that, at the margin, a one-percent increase in life expectancy at age 30 is associated with a hike in wage inequality of 0.9 Gini points on a scale from 0 to 100. Overall, the estimates suggest that pre-tax wage inequality increased by 6.48 Gini points following health gains for working-age individuals during the cardiovascular revolution. Moreover, the analysis reveals that the effect of health improvements and longevity on income inequality declines sharply and even vanishes for higher ages. This finding suggests that increased income dispersion results from health improvements and higher prospective longevity during working ages rather than from population aging per se. In particular, this finding is consistent with reinforced dispersion of incomes over the life-cycle due to higher educational attainment (Cervellati and Sunde, 2013; Hansen and Strulik, 2017), flatter life-cycle earnings profiles at higher ages (Chapter 4), and wage polarization as a consequence of skill-biased technical change (Acemoglu and Autor, 2011; Autor and Dorn, 2013). However, more research is still needed to shed light on the channels that link individual decisions regarding educational attainment and occupational choices to aggregate inequality.

Chapter 6

Income Shocks, Inequality, and Democracy

6.1 Introduction

Since Lipset's famous hypothesis that a sufficiently high level of income is a prerequisite for democracy (Lipset, 1959), the causal effect of income on democracy has been a central theme in the social sciences. While much of the previous literature found evidence consistent with a positive effect of income on the quality of democratic institutions (for example, Barro, 1999), more recent work by Acemoglu et al. (2008, 2009) suggests that the positive association between income and democracy disappears, once relevant systematic differences across countries are accounted for that affect both income and democracy. In particular, their results reveal no significant effect of income on democracy in cross-country panel regressions with country and time fixed effects. This finding has initiated an ongoing debate about the role of income for democracy, with some studies finding evidence for a positive effect based on non-linear estimators or refined methods, others finding substantial heterogeneity in the effect of income, and yet others providing evidence for significant improvements in democratic institutions in response to negative income dynamics or shocks. To date, there is no coherent explanation for this apparently incoherent and contradictory evidence on the effect of income on democracy.

This paper tests the hypothesis that income shocks—rather than minor fluctuations in income—trigger major changes in institutional quality, as reflected by transitions between autocracy and democracy, and that the effect depends crucially on the social environment, as reflected by economic inequality. This hypothesis is rooted in the theoretical literature of democratic transitions under threats of revolutions (see, for example, Acemoglu and Robinson, 2000, 2005) and the alternative of elite-driven transitions to democracy (see, for example, Lizzeri and Persico, 2004).[1] According to this literature, negative economic

[1]See also Cervellati, Fortunato, and Sunde (2014) for a unified theory of different transition scenarios.

shocks might provide an opportunity to overcome autocratic institutions, in particular in an environment with high inequality. Conversely, democracy might emerge for economic reasons in environments of low inequality and, thus, low redistributive conflict.

In light of these predictions, we hypothesize that an appropriate empirical analysis of the income-democracy nexus should focus on economic shocks and non-marginal changes in democratic quality instead of exploiting continuous variation in income and institutional quality. Moreover, the theory suggests that the effects of income shocks crucially differ by the cohesiveness of society, as reflected by economic inequality. In economically highly unequal societies, negative income shocks are likely to trigger revolts and, thereby, open a window of opportunity for democratization, whereas positive income shocks tend to stabilize oligarchic structures. In economically equal societies, by contrast, positive income shocks do not generate much distributive pressure, thus helping to consolidate and improve democratic quality, whereas negative income shocks might erode democracy by creating tensions within the society.

The results provide support for this hypothesis and document the crucial role of inequality for the effects of economic shocks on the quality and stability of political institutions. In particular, the findings show that negative income shocks unfold a negative effect on democracy in countries with low inequality but a positive effect in countries with high inequality.

The paper contributes to the existing literature in several ways. In response to the analysis by Acemoglu et al. (2008, 2009), several studies find a positive effect of income on democracy using non-linear estimators (Heid, Langer, and Larch, 2012; Benhabib, Corvalan, and Spiegel, 2013; Che et al., 2013). At the same time, papers analyzing the effect of exogenous income shocks find both, positive and negative effects on democratic quality. For instance, Papaioannou and Siourounis (2008) document pronounced negative income dynamics before democratization, and Aidt and Franck (2015) show that poverty-related riots led to democratic improvements in 19th Century England. Similarly, Brückner and Ciccone (2011) find that negative income shocks have a positive effect (a window of opportunity) for democratic improvements in Africa. In contrast, Brückner, Ciccone, and Tesei (2012) find that positive income shocks due to increases in oil prices have a positive effect on democratic quality in countries that are net oil exporters. Our findings reconcile earlier results for positive effects of income on democracy with evidence that negative income shocks have a positive effect on democratic improvements: We document an important role of major income fluctuations and a significant asymmetry in interaction with economic inequality. Moreover, this finding also complements evidence from other recent contributions that suggest that income unfolds vastly heterogeneous effects on democratic institutions (for example, Moral-Benito and Bartolucci, 2012, for heterogeneity across low and high income countries, and Cervellati et al., 2014, for heterogeneity with respect to colonial history).

This study most closely relates to work by Dorsch and Maarek (2014a, 2014b) according to which episodes of democratization can be explained by variation in income inequality, whose effect is amplified during economic downturns. In contrast, this paper focuses on the role of economic shocks rather than income inequality. In particular, we show that economic shocks non-monotonically affect democratic institutions conditional on the level of inequality. This result holds irrespective of whether economic shocks are measured by cyclical fluctuations around the long-run trend of income per capita, by the trend itself, or by devaluation of disposable income as a consequence of episodes of high inflation. Moreover, we provide evidence for an asymmetry with respect to the nature of economic shocks: While negative shocks open a window of opportunity for institutional change, no comparable countervailing effect is found for positive shocks. Finally, our empirical analysis considers demographic pressure as another determinant of democratic transitions.

The remainder of this paper is structured as follows. Section 6.2 presents the empirical approach, data sources, and variable construction. Section 6.3 presents the main results and provides a brief discussion of robustness and additional findings. Section 6.4 concludes.

6.2 Empirical Framework and Data

6.2.1 Empirical Framework

The empirical framework required to test the hypothesis of this paper focuses on identifying the effect of major income fluctuations on the quality of political institutions, in isolation as well as in interaction with inequality. To this end, the estimation framework exploits within-country variation over time in a dynamic linear panel model.[2] The estimation equation is given by

$$d_{i,t} = \alpha s_{i,t-k} + \beta x_{i,t-k} + \gamma (s \cdot x)_{i,t-k} + w'_{i,t-k}\delta + \zeta_i + \eta_t + \varepsilon_{i,t}, \tag{6.1}$$

where $d_{i,t}$ denotes democratic quality (or democratization) in country i in year t, measured continuously or by a binary indicator reflecting major changes in the index between $t-k$ and t; $s_{i,t-k}$ indicates whether an economic shock occurred during the time period $t-k$ and $t-(k+l)$; $x_{i,t-k}$ denotes inequality; and $w_{i,t-k}$ is a vector of controls that includes the quality of democratic institutions, the level of GDP per capita as well as education in period $t-k$. In the baseline analysis, we use three-year windows; that is, $k=3$. The specification includes a full set of country and time dummies, ζ_i and η_t, respectively. The coefficients of interest in light of the hypothesis to be tested are α and γ.

Country-specific fixed effects are removed using the within-transformation. The analysis exploits a yearly panel data set ranging from 1950 to 2014 with a maximum number

[2]Similar specifications were used by Acemoglu et al. (2008) and Murtin and Wacziarg (2014).

of 64 time periods. Therefore, the time dimension T is sufficiently large such that the well-known Nickell (1981) bias is of little concern for the identification of the coefficients of interest (Judson and Owen, 1999).

We define economic shocks as cyclical fluctuations in output. Unlike long-run growth trends, these fluctuations are arguably largely unforeseen by individual agents. For this purpose, we use the Hodrick-Prescott (HP) filter to disentangle cyclical fluctuations from long-run trends in economic development. Section 6.2.3 discusses the construction of the shock indicator in more detail.

6.2.2 Data Sources

There is an ongoing debate about the appropriate measurement of institutional quality. Specifically, this debate refers to the information on which the respective indices are based, as well as their measurement on a discrete or continuous scale. While continuous measurements conform more to the slowly-changing nature of institutions described by North (1990), dichotomous measures provide a clearer distinction of the bi-modal distribution of political institutions observed in practice (Cheibub, Gandhi, and Vreeland, 2010). In the absence of a consensus with respect to this question, we report results for both continuous and dichotomous measures from different sources. In particular, we use the composite PolityIV index by Marshall, Jaggers, and Gurr (2013), a composite indicator based on the Freedom House (2014) Political Rights and Civil Liberties measures, and the binary Democracy-Dictatorship index by Cheibub, Gandhi, and Vreeland (2010). For comparability, we normalize all measures of democratic quality to a range from zero (full autocracy) to one (full democracy).[3] Following recent suggestions by Voigt (2013), we additionally construct an artificial indicator based on the principal components of the PolityIV, Freedom House, and Democracy-Dictatorship indicators. This composite index isolates and extracts the common variation among all three measures and combines them in a single indicator that can be interpreted as democratic institutions.[4]

Data for (log) GDP per capita and its growth rate are taken from Penn World Tables by Feenstra, Inklaar, and Timmer (2015). We proxy income inequality with market (that is, pre-tax, pre-transfer) Gini coefficients from the Standardized World Income Inequality

[3]The PolityIV and the composite Freedom House index both constitute broad measures of institutional quality comprising not only features of the political but also the economic domain. Both dimensions are in practice highly correlated but not necessarily identical as Acemoglu and Johnson (2005) and Kotschy and Sunde (2017) point out. We presume that democratic transitions aim to improve institutional quality compared to the previous regime. Hence, we choose rather broad measures of institutional quality in order to capture all possible facets of these transitions.

[4]Factor analysis synthesizes the variation contained in several variables into common, orthogonal factors, or principal components. This way, one can decompose the variation in institutional variables that corresponds more closely to democratic institutions, from variation that corresponds more closely to other institutional dimensions such as those affecting the economic domain. Hence, this artificial index corresponds to a more narrow measure of democratic quality compared to its source indicators. For more information see Voigt (2013, pp. 20–21).

Database by Solt (2009, 2016a). These data provide standardized Gini coefficients that are comparable across countries and time.[5] Despite recent criticism by Jenkins (2015) regarding quality of the data and of the underlying imputation model used for their construction, these inequality data appear as best suited for the purpose of this paper due to their comparability across countries and time.[6] We normalize the observed Gini coefficients to vary between zero and one, where a value of one indicates maximum inequality and zero perfect equality. Moreover, we control for human capital differences using average years of schooling from Barro and Lee (2013).

6.2.3 Economic Shocks and Binary Democracy Indicators

The main hypothesis of this paper is that, in light of the existing evidence and the theoretical motivation, an investigation of the effect of income on democracy should focus on the consequences of major income fluctuations instead of exploiting marginal changes. Likewise, most theoretical models consider dichotomous institutional regimes—democracy and autocracy—suggesting the relevance of using dichotomous measures of the quality of political institutions rather than multi-valued indices or continuous measures. To implement this, we construct measures of economic shocks and binary measures of democratic institutions.

In particular, we code a binary indicator for economic shocks that takes a value of one, if at least once over the past two years an adverse (that is, negative) cyclical income shock occurred, which in absolute terms is larger or equal than five percent of the HP-filtered income per capita trend. Note that negative cyclical shocks of more than five percent are sizable. For example, even during the Great Recession and its aftermath, most Western countries did not experience shocks of more than two to three percent of GDP per capita. In Greece, a country that was hit especially hard by the recession and the following Euro Crisis, the largest shock amounted to a value of -5.33 percent in 2011. As consequence, such events are relatively rare and occur only in less than five percent of the country-year observations of our data. A substantial amount of these shocks occurs in low and middle income countries. Because these shocks pose a sizable strain on incomes and the political discourse within countries for some time, we allow the shock indicator to take a value of one also if the shock occurred in one of the two previous years (that is, $l = 2$). Given a democratic transition begins in period $t-k$, we thus code the shock indicator s_{t-k} to take a value of one, if a cyclical shock occurred in either of the years $t-k$, $t-(k+1)$, or $t-(k+2)$. The so-measured shock indicator overall takes a value of one for roughly twelve percent of the annual observations with economic shocks being concentrated among low

[5] In particular, the SWIID uses imputation procedures to construct a comprehensive set of inequality estimates over time, with the numerous Gini data points varying with respect to their (un-)certainty.

[6] In fact, Solt (2015), refuted the critique on this data set by asserting that "[t]hose pursuing research on income inequality across many countries and over time [...] will often find that the SWIID is their best choice of data source" (Solt, 2015, p. 690).

and middle income countries. Choosing a longer window ($l > 2$) for economic shocks to affect democratic institutions results in a larger number of "shock" observations and, thus, a noisier measure. In contrast, choosing a narrower window ($l < 2$) results in fewer but more concentrated effects of economic shocks on democratic institutions. A more narrow coding is more likely to miss slumps that unfold their full effects only over the span of several years. We present results for alternative windows lengths for the construction of shocks in the robustness analysis.

Following the rule suggested by Ravn and Uhlig (2002), the smoothing parameter for the HP filter is set to $\lambda = 6.25$. The smoothing parameter determines how smooth the trend component of the filtered GDP series is, and, thus, how large the peaks of cyclical fluctuations are. A lower smoothing parameter produces smaller cyclical fluctuations and thus stronger changes in the long-run trend, and vice versa. We consider the sensitivity of our results regarding the choice of this parameter in the robustness section.

In order to investigate whether economic shocks affect not only small but also major changes in democratic institutions, we construct a binary measure of changes in democratic quality (that is, a "democratization" indicator). This indicator takes a value of one, if the *change* in normalized democratic quality exceeds a certain threshold over an interval of $k = 3$ years. In order to account for disparities in variation among the different indicators of democratic quality, we set these thresholds to 0.50 for the PolityIV, 0.30 for the Freedom House, 0.40 for the artificial principal components indicator. This coding choice generates a similar number of roughly 80 democratic transitions across the three continuous democracy indices. By construction, democratization takes a value of one, if the dichotomous Democracy-Dictatorship index changes from zero to one. Due to the binary dependent variable, the empirical framework corresponds to a linear probability model that estimates the likelihood of democratization conditional on economic shocks and inequality. In the robustness section, we explore the sensitivity of the findings with respect to alternative assumptions about the choice of k and the threshold for the required change in the democracy indices.

In the baseline specifications of the empirical model, economic shocks are coded to occur before the democratization process started. Conceptually, the switch from autocracy to democracy and economic shocks might also overlap. However, in such a case, the economic shock might result from a feedback effect from democratization to economic performance and, thus, be endogenous. Therefore, our baseline coding choice presents a cleaner and more conservative view on the effect of income shocks and inequality on democratization. Nevertheless, we also report results for economic shocks that overlap with democratization in the robustness analysis.

The analysis is based on a yearly unbalanced panel of 130 countries for the period 1950–2014 with more than 3,000 country-year observations.[7]

[7]Table E.1 in the Appendix presents descriptive statistics of the data.

6.3 Empirical Results

6.3.1 Income, Income Shocks, Inequality, and Democracy

Income and Income Shocks. We begin the empirical analysis by replicating the standard specification in the related literature with democracy, measured by an index of the quality of political institutions, as dependent variable and income per capita as main regressor, using a country panel with annual observations. In addition, we add income inequality and an interaction between income and income inequality. Panel (a) of Table 6.1 presents the corresponding results, which effectively reproduce the analysis of Acemoglu et al. (2008) extended to the consideration of economic inequality and an interaction between income and inequality. The results confirm their main finding of no effect of income on democracy, once country and time fixed effects are accounted for. Moreover, the estimates provide no evidence for an interaction between income and inequality in shaping democratic institutions. The point estimates for the coefficients of income and the interaction term combined are not different from zero.[8]

Panel (b) of Table 6.1 goes beyond estimating the effect of marginal changes in income levels and investigates the effects of substantial negative economic shocks instead of continuous variations in income. The dependent variable is still an index measure of democratic quality measured on a discrete or continuous support. The results reveal that the occurrence of a negative cyclical income shock possesses a significant, adverse direct effect on democratic quality. According to the estimates in Column (1), a negative income shock leads to a reduction in democratic quality of 0.09 on the Polity IV index, which has been normalized to lie in the interval [0,1]. In contrast, the interaction term suggests that negative income shocks exert a *positive* effect on democratic quality, which increases with the level of income inequality. Given the empirical results from Column (1), the overall effect of a negative economic shock on democratic quality turns positive above a threshold value of roughly 0.47 for the lagged market Gini coefficient. This cutoff corresponds approximately to the 60th percentile of the lagged market Gini coefficient in the estimation sample. Correspondingly, the estimated marginal effect of an adverse cyclical shock on democratic quality would be positive for approximately 40 percent of the observations and negative for the remainder. Similar estimates obtain for other indicators of democratic quality. This result is a first piece of evidence that income shocks rather than minor fluctuations in income levels trigger major changes in institutional quality.

Democratic Quality vs. Democracy. As next step, we consider binary measures of democracy instead of index measures. Hence, the following analysis considers the effect of

[8]Figure E.1 in the Appendix illustrates this finding by comparing the (collapsed) unconditional variation between log GDP per capita and the PolityIV index with the residuals of both variables after partialling out country and time fixed effects.

Table 6.1: Income, Inequality, and Democracy

	Dependent variable is			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
	(a) Acemoglu et al. (2008) with inequality interaction			
Democratic Quality$_{t-1}$	0.86***	0.84***	0.80***	0.84***
	(0.02)	(0.02)	(0.02)	(0.02)
Income$_{t-1}$	0.02	0.01	0.01	0.02
	(0.03)	(0.02)	(0.05)	(0.03)
Inequality$_{t-1}$	0.42	0.19	-0.45	0.25
	(0.50)	(0.35)	(1.01)	(0.50)
(Income·Inequality)$_{t-1}$	-0.06	-0.02	0.01	-0.04
	(0.06)	(0.04)	(0.11)	(0.06)
Controls	✓	✓	✓	✓
Countries	128	133	131	125
Observations	3898	3794	3307	3026
R^2	0.82	0.78	0.71	0.79
	(b) Negative cyclical income shocks and inequality			
Democratic Quality$_{t-1}$	0.86***	0.84***	0.80***	0.84***
	(0.02)	(0.02)	(0.02)	(0.02)
Shock$_{t-1}$	-0.09***	-0.05*	-0.16***	-0.09***
	(0.03)	(0.02)	(0.04)	(0.03)
Inequality$_{t-1}$	-0.11	-0.04	-0.38**	-0.16
	(0.08)	(0.06)	(0.19)	(0.10)
(Shock·Inequality)$_{t-1}$	0.19***	0.12**	0.35***	0.20***
	(0.06)	(0.05)	(0.09)	(0.06)
Controls	✓	✓	✓	✓
Shocks	492	490	400	382
Countries	128	133	131	125
Observations	3881	3782	3290	3015
R^2	0.82	0.79	0.71	0.79

Notes: All regressions include country and time fixed effects. Average years of schooling is included as control for education. In Panel (b), log GDP p.c. is added to as control for income. The shock indicator in Panel (b) takes a value of one, if, within a time interval of three years, there is at least once a negative cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

income shocks on substantial changes in institutional quality instead of marginal changes of the respective index. Table 6.2 presents the baseline results for the effect of negative cyclical shocks, economic inequality, and their interaction on the probability of observing a shift to democracy. The first row of coefficients provide a mirror image of the previous results regarding the autocorrelation of institutional quality: Countries that originally possess high democratic quality are less likely to undergo a transition from a non-democratic to a democratic regime. The point estimates for the direct effect of an adverse cyclical shock

Table 6.2: Negative Cyclical Income Shocks, Inequality, and Democratization

	Democratization indicator based on			
	PolityIV Index (1)	Political Rights & Civil Liberties (2)	Democracy-Dictatorship (3)	Principal Components (4)
Democratic Quality$_{t-3}$	-0.29***	-0.35***	-0.32***	-0.39***
	(0.04)	(0.05)	(0.03)	(0.05)
Shock$_{t-3}$	-0.17***	-0.13**	-0.33***	-0.26***
	(0.05)	(0.06)	(0.08)	(0.07)
Inequality$_{t-3}$	-0.32**	-0.02	-0.82**	-0.54**
	(0.16)	(0.17)	(0.33)	(0.25)
(Shock·Inequality)$_{t-3}$	0.37***	0.31**	0.68***	0.55***
	(0.12)	(0.14)	(0.16)	(0.16)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Shocks	448	450	349	332
Countries	128	133	129	124
Observations	3678	3575	3036	2773
R^2	0.14	0.12	0.21	0.16

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a negative cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

reveal a negative sign and vary between –0.13 and –0.33. Conditional on the occurrence of such a negative income shock, the probability of democratization declines by 13 to 33 percentage points. This direct effect is quantitatively larger compared to a model that instead uses index measures of democratic quality. Again, this direct effect is moderated by the interaction between cyclical shocks and income inequality. The corresponding estimate possesses a positive sign for all measures of democratization and ranges between 0.31 and 0.68, which is again quantitatively larger than before. Throughout all specifications, the point estimates for the direct effect of a negative cyclical shock and its interaction with inequality are significant at the five-percent level.

The implied marginal effect of an economic shock on democratization is thus non-monotonic and can be either positive or negative conditional on the level of economic inequality. The marginal effect can be obtained by computing

$$\mathrm{ME}_{i,t}^{\mathrm{shock}} = \widehat{\beta} + \widehat{\gamma} \cdot x_{i,t-3} \,. \tag{6.2}$$

The point estimates from the first column therefore imply a positive marginal effect of a negative income shock on democratization for values of the market Gini in period $t-3$ above a cutoff of 0.46, and a negative marginal effect otherwise.

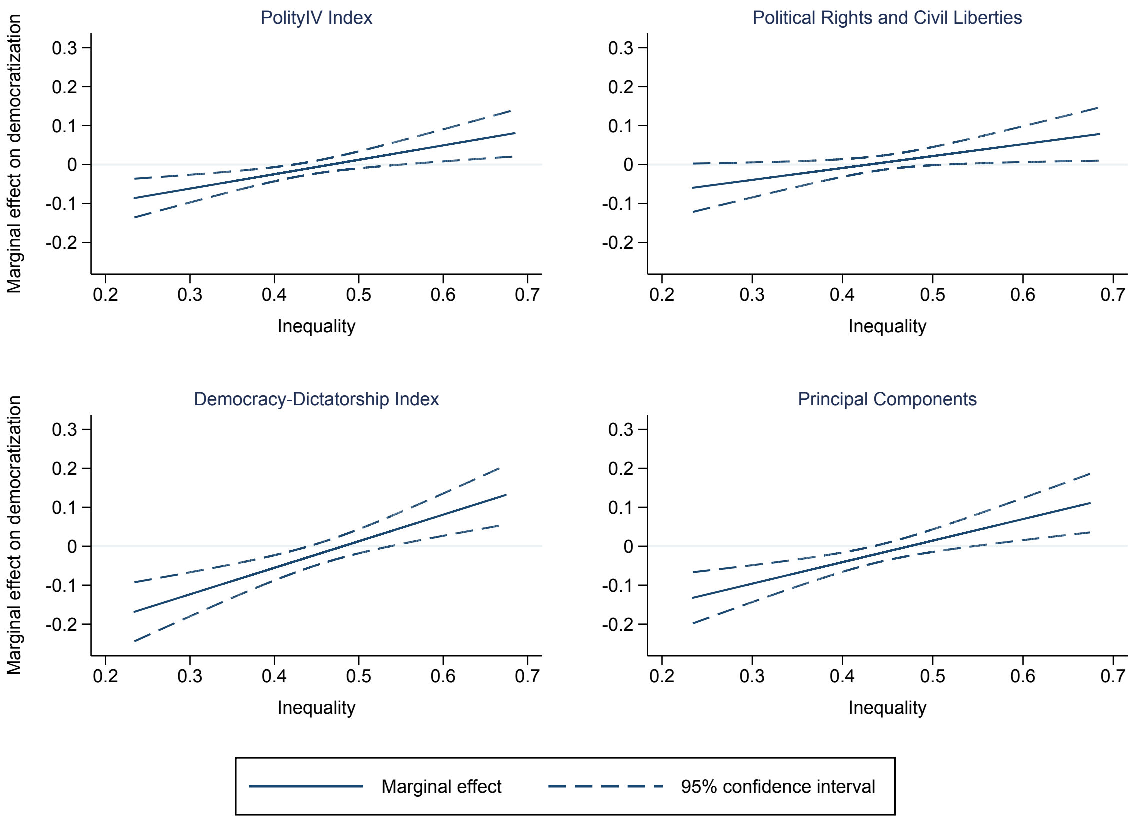


Figure 6.1: Effect of a Negative Income Shock on Changes in Democratic Quality

Notes: The marginal effects of negative income shocks are based on the estimates in Table 6.2.

Figure 6.1 depicts the marginal effects of a negative income shock on changes in democracy conditional on the level of inequality for all four specifications of Table 6.2. The solid line represents the respective marginal effect, whereas the dashed lines depict the corresponding 95-percent confidence intervals.[9] The estimated marginal effects are significant at the five-percent level for a sufficiently equal or unequal distribution of incomes as all panels in Figure 6.1 confirm. For intermediate levels of economic inequality, the marginal effects become quantitatively small and statistically insignificant.[10]

For the sample at hand, these results suggest that negative economic shocks unfold a positive marginal effect on the likelihood of a shift to democracy for roughly 53 percent of the country-year observations and a negative marginal effect for the remaining 47 percent.

[9]For a large sample, the confidence interval of the marginal effect is given by $ME \pm z_{1-\tau/2} \times \widehat{SE}$ with $\widehat{SE} = \sqrt{[Var(\widehat{\beta}) + x_{t-3}^2 \times Var(\widehat{\gamma}) + 2 \times x_{t-3} \times Cov(\widehat{\beta}, \widehat{\gamma})]}$ and $z_{1-\tau/2}$ being the critical value of a two-sided t-test of size τ.

[10]Figure E.2 in the Appendix shows the corresponding marginal effects of economic inequality on the likelihood of democratization conditional on the occurrence of a negative cyclical shock. The figure reveals that in the absence of negative cyclical shocks, a marginal increase in economic inequality reduces the likelihood of democratization. In the presence of an adverse cyclical shock, this hampering effect vanishes, which concurs with a window of opportunity for democratic transitions.

Given a domain of values for the market Gini coefficient in period $t-3$ of roughly $[0.23, 0.68]$, the estimated marginal effect varies across all specifications between $[-0.17, 0.13]$ with a zero effect for intermediate levels of inequality. Thus, in the most extreme case of inequality observed in the data, a negative cyclical shock increases the likelihood of democratization by 13 percentage points. In contrast, for the most equal society observed in the data, a negative cyclical shock reduces the probability of a substantial change in democratic quality by 17 percentage points.

Positive and Negative Shocks. Up to this point, the analysis restricted attention to negative income shocks and confirmed the finding of a non-monotonic effect of negative cyclical income shocks on the likelihood of democratization conditional on the extent of economic inequality. However, similar to the rationale of a window of opportunity for democratization during economic downturns, positive shocks might mute the support for a switch from autocracy to democracy during economic upturns and thereby stabilize autocracies. Analogous to the negative cyclical shock indicator, we thus construct a binary indicator variable for positive economic shocks, which takes a value of one, if, within a time interval of three years, there was at least once a positive cyclical shock of at least plus five percent, as expressed by the cyclical component of the HP-filtered income series. Sizable positive shocks are rare events and occur only in slightly more than five percent of the country-year observations. Following the same coding convention as for negative cyclical shocks, the positive shock indicator takes also value one if a sizable shock occurred during one of the two previous years (that is, $l = 2$). Overall, the shock indicator takes value one for roughly 15 percent of the observations with economic shocks being concentrated among low and middle income countries.

Table 6.3 presents the results for a specification with both positive and negative economic shocks. In order to account for the possibility of heterogeneous effects from positive and negative cyclical shocks, both variables enter the estimation equation separately. Therefore, countries that do not experience either shock constitute the reference category. Throughout all specifications, negative cyclical shocks again exhibit a negative direct effect on the likelihood of democratization. Moreover, the estimated interaction term between negative cyclical shocks and inequality shows a positive sign and is significant at the five-percent level throughout all specifications. Notably, the point estimates for the negative cyclical shock and the corresponding interaction term are quantitatively almost identical to the baseline estimates presented in Table 6.2. Positive cyclical shocks, in contrast, do not affect the likelihood of democratization; the results provide no evidence for either a direct or an indirect effect through the interaction with economic inequality. In light of this result, we focus attention on negative income shocks for the remainder of this paper.

Table 6.3: Cyclical Income Shocks, Inequality, and Democratization

	Democratization indicator based on			
	PolityIV Index (1)	Political Rights & Civil Liberties (2)	Democracy-Dictatorship (3)	Principal Components (4)
Democratic Quality$_{t-3}$	-0.29***	-0.35***	-0.33***	-0.39***
	(0.04)	(0.05)	(0.03)	(0.05)
Negative Shock$_{t-3}$	-0.17***	-0.13**	-0.33***	-0.26***
	(0.05)	(0.06)	(0.08)	(0.07)
Positive Shock$_{t-3}$	-0.04	-0.02	-0.10	-0.02
	(0.07)	(0.08)	(0.10)	(0.09)
Inequality$_{t-3}$	-0.33**	-0.01	-0.84**	-0.55**
	(0.16)	(0.17)	(0.33)	(0.25)
(Negative Shock·Inequality)$_{t-3}$	0.36***	0.31**	0.68***	0.56***
	(0.12)	(0.14)	(0.17)	(0.16)
(Positive Shock·Inequality)$_{t-3}$	0.12	0.08	0.27	0.09
	(0.16)	(0.18)	(0.21)	(0.20)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Negative shocks	448	450	349	332
Positive shocks	534	523	430	394
Countries	128	133	129	124
Observations	3678	3575	3036	2773
R^2	0.14	0.12	0.21	0.16

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a negative/positive cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

6.3.2 Robustness

This section briefly discusses the findings of robustness checks regarding alternative specifications or modifications in the coding of variables. The corresponding tables are presented in the Appendix.

Accounting for Multiple Imputation. The SWIID data set is based on imputation procedures to construct a comprehensive set of inequality estimates over time. As consequence, the data set provides numerous Gini data points are created for a given country and period (100 per year), which vary with respect to their respective (un-)certainty. Accounting for this uncertainty related to multiple imputation requires a time-consuming procedure in which the analysis of interest is conducted 100 times for different potential Gini candidate values. Table E.2 shows that this correction is inessential for the quantitative and qualitative results of our baseline specification.

Length of Time Window for Income Shocks. The construction of a variable of income shocks requires an assumption about the time window l during which the cyclical component is below a certain threshold relative to the trend component. A narrower window choice ($l < 2$) allows for a more precise timing of the effect but might miss out the effects of prolonged economic downturns. Instead choosing a longer window ($l > 2$) results in a larger number of "shock" observations and, thus, a noisier measure. Table E.3 shows that the main results do not change considerably for different coding choices (for example, $l = 0$ or $l = 4$) of the window length of economic shocks.

Sensitivity of the Trend/Cycle Decomposition. The construction of a variable of income shocks using a HP decomposition also requires an assumption about the smoothing parameter. Table E.4 explores how the effect of negative economic shocks on democratic institutions varies for different specifications of the smoothing parameter λ. In particular, cyclical shocks unfold a stronger effect on the likelihood of democratic transitions for a more smoothed long-run trend ($\lambda = 1$) with fewer shocks and a weaker effect for a less smoothed long-run trend ($\lambda = 100$) with more shocks.

Length of Time Window for Changes in Democracy. Ideally, democratization processes that are triggered by economic shocks should show first detectable results within a short period of time. However, it is conceivable that democratization processes require more time to be completed. If the switch from autocracy to democracy spans over a longer period, for instance because it is accompanied by civil conflict, economic shocks are unlikely to be the single most important determinant for the success of democratization. Table E.5 explores how the empirical results react to a different coding of the democratization period (for example, $k = 1$ or $k = 5$).

Thresholds for Changes in Binary Democracy Variable. The construction of a binary democratization variable requires an assumption about a threshold for the required change in the normalized democratic quality index. Table E.6 confirms that the qualitative findings of this paper do not hinge on the difference in thresholds for the construction of our dichotomous democratization variable.

Alternative Binary Democracy Measures. As an alternative robustness check, we apply our coding of negative cyclical shocks and democratic transitions to the concept of a binary democracy variable suggested by Acemoglu et al. (2016). In addition to their democracy indicator, their data additionally contains the binary measure constructed by Papaioannou and Siourounis (2008). Both measures of democratic quality are based on sizable changes in the democracy indicators by PolityIV and Freedom House in conjunction with certain stability criteria for successful democratic transitions. Table E.7 shows that applying our coding choices to the data set of Acemoglu et al. (2016) yields the same qualitative results with quantitatively even slightly larger coefficients. The respective cutoff values for the market Gini coefficient in period $t-3$ are quantitatively almost identical to our baseline specification. The implied marginal effects are slightly larger in absolute terms; that is, they take somewhat higher values for a given level of economic inequality. This result demonstrates that our baseline results do not hinge on coding conventions regarding democratic quality. Moreover, our baseline results provide a conservative view on the role of economic shocks and inequality for the likelihood of democratic transitions.

Time Overlap Between the Occurrence of Shocks and Changes in Democracy. As discussed above, economic shocks are coded to occur before the democratization process started in the baseline specifications of the empirical model. Conceptually, the switch from autocracy to democracy and economic shocks might also overlap, though implying that the baseline specification applies a timing restriction that is too conservative for capturing the full effect. However, this timing excludes direct feedback effects. To explore the robustness of the results with respect to this timing convention, we estimated models that allow for the democratization process and the cyclical shocks to overlap. Explicitly, we still measure changes in democratic quality between t and $t-k$. Economic shocks, however, enter the empirical analysis either in period $t-k+1$ or $t-k+2$, thus generating an overlap between economic shocks and democratic transitions. As expected, in this case the estimated interaction between economic shocks and inequality is, in fact, larger than in the baseline specification as documented in Table E.8.

Logit Estimates. While the linear probability model is preferable in the context of specifications with interactions, the use of a binary dependent variable suggests the use of a logit estimator to check robustness. As shown in Table E.9, such an estimation produces similar qualitative results for a conditional logit regression as for the linear probability model. However, due to the small number of major political transitions and economic shocks in the sample, the logistic regression framework drops a large number of observations for which the estimated likelihood would diverge to infinity. The findings from a logit regression should thus be seen as suggestive at best.

6.3.3 Additional Results

This section reports further results with respect to different measures of economic shocks and discusses the role of demographic characteristics for the likelihood of democratization.

Income Trends. The specifications in Section 6.3.1 considered short-term cyclical income shocks together with economic inequality as key determinant for democratization. Long-run growth trends constitute an alternative angle to approach the definition of economic shocks. A spell of prolonged stagnation or shrinkage presents a potential environment for individual agents to voice their discontent with the existing political

Table 6.4: Negative Long-Run Growth Trends, Inequality, and Democratization

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
Democratic Quality$_{t-3}$	-0.29***	-0.35***	-0.32***	-0.39***
	(0.04)	(0.05)	(0.03)	(0.05)
Trend$_{t-3}$	-0.22***	-0.19*	-0.24*	-0.22**
	(0.07)	(0.11)	(0.13)	(0.09)
Inequality$_{t-3}$	-0.32**	-0.02	-0.84**	-0.56**
	(0.15)	(0.16)	(0.33)	(0.25)
(Trend·Inequality)$_{t-3}$	0.49***	0.49*	0.50*	0.47**
	(0.18)	(0.25)	(0.29)	(0.22)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Trends	256	263	242	237
Countries	128	133	129	124
Observations	3663	3561	3022	2759
R^2	0.14	0.12	0.20	0.16

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The trend indicator takes a value of one, if the real, HP-filtered GDP p.c. series shrinks by at least five percent over a time interval of three years, and zero else. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

institutions. According to this rationale, we code an adverse (that is, negative) economic trend indicator which takes a value of one, if the real, HP-filtered GDP per capita series shrinks by more than five percent over a time period of three years, and zero else. Prolonged trends of shrinking real income occur in roughly seven percent of the country-year observations with again a strong concentration in low and middle income countries.

Table 6.4 reports results for the negative growth trend dummy variable instead of the cyclical shock indicator. The estimated parameters confirm the finding of a non-monotonic effect of economic distress and inequality on the likelihood of democratization. The point estimates show a similar qualitative pattern as those obtained in the baseline model for negative cyclical shocks reported in Table 6.2. In comparison, however, the computed coefficients in Table 6.4 vary less across the different specifications. The estimated parameter for adverse trends and their interaction with economic inequality are significant at the ten-percent level throughout all specifications with the most precise estimates for the PolityIV index. Over all specifications, the implied marginal effects of a negative growth trend on the probability of democratization are quantitatively similar to the baseline estimates and lie within the interval $[-0.13, 0.14]$. Hence, conditional on the level of economic inequality, a prolonged spell of shrinking real per capita income reduces the likelihood of democratization up to 13 percentage points for the most equal society observed in the data. On the other extreme, a negative growth trend increases the likelihood of a democratic transition by 14 percentage points for the most unequal society observed in the sample. For intermediate levels of inequality, the marginal effects are again close to zero and statistically insignificant.

Inflation Shocks. Up to this point, economic shocks were defined as cyclical fluctuations around a long-run growth trend, or the trend itself. Another way of modeling economic shocks is to consider hikes in inflation rates. Many low income countries, for example, heavily rely on the export of certain agricultural products and natural resources. Fluctuations in international commodity prices, however, may put substantial strain on per capita incomes and open a window of opportunity for changes in political institutions.[11] Hence, we construct an economic shock indicator based on price changes derived from the GDP deflator. Because countries, which catch up economically, typically experience high inflation rates during the convergence process, we limit attention to cases with rather drastic price changes. Therefore, the economic shock indicator takes a value of one, if the inflation rate within a given year equals or exceeds 20 percent, and zero otherwise. In order to be consistent with the coding of negative economic shocks in the baseline specification, the indicator also takes a value of one, if the inflation rate exceeded the cutoff value in either of the two previous years. In total, approximately 24 percent of the country-year observations of the data set classify as periods with high inflation. Most of these cases are observed in low and middle income countries with strong persistence regarding the (in-)stability of prices. For data availability, the estimation sample slightly shrinks by 300 to 400 country-year observations.

Table 6.5 reports results for the effect of inflation shocks and inequality on the likelihood of democratization. The point estimates are of similar magnitude as those obtained for the baseline specification. Due to the considerably larger number of shocks, the coefficients are however less precisely estimated compared to the model with negative cyclical shocks. Hence, the empirical model lacks the power to estimate an interaction term that is statistically different from zero for the specifications in Columns (2) and (4). Overall, the findings again confirm the non-monotonic pattern of economic shocks on the democratic transitions conditional on the degree of economic inequality. Given the distribution of market Gini coefficients in period $t-3$, the marginal effects of an inflation shock on the probability of democratization lie within the interval $[-0.14, 0.10]$ which conforms closely to the range of marginal effects obtained for the baseline model. We interpret this result as another piece of evidence that shocks to (disposable) income rather than minor fluctuations in income levels trigger major changes in democratic institutions.

[11] See, for example, work by Brückner, Ciccone, and Tesei (2012), which exploits oil price shocks to identify changes in democratic quality.

Table 6.5: High Inflation, Inequality, and Democratization

	Democratization indicator based on			
	PolityIV Index (1)	Political Rights & Civil Liberties (2)	Democracy-Dictatorship (3)	Principal Components (4)
Democratic Quality$_{t-3}$	-0.31***	-0.33***	-0.33***	-0.38***
	(0.05)	(0.05)	(0.03)	(0.06)
Shock$_{t-3}$	-0.18***	-0.13*	-0.26**	-0.15
	(0.06)	(0.08)	(0.11)	(0.10)
Inequality$_{t-3}$	-0.35*	-0.06	-1.00**	-0.55*
	(0.19)	(0.15)	(0.40)	(0.31)
(Shock·Inequality)$_{t-3}$	0.40***	0.29	0.53**	0.31
	(0.13)	(0.18)	(0.23)	(0.21)
Controls	✓	✓	✓	✓
Transitions	64	61	94	64
Shocks	715	682	639	606
Countries	123	123	118	118
Observations	3239	3132	2547	2440
R^2	0.15	0.10	0.20	0.14

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, at least once within a time interval of three years, the inflation rate, measured by the GDP deflator, exceeds a threshold of 20 percentage points. Inequality refers to the market Gini coefficient. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

The Role of Demography. Finally, the demographic structure of a country represents another possible driving force behind political transitions. In particular, conflict and struggle for power become more likely if young individuals, "those who are in demand of land, jobs, higher education, opportunity, and other kinds of resources in society" (Fuller, 1995, p. 152), constitute a large portion of the total population. Building on Fuller's (1995) insights, the probability of a democratic transition should thus be higher for economically unequal countries that experience a "youth bulge", that is, a large share of youths that causes a bulge in the population tree. Following this reasoning, we interpret a high proportion of youths as measured by the share of 15- to 24-year-olds in the population as demographic pressure.[12] In our sample, the mean share of 15- to 24-year-olds in the population corresponds to approximately 18 percent with values ranging between 9 and 26 percent. Moreover, we interact demographic pressure and economic inequality in order to allow for a window of opportunity.

Columns (1), (3), (5), and (7) of Table 6.6 present the results of a linear probability model, which extends the baseline specification with negative cyclical shocks by including the population share of the 15- to 24-year-olds and its interaction with economic inequality as additional variables to control for demographic pressure.[13] The results confirm a non-

[12]This categorization follows Fuller (1995) and represents "the conventional cutoff for youth in the literature" (Nordås and Davenport, 2013, p. 932).

[13]Table E.10 in the Appendix reports results for an empirical model without negative income shocks.

Table 6.6: Demographic Pressure as Determinant of Democratization

	Democratization indicator based on							
	PolityIV Index		Political Rights & Civil Liberties		Democracy-Dictatorship		Principal Components	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Democratic Quality$_{t-3}$	-0.31***	-0.31***	-0.35***	-0.35***	-0.33***	-0.33***	-0.40***	-0.40***
	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)	(0.03)	(0.06)	(0.06)
Shock$_{t-3}$	-0.18***	-0.23***	-0.14**	-0.19**	-0.33***	-0.33***	-0.27***	-0.28***
	(0.05)	(0.08)	(0.06)	(0.10)	(0.07)	(0.11)	(0.07)	(0.10)
Inequality$_{t-3}$	-1.32***	-1.34***	-0.99*	-1.01*	-0.97	-0.97	-1.27*	-1.27*
	(0.48)	(0.49)	(0.59)	(0.59)	(0.78)	(0.78)	(0.67)	(0.67)
(Shock·Inequality)$_{t-3}$	0.38***	0.37***	0.32**	0.30**	0.68***	0.68***	0.56***	0.56***
	(0.11)	(0.11)	(0.14)	(0.14)	(0.16)	(0.16)	(0.15)	(0.15)
(Sh.15–24)$_{t-3}$	-2.10*	-2.14*	-2.57*	-2.63*	0.17	0.17	-1.61	-1.63
	(1.12)	(1.12)	(1.38)	(1.38)	(1.98)	(1.99)	(1.60)	(1.60)
(Sh.15–24·Inequality)$_{t-3}$	6.36**	6.38**	5.66*	5.71*	1.38	1.38	4.48	4.50
	(2.48)	(2.48)	(3.10)	(3.09)	(4.21)	(4.22)	(3.54)	(3.54)
(Shock·Sh.15–24)$_{t-3}$		0.31		0.35		-0.00		0.07
		(0.29)		(0.39)		(0.50)		(0.38)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Transitions	76	76	81	81	114	114	79	79
Shocks	448	448	450	450	349	349	332	332
Countries	128	128	133	133	129	129	124	124
Observations	3678	3678	3575	3575	3036	3036	2773	2773
R^2	0.15	0.15	0.12	0.12	0.21	0.21	0.16	0.16

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a negative cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Sh.15–24 measures the share of 15- to 24-year-olds in the total population. Inequality refers to the market Gini coefficient. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

monotonic effect of negative cyclical shocks on the likelihood of democratic transitions conditional on the extent of economic inequality. The reported estimates and the implied marginal effects conform almost perfectly to those obtained for the baseline specifications in Table 6.2.

The computed estimates in Columns (1) and (3) parallel the previous qualitative findings of a non-monotonic effect for demographic pressure as potential window of opportunity. Specifically, the likelihood of democratization decreases with an increasing share of youths. However, this effect is moderated by economic inequality through the interaction term. For more unequal societies, for example, above a Gini coefficient of 0.45 according to estimates of the third column, a youth bulge implies an increased likelihood of democratization. Correspondingly, the marginal effect of demographic pressure on the likelihood of democratic transitions is positive for approximately one half of the country-year observations in the sample and negative else. However, the respective

These parsimonious specifications produce parameter estimates for demographic pressure and its interaction with economic inequality that are qualitatively and quantitatively similar to those reported in Table 6.6.

parameter estimates for the youth share and its interaction with economic inequality are not statistically different from zero at conventional significance levels for the specifications in Columns (5) and (7). Hence, the results indicate a potential role of demographic pressure for democratic transitions; though, we view this evidence as too weak to be conclusive.

Finally, Columns (2), (4), (6), and (8) report results for an extended model in which economic shocks and demographic pressure may interact. Throughout all specifications, the estimated coefficient of the interaction is small and not statistically different from zero at the conventional significance levels. Moreover, including this interaction term does not considerably change the qualitative and quantitative findings. Hence, economic shocks and demographic pressure operate not jointly but in parallel to each other.[14]

6.4 Conclusion

This paper documents novel cross-country panel evidence for a non-monotonic effect of income on democracy. In particular, the evidence suggests that when focusing on major fluctuations in income rather than continuous variation in income levels, income shocks have a significant effect on democratic quality. Moreover, the results reveal an important asymmetry of this effect. Negative income shocks exhibit a significantly negative effect on democracy, whereas no comparable countervailing effect is found for positive income shocks. In the absence of shocks, inequality has a deteriorating effect on democracy. Additionally, negative income shocks reveal an important interaction effect with economic inequality. Negative income shocks lead to a deterioration of democratic quality in equal societies, whereas they entail an improvement in democratic quality in unequal societies, as illustrated in Figure 6.1. No such interaction is found for positive income shocks. This suggests that negative economic shocks might initiate democratic movements in an environment with high inequality. By highlighting the role of asymmetric shocks and their interaction with inequality, the results shed new light on the intricate relation between income and democracy and on the seemingly contradictory findings in the literature.

[14]Table E.11 in the Appendix presents results for an empirical model that additionally incorporates a triple interaction term between adverse cyclical shocks, demographic pressure, and economic inequality. The estimated triple interaction term, however, possesses large confidence intervals and is thus insignificant at conventional significance levels throughout all specifications. In combination with the statistically insignificant interaction between demographic pressure and cyclical shocks reported in even columns of Table 6.6, the results preclude a considerable interaction between demographic pressure and economic shocks. Therefore, this evidence further indicates that both channels operate in isolation.

Chapter 7

Democracy, Inequality, and Institutional Quality[1]

7.1 Introduction

The quality of institutions granting economic freedom and liberties is generally thought to be a crucial factor for long-run development. A recent empirical literature has identified two key determinants of institutional quality. First, democracy provides constraints on those in power and is often used as a proxy for the quality of economic institutions (Acemoglu, Johnson, and Robinson, 2005; Acemoglu and Johnson, 2005; Acemoglu and Robinson, 2005; Acemoglu, 2008). Second, a moderate inequality of income and economic resources limits the distributive pressures that might erode institutional quality through influence activities and informality (Chong and Calderon, 2000; Chong and Gradstein, 2007a, 2007b).

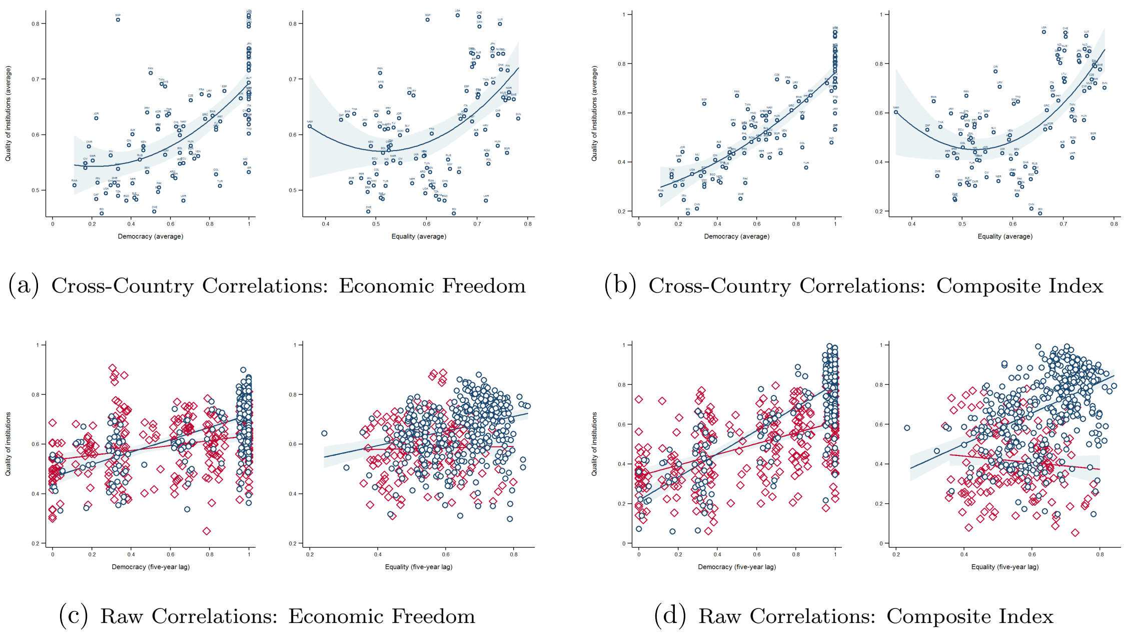
Based on intuitive reasoning, however, it has been claimed repeatedly that the beneficial effect of democracy on the quality of economic institutions might be eroded by excessive inequality. In fact, as is discussed in more detail below, there is a considerable theoretical literature that underpins this conjecture and suggests that democracy and inequality do not affect the quality of economic institutions independently from each other. Instead, there might be important interactions between inequality and political institutions in determining institutional quality. This conjecture seems to have been present in the works by De Tocqueville (1835) as well as Lipset (1959), who saw economic equality as a central co-factor of democracy in generating good institutions. More recently, this conjecture has been rephrased in the context of the debate regarding the increasing inequality and the consequences for civil liberties and social peace (Piketty, 2014). Yet, to date, there exists little to no evidence on the question as to whether the beneficial effect of democracy on institutional quality is eroded by excessive inequality.

[1] The final version of this paper is published in the *European Economic Review*, 91, 209–228. Please refer to the published version.

This paper tests if such an interaction between democracy and equality exists in shaping the de facto quality of economic institutions. The starting point of the empirical analysis is the conceptual distinction between political and economic institutions. There is a widespread perception in the literature that democracy, as measured by de jure measures, such as constraints on the executive or political freedom, is equivalent to high de facto quality of institutions as experienced by individuals, particularly in the economic domain. While both conform to North's notion of "rules of the game" (North, 1990, p. 3), political institutions describe the extent to which individuals can engage and participate in the political process via elections and referendums, economic institutions comprise aspects of de facto economic freedom, as well as institutional features that directly affect the incentives for entrepreneurial activities and investment, such as bureaucratic efficiency and impartiality of the judiciary. There are important conceptual differences, which refer to the nature of institutions, as well as their perception. Economic institutions are mostly implemented by laws that have been passed by the government and reflect de facto liberties of individual citizens in the economic domain. In contrast, political institutions, in terms of democracy, the constraints on the executive, or the ability to vote, reflect legally codified, constitutional rules. In this sense, political institutions can be seen as determinants of economic institutions but not vice versa (see, for example, Acemoglu, Johnson, and Robinson, 2005). Likewise, some authors have pointed out the inherently different nature of institutions by emphasizing the different functions of political institutions as constraints for politicians and the government, as opposed to economic institutions that enable private actors to interact and achieve their goals (see, for example, Voigt, 2013).

Figure 7.1 provides a first indication that democracy and institutional quality are not necessarily equivalent. Panels (a) and (b) illustrate the correlation of country averages of democracy and equality with institutional quality over the time period 1970 to 2010. Institutional quality is measured either in terms of an index of Economic Freedom or in terms a composite index that is based on the first principal component of the Economic Freedom and the Civil Liberties indices. Panels (c) and (d) plot the respective raw data for the same time frame in terms of five-year averages, which is the variation in the data used in the empirical analysis below (jittered for better visibility). The figure on the left in each panel plots the quality of economic institutions against the extent of democracy, proxied by Constraints on the Executive. In the figure on the right, the quality of economic institutions is plotted against equality, proxied by $(1 - \text{Gini})$. Henceforth, we refer to this measure as the reversed Gini coefficient. All variables are normalized to range between 0 and 1 with higher values indicating either a higher institutional quality, more democracy, or more equality.

The data reveal the expected positive correlation between institutional quality and democracy, but contrary to the widespread assumption that democracy and high quality institutions are essentially synonyms, there is substantial variation in institutional quality,



(a) Cross-Country Correlations: Economic Freedom

(b) Cross-Country Correlations: Composite Index

(c) Raw Correlations: Economic Freedom

(d) Raw Correlations: Composite Index

Figure 7.1: Correlation Between Institutional Quality and Democracy or Income Equality

Notes: Institutional Quality is measured either by the Economic Freedom index or by a composite index based on the first principal component of the Economic Freedom and the Civil Liberties indices, respectively. Democracy is measured in terms of constraints on the executive. Income Equality is $(1 - \text{Gini})$. Variables and data sources are described in more detail in Section 7.3, see also Table F.1.

even conditional on the same level of democracy. In particular, the figures illustrate that there are countries that exhibit very democratic political institutions but that at the same time have low scores of institutional quality, such as India. Other countries with autocratic regimes (that is, low democracy scores) score high in terms of the quality of economic institutions. Examples are Jordan, Panama, or Singapore, which achieved a considerably high quality of institutions despite relatively low levels of democracy. When looking beyond country averages, another interesting case is Brazil, which scored high on institutional quality during the military dictatorship during the 1980s.[2] Incidentally, these countries exhibit relatively moderate levels of equality, as measured by the world-wide distribution of reversed Gini coefficients.

The figures on the right in each panel of Figure 7.1 show the corresponding relationship between institutional quality and the degree of income equality. There is no clear correlation pattern. However, as suggested by the non-linear fit in panels (a) and (b) or a categorization into democratic and non-democratic observations in panels (c) and (d), the relationship between equality and institutional quality appears to be heterogeneous or even non-monotonic. In particular, the data patterns suggest a positive correlation between equality and economic freedom for democracies but a negative correlation for non-democracies.[3]

[2] Similar levels of economic institutions have been achieved only during the boom years between 2005 and 2010, fueled by revenues from the state owned oil company. The period after democratization, on the other hand, had been characterized by high corruption, involving bribes of high-ranking politicians through the state owned oil company and large construction companies.

[3] The democracy split is based on the binary Democracy-Dictatorship index constructed by Cheibub,

This is mirrored in the correlation between democracy and institutional quality. This correlation is substantially weaker when economic inequality is high as compared to when inequality is low.

Taken together, the data depicted in Figure 7.1 suggest that democracy and equality might potentially reinforce each other in shaping institutional quality. Observations exhibiting high scores of democracy and equality are likely to exhibit the highest institutional quality, whereas institutional quality might also be (relatively) high for observations characterized by low equality and low scores of democracy. In particular, the two graphs suggest the possibility of a positive interaction between democracy and equality in determining institutional quality. For example, for India or Brazil, or some other Latin American countries, there appears not to have been a clear positive correlation between democratic quality and institutional quality, while at the same time, inequality was high. At the opposite ends of the spectrum, there are examples or countries with low inequality, where democratization was conducive to the quality of economic institutions, such as several former socialist states in Eastern Europe.[4]

The main contribution of this paper is the identification of a robust empirical interaction between democracy and equality in shaping the quality of the economic institutions. The empirical analysis is based on cross-country panel data ranging over the period 1970 to 2010, which are behind the patterns shown in Figure 7.1. Corresponding to the theoretical arguments in the existing literature, the empirical framework treats economic institutions as the dependent variable that is determined by the quality of political institutions in terms of democracy and constraints on the executive, as well by the prevailing level of economic (in-)equality.[5] The empirical strategy exploits variation in democratic quality and income equality within countries over time, thereby conditioning on country-specific and time-specific unobserved heterogeneity that might influence institutional quality, and on controls for institutional quality in the past. Different estimation methods are employed in order to account for the well-known problems in dynamic panels. Irrespective of the estimation method and the underlying identification assumptions, the results reveal a robust, significant positive interaction between democracy and equality in shaping the quality of economic institutions. In terms of size, this interaction term is large enough to render the effect of democracy on institutional quality negative for high levels of inequality within the range of what is observed in the data. Hence, the results provide evidence for a relevant dimension of heterogeneity in the way income inequality and democracy affect the quality of economic institutions.

Gandhi, and Vreeland (2010). Similar results obtain for splits based on alternative democracy indices.

[4]Likewise, there are highly unequal, non-democratic countries, like the Gulf states or some East Asian countries, that appear to implement high quality economic institutions. Due to missing inequality data, these countries are not included in the data sample, however.

[5]Clearly, there are indirect feedback effects from economic institutions through income growth, unemployment, etc. to political institutions. These effects need time to work and are accounted for accordingly in the empirical model, as discussed in detail below.

By documenting a robust interaction between political institutions and inequality in shaping institutional quality, this paper contributes to several branches of the literature. The analysis provides evidence for an untested implication of an entire strand of theoretical papers that will be discussed in the next section. The results also contribute to a small empirical literature that has tried to identify the role of democracy and inequality for institutional quality. Chong and Gradstein (2007b) provide evidence for a two-way causality between the quality of institutions and a more equal distribution of income. In related work, Chong and Gradstein (2007a) show that greater inequality can erode institutional quality through fostering inequality. However, both papers do not investigate potential interaction effects between inequality and democracy in affecting institutional quality. The empirical contribution closest to the present investigation is by Sunde, Cervellati, and Fortunato (2008). They provide evidence for an interaction effect when using cross-country variation. Given the problems of omitted variables and unobserved heterogeneity that abound in cross-section regressions, their evidence can at best be seen as suggestive. The present study shows that the finding of an interaction effect also emerges in a panel setting after eliminating unobserved time-invariant cross-country heterogeneity, accounting for past institutional outcomes, and applying state-of-the-art panel data estimation methods. Besides this methodological contribution, our analysis also provides a more direct test of the theoretical prediction, which stipulates an interaction of democracy and (in-)equality over the development path, that is, within the same country over time. The empirical findings also have implications for the interpretation of some of the earlier results on the role of institutions for development—for example, Rodrik, Subramanian, and Trebbi (2004)—because they suggest a potentially more important role than previously thought of political institutions and inequality. Moreover, this paper contributes to an ongoing debate about how institutions should be measured, and what is comprised by different measures that are frequently used in the literature. In particular, while there is a common perception that different measures capture similar underlying institutional features, our findings document that there is substantial variation across measures of political and economic institutions, and that the correlation is lower than commonly thought, complementing the conceptual arguments by Voigt (2013). To our knowledge, our analysis is among the first to open this black box and test a theoretical prediction using a variety of different measures of political and economic institutions. The results indicate a very robust interaction between political institutions and (in-)equality within and across countries that has gone largely unnoticed in the existing literature and that holds across combinations of various measures.

The remainder of the paper is structured as follows. Section 7.2 presents the theoretical background that motivates the empirical analysis. Section 7.3 describes the data and their sources. Section 7.4 presents the empirical model and identification strategy. Section 7.5 presents the main empirical results as well as the results of extensive robustness checks. Section 7.6 concludes.

7.2 Theoretical Background

The hypothesis that political institutions and inequality interact in shaping institutional quality goes back as far as De Tocqueville (1835), who recognized the possible problems associated with democracy in societies characterized by large economic inequality, in particular, the possibility of a deterioration of equality of rights. At the same time, he conjectured that equality was a prerequisite for a functioning democracy, because "(w)hen no member of the community has much power or much wealth, tyranny is, as it were, without opportunities"(De Tocqueville, 1835, Book 4, Chapter 6).[6] Intuitive arguments suggesting the existence of an interaction between democracy and inequality in shaping institutional quality like those by De Tocqueville (1835) also seem to closely resemble the view by Lipset (1959), according to whom not only income but also equality are the key prerequisites of universal civil (and presumably economic) freedom.[7] More recently, this view has been been extended by arguments that excessive inequality leads to a breakdown of institutional quality in democracies, see, for example, Piketty (2014).

The starting point for the empirical analysis is the commonly accepted perception of the hierarchy of institutions, as reflected in the schematic model displayed in Figure 7.2. This conceptualization of the interplay between political institutions, inequality, and institutional quality experienced by the citizens of a country adapts the framework of institutional and economic development described by Acemoglu, Johnson, and Robinson (2005). They view political institutions and the distribution of resources as the two state variables for analyzing institutional and economic dynamics. Through their influence on de facto political power, these state variables determine the actual quality of institutions that individuals experience and take as a basis for their decisions, in terms of, for example, economic institutions or economic freedom. The consequences materialize both in terms of economic performance and the future distribution of resources.[8] This framework explicitly

[6]Alternatively, in Book 2, Chapter 3, de Tocqueville writes: "When the conditions of men are very unequal, and inequality itself is the permanent state of society, individual men gradually become so dissimilar that each class assumes the aspect of a distinct race: only one of these classes is ever in view at the same instant; and losing sight of that general tie which binds them all within the vast bosom of mankind, the observation invariably rests not on man, but on certain men. Those who live in this aristocratic state of society never, therefore, conceive very general ideas respecting themselves, and that is enough to imbue them with an habitual distrust of such ideas, and an instinctive aversion of them. He, on the contrary, who inhabits a democratic country, sees around him, on every hand, men differing but little from each other; he cannot turn his mind to any one portion of mankind, without expanding and dilating his thought till it embraces the whole. All the truths which are applicable to himself, appear to him equally and similarly applicable to each of his fellow-citizens and fellow-men."

[7]In his famous article, Lipset writes that low inequality is consistent with democracy and individual freedom: "From Aristotle down to the present, men have argued that only in a wealthy society in which relatively few citizens lived in real poverty could a situation exist in which the mass of the population could intelligently participate in politics [...] A society divided between a large impoverished mass and a small favored elite would result either in oligarchy [...] or in tyranny." (Lipset, 1959, p. 77).

[8]The framework is an adaptation of that presented by Acemoglu, Johnson, and Robinson (2005), who write that "(t)he two state variables are political institutions and the distribution of resources, and the knowledge of these two variables at time t is sufficient to determine all the other variables in the system.

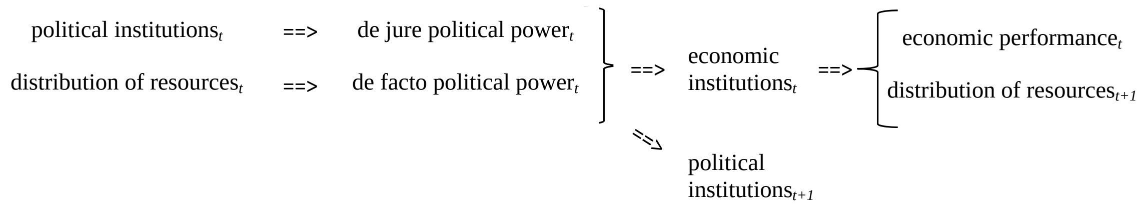


Figure 7.2: A Conceptual Framework of Institutional and Economic Dynamics (Adapted from Acemoglu, Johnson, and Robinson, 2005, p. 392)

accounts for the possibility that economic development and wealth accumulation affect political institutions, as argued in the political economy literature (see, for example, Olson, 1982). This framework can be implemented empirically by exploiting the difference in the available empirical measures of institutions between the de jure political institutions and the de facto liberties and freedoms in the economic domain. According to the view of political institutions and the distribution of resources as state variables in the dynamic process, the empirical framework treats economic institutions as dependent variable and the quality of (de jure) political institutions as well as the level of economic inequality as explanatory variables.[9]

This conceptual framework does not yet provide testable hypotheses regarding the signs and form of the influence of the state variables on institutional quality. The commonly accepted view in the literature, including Acemoglu, Johnson, and Robinson (2005), is that inclusive political institutions, such as democracy, as well as moderate inequality in the distribution of resources, are conducive to high-quality economic institutions and, thus, a favorable environment for economic performance.

A series of formal models that study the institutional and economic dynamics in more detail also predict an interaction of political institutions, in terms of democracy, and inequality in determining the quality of economic institutions. For instance, Cervellati, Fortunato, and Sunde (2008) investigate the interactions between political regime (oligarchy versus democracy) and inequality in shaping the social contract in terms of rule of law, and derive conditions under which rule of law emerges in equilibrium. In their model, rule of law can emerge and persist in oligarchies, if economic inequality is sufficiently large. Under these conditions, this equilibrium is preferred by all groups of society to an equilibrium involving a democratic regime but wasteful social conflict. The reason is the excessive distributive pressure that reduces the incentives for the protection of private property rights by the state. A direct consequence is that democracy is not sufficient to

While political institutions determine the distribution of de jure political power in society, the distribution of resources influences the distribution of de facto political power at time t. These two sources of political power, in turn, affect the choice of economic institutions and influence the future evolution of political institutions" (Acemoglu, Johnson, and Robinson, 2005, p. 392).

[9]In spirit, the framework captures the main elements of the framework of Acemoglu, Johnson, and Robinson (2005), which suffers from the shortcoming that de facto political power is hard to measure empirically using the available measures.

implement good institutions, and can, in the context of excessive inequality, lead to even lower quality of the economic institutions. At the same time, democracy is instrumental for implementing high institutional quality when inequality is low. A similar prediction emerges out of dynamic models of franchise extension that allow for different scenarios of democratization and show that peaceful transitions to democracy are more likely to occur in societies in which resources are distributed more equally. In turn, peaceful transitions to democracy lead to greater improvements in institutional quality, as well as greater stability of democracy, see, for example, Cervellati, Fortunato, and Sunde (2012, 2014).

A non-monotonic effect of democracy on institutional quality is also consistent with the predictions of the model by Acemoglu and Robinson (2008), which is based on the distinction between de jure political power reflected in democratic institutions, and the distribution of de facto political power, which is the result of economic power as reflected by inequality of income and wealth. The model illustrates how an elite can use their wealth to influence electoral outcomes. An elite that commands a large fraction of income and resources in the economy can use these to capture the state and enforce institutions that serve their own interests. If economic resources are relatively evenly spread among the population, it is difficult for the elite to direct sufficient funds and amass enough de facto power to bias institutions in their favor, thereby eroding the quality of institutions. Consequently, oligarchic structures and high inequality complement each other in restricting institutional quality and civil liberties for the masses, and vice versa.

Alternatively, Acemoglu, Robinson, and Torvik (2013) consider a political environment, where groups of the population can bribe the elected government in order to achieve the desired institutional outcome, provided they overcome a collective action problem. In weakly institutionalized polities, where such bribe payments may play a role, voters prefer to allow rents for the government in order to reduce incentives for bribes to be accepted. Consequently, the sovereign will deliberately remove some of the checks and balances so that the elite cannot successfully influence the executive and legislative. This leads to two possible equilibrium constellations. On the one hand, there is a properly working democracy, where the elite is not influential enough such that bribes do not play a significant role for political decisions; in this scenario citizens can exert their democratic rights to the full extent. On the other hand, if the elite is influential enough to effectively control the government via bribes and implement policies favoring themselves, the government is either de jure democratic but de facto autocratic, or the people remove some checks and balances to reduce the incentives for corruption, rendering the government de jure autocratic. Hence, "checks and balances are less likely to emerge [...] when inequality and taxes are quite high" (Acemoglu, Robinson and Torvik, 2013, p. 845). Hence, institutional quality will depend on an interaction of inequality and the constraints imposed on the executive, de facto or de jure, consistent with the empirical hypothesis of a non-monotonic effect of democracy on institutional quality conditional on equality.

A common feature of these theoretical considerations is that inequality is a determinant of, or even a proxy for, the de facto power of the elite as well as a key factor behind the redistributive pressure faced by the elite. If inequality is low, a democracy that provides efficient economic institutions can be implemented, because the elite is relatively weak and has no incentive for influencing the institutions. On the other hand, if inequality is high, the elite either has the means and the incentive to control the government indirectly via de facto power, or it can even directly take over because the poor majority accepts a social contract with limited franchise in exchange for more efficient economic institutions. The consequence is low institutional quality in terms of civil liberties and economic freedom for the broad majority.

The non-monotonicity immanent in each of these theories implies the testable hypothesis of a positive interaction of equality and political institutions, in shaping institutional quality. The following empirical analysis is devoted to testing this hypothesis.

7.3 Data

In order to empirically assess the effects of equality and democracy on institutional quality, we construct an unbalanced panel of 96 countries over the period 1970 to 2010.[10] Because institutions and equality exhibit limited variation over time, the data are used in five-year periods, similar to Acemoglu, Johnson, Robinson, and Yared (2008). This section gives a short overview over the construction and coding of the main variables. A detailed description of the data and their sources can be found in Table F.1 in the Appendix. Table F.2 in the Appendix provides descriptive statistics of the main variables used in the empirical analysis.[11]

7.3.1 Institutional Quality

The two main measures of (de facto) institutional quality are based on the Economic Freedom of the World Index provided by Gwartney, Lawson, and Hall (2013) and the Freedom House (2014) Civil Liberties indicator. The Economic Freedom index (EF) is composed of 42 distinct variables in five general categories: size of government and taxation; private property and the rule of law; soundness of money; trade regulation and tariffs; regulation of business, labor, and capital markets. The Civil Liberties index (CL) is based on 15 items for four subcategories: freedom of expression and belief; associational and organizational rights; rule of law; and personal autonomy and individual rights.

[10] During earlier periods, developed countries of Europe and North America are oversampled due to data availability reasons. This selection is less severe for the data starting at 1970. A possibility to further circumvent this problem would be to impute the data set and work with a balanced panel. This would require additional (strong) assumptions on the missing data points, however.

[11] Table F.3 in the Appendix reports simple (pairwise Pearson) correlation coefficients for the variables of main interest.

These comprise dimensions such as freedom of residence, employment, and education; the right to own property and establish businesses; personal social freedom including marriage partners and size of family; and equality of opportunity and the absence of exploitation. Both measures are conceptually somewhat broader than only reflecting economic institutions. Nevertheless, Economic Freedom as well as some components of the Civil Liberties index reflect the de facto institutional quality as implied by the conceptual framework in Figure 7.2. The empirical analysis, therefore, uses the Economic Freedom index as a baseline measure of institutional quality. In addition, we construct a composite measure of economic institutional quality by extracting the first principal component from the Economic Freedom and Civil Liberties indices, normalized to range from 0 to 1 (see Table F.2 in the Appendix). This composite index extracts the variation in institutional quality in the economic domain contained in the two variables, and, in particular, it isolates common factors among institutional indicators that can be interpreted as economic institutions.[12] We view this composite index as more suited for the purpose of measuring institutional quality in the sense of the theoretical framework underlying the empirical analysis than the plain Civil Liberties index, which also comprises aspects of political freedom and is sometimes even used as proxy for political institutions. The composite measure of institutional quality provides a more clear-cut distinction of institutional quality from the measures of political institutions. The use of such a composite measure also follows recent suggestions in the literature (see, for example, Voigt, 2013) for the construction of a suitable measure of the institutional quality. Below, we also report the results of extensive robustness checks regarding the measure of institutional quality. In addition, the robustness analysis makes use of information about the Civil Liberties index, property rights protection (Investment Profile), and protection against corruption (Corruption) provided by the International Country Risk Guide (ICRG).[13] These data are available only for the shorter time span from 1985 to 2010. Results are also reported for subcomponents of the Economic Freedom index.[14] All indicators are normalized and adjusted to range from 0 (low quality) to 1 (high quality).

7.3.2 Democracy

To measure political institutions (in terms of democracy), we apply two of the most widely used indicators of (de jure) political institutions in the literature, namely the Constraints on Executive indicator and the composite PolityIV index provided by the

[12] Factor analysis synthesizes the variation contained in several variables into common, orthogonal factors, or principal components. This way, one can decompose the variation in institutional variables that corresponds more closely to economic institutions, from variation that corresponds more closely to political institutions. For more information see Voigt (2013, pp. 20–21).

[13] `http://epub.prsgroup.com/products/international-country-risk-guide-icrg`.

[14] These are, in particular, the regulation of credit, labor and business (regulation), and the soundness of money.

PolityIV data set by Marshall, Jaggers, and Gurr (2013). Constraints on the executive reflect the extent of institutionalized constraints on the decision-making powers of chief executives, whether individuals or collectives.[15] The composite PolityIV index is based on the evaluation of competitiveness of executive recruitment, openness of executive recruitment, constraints on the chief executive, competitiveness of political participation, and regulation of participation.[16] In contrast to the measures of institutional quality above, political institutions indicators are based on comparably more objective criteria and reflect more closely the de jure nature of political institutions. The indicators are rescaled to range from 0 (full autocracy) to 1 (full democracy).

For the robustness exercises we use the Freedom House (2014) Political Rights indicator, the Index of Democracy (as well as its components) constructed by Vanhanen and Lundell (2014), the binary Democracy-Dictatorship indicator of Cheibub, Gandhi, and Vreeland (2010), as well as artificially created indicators based on principal component analysis. As with institutional quality, we obtain alternative measures of political institutions by decomposing the variation contained in various indices using factor analysis. All alternative indicators are also rescaled to range from 0 (full autocracy) to 1 (full democracy).

The categorization of institutional measures in terms of institutional quality and democracy can also be justified in terms of the distinction into de jure and de facto institutional quality, consistent with the conceptual framework of Figure 7.2. In particular, both measures of institutional quality, economic freedom and civil liberties, represent measures of the de facto institutional quality as experienced by individuals, predominantly (or to a large extent) related to the economic domain. The democracy measures, in contrast, more closely reflect a de jure aspect of political institutions. This is also reflected in the relatively more objective criteria on which these measures are based.[17]

7.3.3 (In-)Equality

The degree of economic (in-)equality within a country is measured using the (reversed) income Gini coefficient, $(1 - \text{Gini})$, based on net incomes from the Standardized World Income Inequality Database (SWIID) constructed by Solt (2009, 2016b).[18] We reverse the scale of the Gini coefficients and normalize them so that values range from 0 (perfect inequality) to 1 (perfect equality). This reversion of the scale of the Gini coefficient into

[15]The raw data assign a value from 1 to 7. The empirical analysis uses a rescaled measure that takes values between 0 and 1, with 1 corresponding to the highest constraints, see Table F.1 in the Appendix for details.

[16]The raw composite polity score ranges from –10 to 10. The empirical analysis uses a rescaled measure that takes values between 0 and 1, with 1 corresponding to most democratic institutions, see Table F.1 in the Appendix for details.

[17]See also Table F.1 in the Appendix for details.

[18]The distinctive feature of this data set compared to other inequality data is that the standardization of Gini coefficients across countries allows for cross-country comparisons, while the multiple imputation used in its construction provides a comprehensive data set with global coverage. In contrast to many non-standardized data sets often used in applied work, this data set is ideal for the purposes of this study.

a measure of equality rather than inequality, which might seem counterintuitive at first glance, facilitates the interpretation of the interaction term with political institutions. In order to limit the gaps in the data set, the SWIID uses multiple imputation procedures to recover missing values. For this reason, the data set provides 100 values of (in-)equality for each country-year cell, which can be used for the correction of standard errors to account for multiple imputation. Following the standard in the literature, the empirical analysis in this study uses the simple mean of the reversed Gini.

In the robustness analysis, we provide results based on standard errors that are adjusted to account for the multiple imputation of the inequality measures. In the robustness analysis, we also use alternative measures of (in-)equality, in terms of the gross Gini coefficient by Solt (2009, 2016b), a binary indicator of equality (obtained by a median split), and a human capital Gini coefficient which reflects (in-)equality in the distribution of skills and human capital available for production.[19] To facilitate the interpretation of the results, all indicators are normalized to range from 0 (perfect inequality) to 1 (perfect equality).

7.3.4 Controls

In order to explore problems of omitted variables and reduce the systematic variation captured in the country fixed effect, the analysis is also conducted with specifications of the empirical model with additional controls for log GDP per capita and GDP per capita growth from Penn World Tables 8.1 by Feenstra, Inklaar, and Timmer (2015); average years of schooling by Barro and Lee (2013); log population size by Barro and Lee (2013); ethnic polarization by Reynal-Querol and Montalvo (2005); colonial history by the CEPII data set by Mayer and Zignago (2011); an inflation and a deflation dummy variable constructed by the inflation rate of World Development Indicators by the World Bank (2014); as well as binary indicator variables for oil exports and for former socialist countries.

[19] Following Castelló and Doménech (2002), a human capital Gini coefficient is constructed based on average years of schooling for primary, secondary, and tertiary education using the Barro and Lee (2013) data set. The human capital Gini coefficient is computed as

$$G^h = \frac{1}{2\bar{H}} \sum_{i=0}^{3} \sum_{j=0}^{3} |\hat{x}_i - \hat{x}_j| n_i n_j,$$

where $\bar{H}$ are average years of schooling of the entire population aged 15 and older, x_i and x_j are the cumulative average years of schooling associated with four different educational levels (no formal, primary, secondary, and tertiary education, correspondingly) with their respective population shares n_i and n_j.

7.4 Empirical Approach

7.4.1 Empirical Framework

The empirical analysis is rooted in the conceptual framework illustrated in Figure 7.2. This framework stipulates an interaction between political institutions and inequality in shaping institutional quality in a dynamic setting, that is, when exploiting variation within a country over time. The empirical analysis is therefore based on a standard dynamic linear panel model:[20]

$$\begin{aligned} inst_{i,t} = \; & \alpha \, inst_{i,t-1} + \beta_e \, eq_{i,t-1} + \beta_d \, demo_{i,t-1} + \gamma \, (eq_{i,t-1} \times demo_{i,t-1}) \\ & + z'_{i,t-1}\delta + \eta_t + \eta_i + \epsilon_{i,t}, \end{aligned} \tag{7.1}$$

where $inst_{i,t}$ measures the quality of (economic) institutions for country i in period t; $eq_{i,t-1}$ the indicator for economic equality; $demo_{i,t-1}$ is the lagged democracy index; $z'_{i,t-1}\delta$ is a vector of lagged control variables and their corresponding coefficients; η_t and η_i denote a time effect and a time-invariant country-specific fixed effect; and $\epsilon_{i,t}$ is an idiosyncratic error term. The novel element is the inclusion of the lagged interaction of democracy and equality ($eq_{i,t-1} \times demo_{i,t-1}$). Time periods are five-year intervals from 1970 to 2010. Because the evolution of economic institutions and civil liberties is strongly persistent, the specification also includes a lagged value of the dependent variable.[21] Political participation and equality also vary slowly over time. Therefore, using first lags instead of a model with contemporaneous variables seems appropriate. General time trends as well as country-specific features that are constant over time like culture or ethnic and linguistic fractionalization are captured by the period and country effects. In addition, dummy variables for oil production and former socialist countries are included for models that use level equations in order to reduce the risk of omitted variables and to diminish the variance explained by the country fixed effect and, thus, the risk that instruments are weak in some of the specifications (explained in more detail below). To account for unobserved heterogeneity and slow-moving dynamics in institutional quality, the empirical specification also includes a lagged value of the measure of institutional quality as explanatory variable in the empirical analysis unless stated otherwise. Finally, the specification also accounts for country- and period-specific variation in the measurement of the institutional indicators used in the analysis, such as changes in the questionnaires underlying the indices or discrete shifts in the judgment of institutional quality over time (for example, corruption). This point is of particular importance given the construction of some of the indicators is based

[20] Similar specifications are used in Acemoglu, Johnson, Robinson and Yared (2008, 2014) and Murtin and Wacziarg (2014).

[21] This specification measures the short-run effect of democracy on institutional quality, conditional on the degree of economic equality. The long-run effect can be computed imposing a steady-state condition and solving for *inst*.

on subjective assessments, and occasional changes in the survey methodology from one wave to another. The panel estimation thus also restricts the problem of measurement error.

The main interest of the empirical analysis lies in the interaction of political institutions (democracy) with economic equality, and the corresponding heterogeneity or even non-monotonicity of the effect of democracy and equality on the quality of economic institutions. Correspondingly, the total effect of democracy can be computed from the values of β_d and γ, conditional on the level of equality. In general, the specification allows for the possibility that the effect of democracy is positive for some country-period observations (for example, those with a very equal income distribution) and negative for others (those with high inequality).

7.4.2 Estimation and Identification

Due to the lack of quasi-experimental variation in democracy and equality, five different dynamic panel estimators are applied in order to estimate the coefficients of interest: random effects, fixed effects, bias-corrected fixed effects, differences GMM, and system GMM. Each of these estimators relies on a set of identification assumptions and therefore has advantages and disadvantages. The main contribution of this paper is based on a comparison of the coefficient estimates obtained from these different estimation methods. Even if each specification by itself may suffer from biases due to mechanical endogeneity, overly restrictive exogeneity, or inappropriate stationarity assumptions, the overall pattern of estimates is informative with respect to the bounds of the true coefficient and the potential relevance of problems regarding the identification assumptions. In particular, finding a robust significant interaction effect between democracy and equality on institutional quality throughout all models would provide qualitative evidence for the existence of such an interaction, and the pattern of estimates can be informative regarding the quantitative relevance of the effect in terms of the appropriateness of different sets of identification assumptions. In the following, we provide a brief summary of the pros and cons of each estimator, of the caveats regarding the validity in the present application, and of the interpretation of the estimation results.

Without the lagged dependent variable as explanatory variable, random and fixed effects models share the requirement of strong exogeneity of the error terms; that is, all explanatory variables, including lagged and future realizations of the explanatory variables, must be uncorrelated with current-period error terms. For the random effects model, the regressors also have to be uncorrelated with the unobserved individual fixed effect to obtain unbiased estimates. In the fixed effects model, this problem can be avoided by eliminating the unobserved fixed effect through performing a within-transformation. Correspondingly, a potential omitted variable would have to be correlated with the demeaned regressors to bias the estimates. When time effects are accounted for, an omitted variable would have to be both time- and country-varying, which restricts the set of potential confounds considerably.

Dynamic fixed and random effects panel estimates (obtained with specifications that include the lagged dependent variable) are biased in short panels because of mechanical correlation between the lagged dependent variable and the error term. This bias vanishes as the number of time periods increases (Nickell, 1981). One way to deal with this inherent endogeneity problem is to perform a bias correction, as proposed by Bun and Kiviet (2003) for balanced panels, or by Bruno (2005) for unbalanced panels. The resulting estimator approximates and eliminates the bias that results from the fixed panel length, thereby yielding more credible results than the fixed effects estimator without requiring additional, potentially more restrictive assumptions such as stationarity.[22]

In the context of dynamic panels, the most frequently used solution to the endogeneity problem is the application of estimators based on generalized methods of moments (GMM). The difference GMM estimator identifies the coefficients of interest from changes in the explanatory variables, exploiting the lagged values of the explanatory variables as instruments for the endogenous regressors (Arellano and Bond, 1991). Panel GMM models require weaker assumptions than the strong exogeneity condition needed in the fixed effects case. In particular, GMM only requires lagged instruments to be uncorrelated with the current-period error terms. A well-known problem that can arise when using lagged values as instruments is that these may only be weakly correlated after country fixed effects are removed. This weak instruments problem, which is more severe in persistent time series, implies that the estimates for the autoregressive parameter α are biased toward the fixed effects estimator, and can thus be seen as a lower bound for the true parameter. The bias-corrected fixed effects estimator can serve as a rough guidance in this case.

The system GMM estimator achieves more efficient estimates and provides more stable results for highly persistent variables, that is, for estimates of α close to unity, by including levels as additional instruments (Blundell and Bond, 1998). In order to further reduce the danger of weak instruments in the SGMM specification and the relative variance of the country fixed effects to the idiosyncratic error term, also time-invariant variables can be included in the level regressions.[23] A comparison of results of differences and system GMM also provides further information as to whether instruments are weak. The parameter estimate of SGMM is a weighted sum of the difference and level equation with more weight being put on the differenced model if identification is strong. Therefore, quantitatively

[22] See also Bun and Carree (2005) and Bun and Kiviet (2006).

[23] Bun and Windmeijer (2010) point out that, if the variance of the unobserved individual fixed effect is high compared to the variance of the idiosyncratic error term, there may be severe downward bias in the autoregressive parameter for the SGMM estimator. Adding level controls that absorb some of the variation that is explained by the country fixed effect helps to alleviate this problem. In our setting, the variation explained by fixed effects and the unexplained variation are of approximately similar size.

similar estimates from DGMM and SGMM are an indication for instrument relevance. However, in addition to the identifying assumptions of DGMM, SGMM also demands a stationarity assumption requiring changes in the dependent variable be uncorrelated with the country fixed effect. This condition likely holds only approximately in the present application, where democratic transitions may affect the long-run steady state. The condition is, however, still satisfied as long as countries are sufficiently close to their long-run steady state at the end of the panel period (see also Roodman, 2009, p. 144).

An important aspect of both GMM estimators is the choice of the number of instruments employed in the estimation. The number of moments available for estimation increases efficiency without affecting the consistency of the estimated parameters provided that the identifying assumptions are valid. A high number of instruments can lead to bias in the Hansen J-test whether instruments provide statistically similar results in the case of overidentification. Using lags that have little explanatory power for contemporaneous variables can weaken the set of instruments. Hence, there is a trade-off between the efficiency of results and the strength of instruments as well as the validity of the test statistics for the validation of the identification assumptions. According to Roodman (2009), as a rule of thumb, the number of instruments should not exceed the number of cross-sectional units (here: countries) in the sample.[24] To determine which lags are most informative and prima facie causal, we use Granger (1969, 1980) tests of different lag specifications in combination with the Akaike and Bayesian Information Criteria. While, in general, the choice of the transformation matrix that removes the unobserved individual fixed effects does not matter, if the complete instrument set is used (Arellano and Bover, 1995), the transformation matrix may be of importance in light of the need to limit the number of lags used as instruments. In this respect, the forward orthogonalized deviations method is useful, because it eliminates less data than first differences.[25]

In sum, all estimators have their strengths and caveats. Appendix A in the supplementary material provides a more detailed description and discussion of the identification assumptions of each estimator. In the context of the current application, the bias-corrected fixed effects estimator as well as the difference GMM estimator appear to be the most conservative and reliable estimators.

7.5 Empirical Results

7.5.1 Baseline Results

Table 7.1 presents the findings for the baseline specification (7.1), including the lagged dependent variable as explanatory variable. This specification represents the most con-

[24]Note that this rule of thumb is not conservative such that instrument counts close to the respective threshold are no guarantee that test statistics are not biased.

[25]See Hayakawa (2009) for a comparison of different transformation techniques.

servative approach to the determinants of institutional quality, as it directly accounts for unobservable heterogeneity that is reflected in past institutional quality. Columns (1) and (2) report results from random effects, Columns (3) and (4) present fixed effects estimates, Columns (5) and (6) bias-corrected fixed effects, Columns (7) and (8) difference GMM, and Columns (9) and (10) contain system GMM estimates. In addition to country and time effects, the empirical specifications include log per capita income and human capital in terms of average years of schooling as controls. The level equations of random effects and system GMM additionally include dummies for former socialism and oil production. For every specification, we report the count of observations, countries, instruments as well as the p-values of the AR(2) and the Hansen J-test for the GMM models below the coefficient estimates of interest. Moreover, we report the p-values of the difference-in-Hansen test for additional level instruments for system GMM models. To determine the lags used as instruments in the GMM estimates, we conducted Granger (1969, 1980) tests of different lag specifications in combination with the Akaike and Bayesian Information Criteria.[26]

Panel (a) of Table 7.1 reports the results for the Economic Freedom measure as the dependent variable. The measure of political institutions (denoted as democracy) is the index of Constraints on the Executive; equality is measured by $(1 - Gini)$, where $Gini$ is the Gini index based on net incomes. The estimation results reveal moderate persistence in institutional quality, as expressed by highly significant and positive coefficient estimates of α between 0.47 and 0.74 across all specifications. Only for the corrected fixed effects estimates, the coefficient estimate is higher (around 0.83). The comparably low parameter estimates obtained with the fixed effects estimator might be the result of the Nickell-bias discussed above and can be seen as a lower bound of the estimate.

In the specifications without interaction term, the parameter $\widehat{\beta}_e$ of economic equality, which is measured by the normalized reversed net Gini coefficient, is not significantly different from zero, which might suggest that equality does not have a significant direct effect on institutional quality. The point estimates are typically positive. The estimated effect of democracy $\widehat{\beta}_d$ for the models without interaction term is positive and significant at the five-percent level, with the exception of the random effects estimates in Column (1) and the system GMM estimates in Column (9). In line with some results in the literature, these findings support the hypothesis that democracy alleviates the creation of high-quality

[26]The first lag of the dependent variable has to be instrumented, because it is mechanically correlated with the error term in short panels, as discussed above. The level lags y_{t-2} and y_{t-3} are thus chosen as instruments for the difference $(y_{t-1} - y_{t-2})$ in Differences GMM. Instrumenting explanatory variables can improve efficiency in case of predetermined variables. Hence, the first three lags $x_{t-1}, x_{t-2}, x_{t-3}$ are also used as instruments for differences GMM. For system GMM, y_{t-2} and x_{t-1}, x_{t-2} are used as instruments in the differences equation; further lags are not used to limit the instrument count, because a large number of may lead to severe bias in the Hansen J-test of overidentifying restrictions and weaken the instrument set. The level equation of system GMM uses the first lagged difference of the dependent variable and the explanatory variables as instruments so that the same number of time periods is exploited in differenced and level equations. Robustness results with respect to different lags specifications and the choice of the transformation matrix to eliminate the country fixed effects are reported in the Supplementary Material.

Table 7.1: Effect of Democracy and Equality on Institutional Quality

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Economic Freedom									
L.Inst. Quality	0.71***	0.71***	0.48***	0.47***	0.83***	0.81***	0.69***	0.65***	0.74***	0.68***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.09)	(0.09)	(0.06)	(0.05)	(0.06)	(0.04)
L.Equality	0.03	-0.05	0.03	-0.14	-0.01	-0.15**	0.03	-0.16	0.00	-0.21**
	(0.02)	(0.08)	(0.06)	(0.10)	(0.06)	(0.07)	(0.07)	(0.11)	(0.06)	(0.09)
L.Democracy	0.02	-0.05	0.04***	-0.15**	0.04***	-0.13**	0.04**	-0.14	0.02	-0.16**
	(0.01)	(0.06)	(0.02)	(0.07)	(0.01)	(0.06)	(0.02)	(0.08)	(0.02)	(0.08)
L.(Eq×Demo)		0.12		0.35***		0.31***		0.32**		0.35***
		(0.10)		(0.12)		(0.11)		(0.14)		(0.13)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
AR(2) p–value							0.80	0.71	0.86	0.66
Hansen p–value							0.15	0.39	0.09	0.24
Diff.–in–Hansen p–value									0.70	0.85
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543
	(b) Dependent Variable: Principal Component of Economic Freedom and Civil Liberties									
L.Inst. Quality	0.78***	0.77***	0.56***	0.54***	0.89***	0.83***	0.75***	0.69***	0.78***	0.71***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)	(0.05)
L.Equality	0.08**	-0.06	0.00	-0.20*	-0.03	-0.16	0.08	-0.16	0.12*	-0.12
	(0.04)	(0.08)	(0.09)	(0.12)	(0.08)	(0.10)	(0.10)	(0.15)	(0.07)	(0.12)
L.Democracy	0.02	-0.10*	0.02	-0.20***	-0.02	-0.17**	-0.01	-0.18*	-0.01	-0.22***
	(0.01)	(0.05)	(0.02)	(0.07)	(0.02)	(0.09)	(0.03)	(0.10)	(0.02)	(0.09)
L.(Eq×Demo)		0.20**		0.42***		0.30*		0.33*		0.39**
		(0.09)		(0.13)		(0.15)		(0.19)		(0.15)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
AR(2) p–value							0.16	0.14	0.13	0.10
Hansen p–value							0.02	0.13	0.07	0.20
Diff.–in–Hansen p–value									0.87	0.73
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of the bias-corrected fixed effects estimator is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first two lags of the explanatory variables. In the level equation, the lagged difference of the regressors is used so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

economic institutions. Overall, the direct effect seems to be small and not overly robust.

Inserting an interaction between equality and political institutions, however, delivers a positive and (with the exception of the random effects estimates) significant coefficient estimate of the interaction term $\widehat{\gamma}$. The point estimate varies between 0.31 and 0.35 and is quantitatively very similar throughout the different specifications.

Panel (b) of Table 7.1 reports the corresponding results for the composite index of institutional quality as the dependent variable. The results are qualitatively and quantitatively very similar. The persistence in institutional quality, as indicated by the coefficient estimate for the lagged dependent variable, is quantitatively almost identical to that obtained for Economic Freedom in Panel (a). Also the positive and significant effect of the interaction between equality and democracy, as well as the negative main effects for democracy and equality, are found consistently throughout all specifications.

Taken together, the empirical results provide evidence that is in line with the conjecture of an interaction, regardless of the estimation method and the measure of institutional quality. The finding of this significant interaction constitutes the main contribution of this study and suggests that democracy and equality complement each other in the ability of a country to establish high-quality economic institutions. In line with the hypothesis of several models in the literature discussed in Section 7.2, these findings suggest that the positive effect of democracy on institutional quality is reinforced by an equal distribution of incomes. In contrast, the effect of democracy is eroded, if inequality becomes too high. This is indicated by the estimated coefficients of democracy $\widehat{\beta}_d$, which turns negative and even significant in most specifications. This result, which seemingly contradicts the conventional wisdom and most previous empirical findings in the literature, reflects the effect of democracy in a society with an entirely unequal distribution of income (a reversed Gini coefficient of 0), and is moderated by the significant positive interaction term $\widehat{\gamma}$. A similar finding applies to equality, which also displays a negative, and sometimes significant, coefficient $\widehat{\beta}_e$ for the direct effect. Given the empirical specification, this reflects the effect of an increase in equality under entirely undemocratic political institutions.

Table 7.2 presents the results for the baseline specification (7.1), under the constraint that the coefficient of the lagged dependent variable, α is set to zero. This static specification corresponds to the typical specification of empirical models found in the existing literature. Again, Panel (a) presents the results with the Economic Freedom index as measure of institutional quality, Panel (b) presents the results for the composite index of institutional quality; democracy is measured by the index of Constraints on the Executive, and equality is measured by the reverse Gini index. For better comparability, the structure of the Table is the same as in Table 7.1. Columns (1) and (2) report random effects estimates, Columns (3) and (4) are fixed effects estimates, Columns (5) and (6) are DGMM, and Columns (7) and (8) contain SGMM estimates.[27] Again, all empirical models include country and time effects, as well as log per capita income and human capital in terms of average years of schooling as additional controls. As in the baseline specification, the level equations of random effects and system GMM additionally include binary indicators for former socialism and oil production.[28]

[27]Due to the exclusion of the lagged dependent variable from the empirical model, the corrected fixed effects estimator, reported in Columns (5) and (6) of Table 7.1 is redundant in this Table.

[28]Although the specification does not include a dynamic component, the GMM estimators provide useful information as they use additional moment conditions for identification. As before, the count of observations, countries, instruments as well as the p-values of the AR(2), the Hansen J-test for both GMM models, and the difference-in-Hansen test for additional level lags for SGMM are reported below the coefficient estimates of interest. The slightly increased number of observations is due to countries for

Table 7.2: Static Model: Effect of Democracy and Equality on Institutional Quality

	Random Effects		Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(a) Dependent Variable: Economic Freedom							
L.Equality	0.10*	-0.11	0.07	-0.13	0.03	-0.21*	0.04	-0.19
	(0.05)	(0.10)	(0.08)	(0.11)	(0.07)	(0.11)	(0.06)	(0.13)
L.Democracy	0.05**	-0.15**	0.05**	-0.18**	0.04*	-0.18**	0.03	-0.23***
	(0.02)	(0.07)	(0.02)	(0.08)	(0.02)	(0.09)	(0.02)	(0.08)
L.(Eq×Demo)		0.35***		0.41***		0.40***		0.49***
		(0.13)		(0.14)		(0.15)		(0.14)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓					✓	✓
AR(2) p–value					0.31	0.22	0.39	0.25
Hansen p–value					0.12	0.21	0.32	0.39
Diff.–in–Hansen p–value							0.54	0.96
Instruments					49	69	54	75
Groups	96	96	96	96	96	96	96	96
Observations	553	553	553	553	457	457	553	553
	(b) Dependent Variable: Principal Component of Economic Freedom and Civil Liberties							
L.Equality	0.18**	-0.29**	0.06	-0.31**	-0.02	-0.38**	0.19*	-0.21*
	(0.09)	(0.13)	(0.11)	(0.13)	(0.13)	(0.17)	(0.10)	(0.11)
L.Democracy	0.16***	-0.28***	0.13***	-0.28***	0.07*	-0.31***	0.09**	-0.38***
	(0.03)	(0.08)	(0.03)	(0.09)	(0.04)	(0.08)	(0.04)	(0.10)
L.(Eq×Demo)		0.79***		0.76***		0.75***		0.87***
		(0.15)		(0.16)		(0.17)		(0.19)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓					✓	✓
AR(2) p–value					0.31	0.10	0.17	0.06
Hansen p–value					0.08	0.09	0.08	0.15
Diff.–in–Hansen p–value							0.15	0.50
Instruments					49	69	54	75
Groups	96	96	96	96	96	96	96	96
Observations	553	553	553	553	457	457	553	553

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by $(1 - \text{Net Gini})$. Control variables are log GDP p.c., average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first two lags of the explanatory variables. In the level equation, the lagged difference of the regressors is used so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

In the specifications without interaction term, democracy exhibits a positive effect $\widehat{\beta}_d$, which is significant in most cases, while equality has a positive but typically insignificant (and in some cases in Panel (b) even negative) effect on institutional quality. These results resemble the findings in the previous literature, where democracy has been associated with better institutional quality, and where the evidence for equality is more scarce.[29]

Most importantly, however, the estimates obtained with a specification without a lagged

which data for explanatory variables are available during the first period but no data for institutional quality. Details are available upon request.

[29]The findings by Chong and Gradstein (2007b) for the years 1960–2000 are a notable exception. Possible explanations for the slight difference in the estimates are the composition of the data set as well as problems with the cross-country comparability of earlier inequality indicators, such as the one by Deininger and Squire (1996), as was pointed out in Section 7.3.

dependent variable also reveal a positive and significant coefficient $\widehat{\gamma}$ for the interaction term between equality and democracy, regardless of the estimator employed. The effect is qualitatively very similar, and quantitatively even slightly larger, when compared to the estimates obtained in specifications that include a lagged dependent variable.

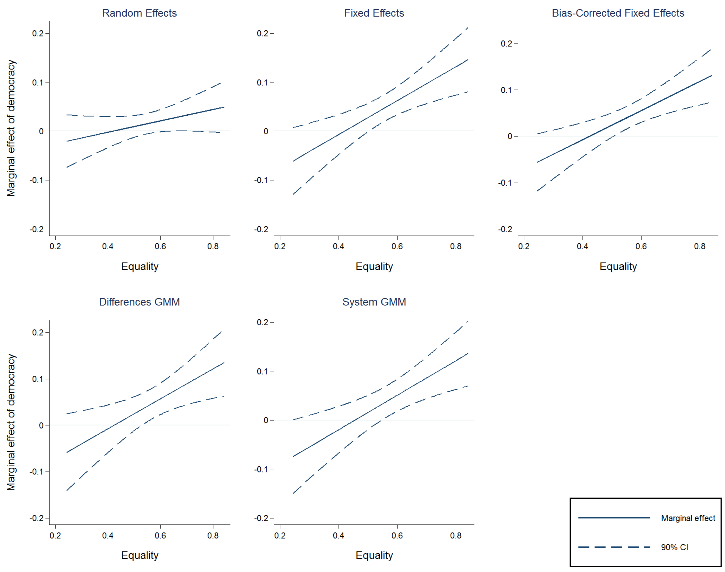
The results from both sets of estimates suggest that democratic political institutions can even have a non-monotonic effect on institutional quality, the sign of which depends on the degree of economic equality. For countries with an equal distribution of incomes, the overall effect of a marginal increase in democratic quality on institutional quality is positive, while for highly unequal societies the total effect is negative. Likewise, an increase in equality appears to have a detrimental effect on institutional quality in non-democratic countries, whereas the opposite is true in democracies. Given the estimates for political institutions, income equality, and the corresponding interaction term, the marginal effects (ME) of democracy and equality are given by

$$ME_{i,t}^{demo} = \widehat{\beta}_d + \widehat{\gamma} \times eq_{i,t-1} \quad \text{and} \quad ME_{i,t}^{eq} = \widehat{\beta}_e + \widehat{\gamma} \times demo_{i,t-1} \,. \tag{7.2}$$

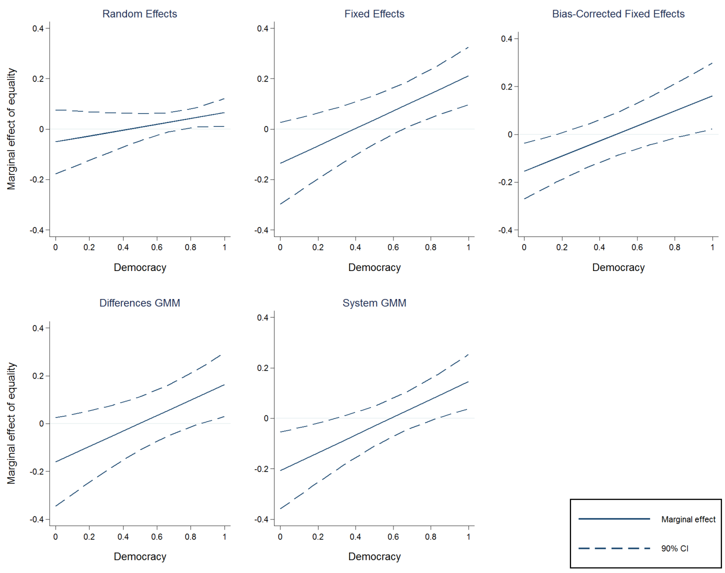
Using this expression, one can calculate a threshold of equality that indicates whether a marginal change in democracy has a positive or negative effect on economic institutions. Taking the estimated parameters from bias-corrected fixed effects and GMM in Columns (6), (8), and (10) in Table 7.1, the corresponding threshold values are approximately 0.45 for Economic Freedom and 0.55 for the composite index of institutional quality. For a reversed Gini coefficient that is higher than this respective threshold, that is, for sufficiently high equality, the effect of a marginal change in the democracy score is positive. This is the case for about 92 percent and 70 percent of the observations in the sample, respectively. On the other hand, the effect of a marginal change in democracy on institutional quality is negative for about 8 to 30 percent of the observations in the sample. Similarly, the threshold of democracy that indicates whether a marginal increase in equality has a positive or negative effect is 0.55 for Economic Freedom and 0.48 for the composite index, respectively. Consequently, an increase in equality has a positive effect on the respective institutional quality variable when the democracy index is above this threshold. This is the case for 65 and 67 percent of the sample. The effect is therefore negative for the remaining countries.

Figure 7.3(a) plots the marginal effect of democracy conditional on the level of equality for the different estimates reported in Table 7.1. Figure 7.3(b) presents the corresponding marginal effect of equality conditional on the level of democracy.[30] The solid lines depict the marginal effect of democracy on institutional quality given by equation 7.2. The

[30] Figure A.1 in the Supplementary Material provides the corresponding graph for the results for the composite index of institutional quality. The heterogeneity is even more pronounced when plotting the corresponding graphs based on the estimation results in Table 7.2. The figures are available on request.



(a) Democracy and Institutional Quality: ME^{demo}

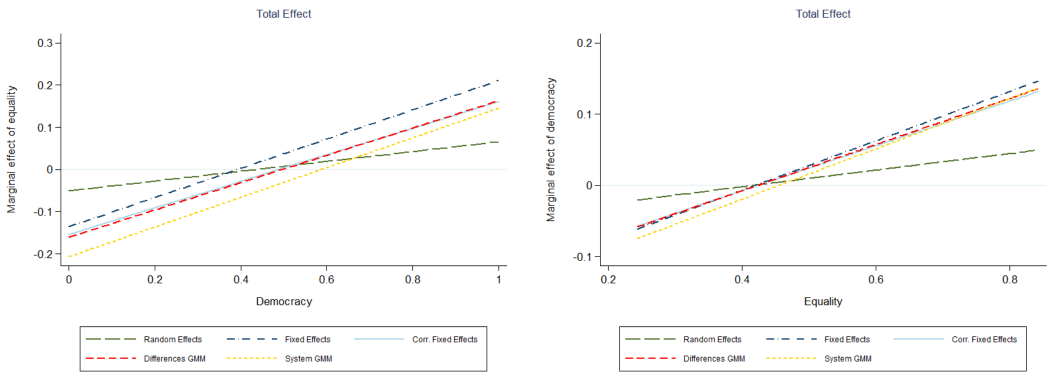


(b) Equality and Institutional Quality: ME^{eq}

Figure 7.3: Marginal Effects of Estimates in Table 7.1(a)

dashed lines represent the +/− five-percent confidence intervals of the composite effect.[31] The graphs show that democracy has a heterogeneous and even non-monotonic effect, which strongly depends on the level of equality. The effect is largest in absolute value for extreme degrees of (in-)equality and zero for intermediate levels. For low equality, the effect of democracy is even negative, suggesting that inequality may erode the possibility of

[31] For a large sample, the confidence interval of the composite effect is given by $ME \pm z_{1-\tau/2} \times \widehat{SE}$ with $\widehat{SE} = \sqrt{[Var(\widehat{\beta}_d) + eq_{t-1}^2 \times Var(\widehat{\gamma}) + 2 \times eq_{t-1} \times Cov(\widehat{\beta}_d, \widehat{\gamma})]}$ and $z_{1-\tau/2}$ being the critical value of a two-sided t-test of size τ.



(a) Equality and Institutional Quality (b) Democracy and Institutional Quality

Figure 7.4: Marginal Effects Across Estimators, Table 7.1(a)

democracies to implement high institutional quality. Likewise, the effect of an increase in equality is highest in democracies, and lower, potentially even negative, in non-democracies.

Finally, Figure 7.4 illustrates the composite effect of democracy and equality on institutional quality for the five different estimators in Table 7.1.[32] The effect of democracy is heterogeneous and non-monotonic throughout all estimators. In addition, all estimators yield comparable estimates for the total effect, except for random effects which shows a somewhat flatter slope. A similar comment applies to the effect of equality. The finding of a significant interaction between democracy and equality, which implies a non-monotonic effect of political institutions on institutional quality, thus appears not to be driven by the specific identification assumptions for a particular estimator. The finding that the total effect is quantitatively similar across specifications provides additional support for the hypothesis that there is substantial heterogeneity and even non-monotonicity in the effect of democracy on institutional quality.

7.5.2 Robustness

The baseline results in Tables 7.1 and 7.2 indicate an interaction between democracy, measured by Constraints on the Executive, and equality, regardless of whether institutional quality is measured by the Economic Freedom index or the composite index of institutional quality. One potential concern with the results is multicollinearity, if the quality of political institutions is strongly correlated with the extent of income equality. In addition, this problem might be more severe the less variation over time the respective variables exhibit. The correlations, however, are rather moderate, indicating that this is not likely to affect the results.[33] In the data, which are measured in five-year intervals, the pairwise correlation of the Constraints on the Executive and the reversed Gini coefficient is about

[32]The corresponding plots for the estimates in Table 7.2 are very similar and available upon request.
[33]See Table F.3 in the Appendix.

Table 7.3: Robustness: Economic Institutions (CFE)

	Civil Liberties	Regulation	Soundness of Money	Investment Profile	Protection vs. Corruption
	(1)	(2)	(3)	(4)	(5)
L.Inst. Quality	0.73***	0.79***	0.61***	0.32***	0.46***
	(0.07)	(0.09)	(0.06)	(0.08)	(0.08)
L.Equality	-0.27*	-0.14	-0.33*	-0.56**	0.21
	(0.16)	(0.10)	(0.19)	(0.26)	(0.24)
L.Democracy	-0.30**	-0.17**	-0.38**	-0.36*	-0.08
	(0.14)	(0.08)	(0.15)	(0.21)	(0.19)
L.(Eq×Demo)	0.48**	0.29**	0.78***	0.72**	0.17
	(0.24)	(0.14)	(0.26)	(0.35)	(0.32)
Controls	✓	✓	✓	✓	✓
Groups	112	96	96	90	90
Observations	610	539	557	407	407

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. Results are computed using the bias-corrected fixed effects estimator by Bruno (2005). Dependent variables in the respective columns are the principal component of Economic Freedom and Civil Liberties; Civil Liberties; Regulation of Credit, Labor and Business; Soundness of Money; Investment Profile (Property Rights); and Protection against Corruption. Democracy is measured by the Constraints on Executive indicator. Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

0.37 unconditionally, and 0.12 when conditioning on country-fixed and time effects. In addition, simple diagnostic tools such as the variance inflation factor suggest that the problem of multicollinearity does not pose a threat to the findings.

The remainder of this section reports the results of several robustness checks that explore the sensitivity of the results. A first set of robustness checks replicates the analysis using several alternative measures of institutional quality. These alternative measures include the Civil Liberties index; an index of the regulation of credit, labor, and business, and an index of the soundness of money as a measure of reliability of the economic environment (both of these measures are taken from subindices underlying the Economic Freedom index); an index of the quality of property rights, and finally an index of the protection against corruption (both taken from the ICRG data set). These measures capture different dimensions of institutional quality. For instance, the index of Civil Liberties is more closely linked to political liberties than the Economic Freedom index. Moreover, some of these measures are only available for a substantially shorter time span (1985–2010), which limits the variation over time. Nevertheless, the empirical estimates obtained with these institutional measures as the dependent variable also deliver evidence for a positive interaction between democracy and equality, which is significant in all but one specification, as shown in Table 7.3.[34]

A second set of robustness checks replicates the analysis for several alternative measures of political institutions. In particular, in analogy to the construction of the institutional quality index, we constructed a composite measure based on a principal component analysis of political institutions. This measure is based on the index of Constraints on the

[34]The estimates are based on the bias-corrected fixed effects estimator. Additional results for the other estimators are contained in the Supplementary Material.

Table 7.4: Robustness: Political Institutions (CFE)

	Principal Comp. XC & VHC (1)	PolityIV Index (2)	Political Rights (3)	Principal Comp. PIV & PR & VH (4)	Democracy-Dictatorship (5)
L.Inst. Quality	0.80***	0.79***	0.79***	0.80***	0.83***
	(0.09)	(0.10)	(0.09)	(0.09)	(0.07)
L.Equality	-0.15**	-0.10	-0.20**	-0.13	-0.09
	(0.07)	(0.08)	(0.08)	(0.08)	(0.07)
L.Democracy	-0.14**	-0.05	-0.18***	-0.12	-0.10**
	(0.06)	(0.06)	(0.07)	(0.08)	(0.04)
L.(Eq×Demo)	0.35***	0.20*	0.40***	0.34**	0.22***
	(0.11)	(0.11)	(0.11)	(0.14)	(0.07)
Controls	✓	✓	✓	✓	✓
Groups	96	96	96	96	96
Observations	543	550	550	550	546

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. Results are computed using the bias-corrected fixed effects estimator by Bruno (2005). The dependent variable is Economic Freedom. The democracy proxies in the respective columns are the principal component of Constraints on Executive and the Vanhanen Competition indicator; the combined PolityIV indicator; Political Rights; the principal component of the combined PolityIV indicator, Political Rights and the composite Vanhanen Democracy indicator; and the dichotomous Democracy-Dictatorship indicator. Equality is proxied by $(1 - \text{Net Gini})$. Control variables are log GDP p.c. and average years of schooling. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Executive and the political competition component of the democracy index by Vanhanen and Lundell (2014)—see, for example, Voigt, 2013, for a discussion of the advantages of such a composite measure. Alternatively, we employ the PolityIV composite index (for democracy and autocracy); the Political Rights index constructed by Freedom House; a composite measure of political institutions based on the principal component of the index constructed by Vanhanen and Lundell (2014), the Polity IV index and the Political Rights index; and finally a binary indicator of democracy vs. dictatorship constructed by Cheibub, Gandhi, and Vreeland (2010), as advocated by some scholars (for example, Persson and Tabellini, 2006, and Papaioannou and Siourounis, 2008). The results, which are shown in Table 7.4, are qualitatively and quantitatively similar to the baseline results.[35] Overall, these results also indicate that the finding of a non-monotonic effect of democracy in relation to equality is not sensitive to the particular democracy indicator that is used in the estimation.

A third set of robustness checks refers to the measure of economic (in-)equality. The SWIID data reflect the state of the art when it comes to comparative panel data on inequality. Nevertheless, recent criticism regarding data comparability, quality, and the imputation model used for their construction has led some to question their suitability for cross-country comparative studies (see Jenkins, 2015). This criticism has been convincingly refuted by Solt (2015), who concludes that "[t]hose pursuing research on income inequality across many countries and over time [...] will often find that the SWIID is their best choice of data source" (Solt, 2015, p. 690). Nevertheless, some features of the inequality measure

[35]Again, the estimates are based on the bias-corrected fixed effects estimator. Additional results for the other estimators as well for other measures that have been suggested in the literature are contained in the Supplementary Material.

Table 7.5: Robustness: Economic (In-)Equality (CFE)

	Multiple Imputation (1)	Gross Gini (2)	Binary Indicator (3)	Human Cap. Gini (4)	Lowest Quintile (5)	2nd–4th Quintile (6)	Highest Quintile (7)
L.Inst. Quality	0.81***	0.81***	0.80***	0.83***	1.23**	0.82***	1.14
	(0.08)	(0.09)	(0.09)	(0.08)	(0.53)	(0.17)	(1.28)
L.Equality	-0.14*	-0.09	-0.03**	0.08			
	(0.07)	(0.07)	(0.01)	(0.07)			
L.Democracy	-0.12*	-0.09	0.02	-0.01	-0.02	0.07**	0.10
	(0.06)	(0.06)	(0.02)	(0.03)	(0.10)	(0.03)	(0.17)
L.(Eq×Demo)	0.29**	0.24**	0.07***	0.08			
	(0.10)	(0.10)	(0.03)	(0.05)			
Controls	✓	✓	✓	✓	✓	✓	✓
Groups	96	96	96	96	38	80	29
Observations	553	543	543	639	126	324	114

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. Results are computed using the bias-corrected fixed effects estimator by Bruno (2005). The dependent variable is Economic Freedom. Democracy is measured by the Constraints on Executive indicator. Equality measures in the respective columns are (1 − Net Gini); (1 − Gross Gini); a binary equality indicator which takes a value of 1 if (1 − Net Gini) is above the median of the respective distribution for a given year and 0 else; (1 − Human Capital Gini Gini). Columns (5)–(7) split the sample with respect to quintiles of the distribution of (1 − Net Gini): results are shown for the lowest quintile in Column (5), the second to fourth quintile in Column (6), and the highest quintile in Column (7). Control variables are log GDP p.c. and average years of schooling. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

lend themselves to additional robustness checks. The first check explicitly accounts for the concern that the results are affected by the multiple imputation techniques that were used by Solt (2009, 2016b) in order to reduce the number of missing values in the SWIID data set. In particular, we use the multiple observations for each country-year cell provided in the SWIID data to estimate the coefficients multiple times (for every single Gini realization) and then take the average over all estimated coefficients to compute the respective coefficients and standard errors.[36] Second, instead of measuring economic inequality using net incomes as in the baseline, we use a measure of equality based on market incomes. Arguably, this might be a better proxy for the de facto political power, and, hence, the relevant determinants of economic institutions. Third, to alleviate the concern that the results are driven by measurement error due to the interpolation procedure, we employ a binary measure of inequality. Because economic inequality is also reflected by the distribution of skills, the fourth alternative measure of equality is provided by 1 – the human capital Gini coefficient described in Section 7.3.3. Consistent with the earlier findings, the results reveal a positive interaction effect between equality

[36] This procedure reflects the underlying uncertainty in the equality measure that has been introduced by the imputation procedure. Intuitively, instead of using the mean over all 100 realizations in the estimation, this procedure estimates each model 100 times and takes the mean over the estimated parameters thereafter. The robustness check uses all available 100 Gini realizations, even though a smaller number is typically already sufficient. For a more technical discussion see Rubin (1996) and the Supplementary Material. To the knowledge of the authors only few existing papers account for multiple imputation or interpolation of inequality data. To the extent that the effect for many variables might be overstated without accounting for the imputation noise, the analysis in this study also provides a contribution in this respect by applying more extensive corrections for standard errors than what is usually done in the literature. At the same time, the analysis abstracts from the problem of sample selection that also arises from imputation procedures, if the pattern of missing observations is not random and not adequately modeled in the imputation procedure, see Cameron and Trivedi (2005).

and democracy, which is significant in all specifications except for the human capital Gini, see Table 7.5.[37] Finally, we estimate the empirical model separately on subsamples that were split by equality quintiles (into the lowest quintile, the three intermediate quintiles, and the highest quintile) without an interaction term. Here, the effect of democracy on institutional quality is negative (although not significant) in the subsample in which equality is lowest, whereas the effect becomes increasingly positive in the subsamples reflecting intermediate inequality, and is most positive in the subsample where equality is highest.[38] Overall, the results deliver a coherent pattern of heterogeneity in the effect of democracy on institutional quality that is related to equality.

Additional results reported in the Appendix confirm the robustness of the results for different estimators, different specifications with additional or fewer control variables, lag specifications and transformations in the context of the GMM estimators, and additional interactions of democracy with variables like income per capita or years of schooling in order to account for other factors that have been identified in the previous literature.[39] The findings from these robustness checks suggest that democracy has a robust heterogeneous effect on institutional quality conditional on the degree of economic equality.

Additional results also provide some tentative evidence for the mechanisms underlying the interaction term. One set of results points to a negative interaction effect between democracy and equality in an estimation framework with redistribution as the dependent variable.[40] The arguments mentioned in the Introduction regarding the question of stability of democracy also imply that democracies might become unstable, if inequality becomes too large. Estimations with the level of political institutions as dependent variable using the baseline sample 1970–2010 in five-year intervals, as well as a sample over the longer period from 1870–2010 in ten-year intervals, deliver a positive interaction effect of democracy and equality on political stability, particularly over the longer horizon. Democratic institutions are thus self-reinforcing, but this effect is stronger the greater the level of equality in society. This evidence is suggestive, or at least not inconsistent, with the theories underlying the main hypothesis of this paper. In an attempt to investigate the possibility of heterogeneity

[37] The coefficient is smaller in size and estimated with less precision, which might indicate that this measure involves more measurement error. Interestingly, however, the findings are consistent with recent evidence by Castelló-Climent and Doménech (2014), who find divergent trends in income inequality and education inequality.

[38] Additional results for the other estimators as well for additional alternative measures, including the top-10-percent income share by Piketty (2014), are contained in the Supplementary Material.

[39] These estimates account for the view that a multitude of factors might be relevant for institutions to work successfully. For example, motivated by the arguments forwarded by Lipset (1959), Acemoglu, Johnson, Robinson and Yared (2008, 2009) investigate the importance of income for political institutions, while Murtin and Wacziarg (2014) and Fortunato and Panizza (2015) identify human capital as a central determinant of institutional quality.

[40] This analysis is limited by data availability on redistribution, however. Additional unreported results using a relative political extraction indicator for industrial countries by Hendrix (2010) and Arbetman-Rabinowitz et al. (2013) as measure for taxation and redistribution reveal a positive interaction effect, but the coefficient is estimated with insufficient precision to deliver statistically significant results.

in the interaction effect, we also conducted the estimation separately for countries that democratized before and after 1974, following the classification by Huntington (1993). While the interaction appears positive throughout, the coefficient estimate is indeed somewhat larger and more significant for the sample that democratized after 1974 (the "third wave" of democratization). This suggests that the interaction between a democratic political regime and an equal distribution of income had a particularly large impact on institutional quality in countries in which the institution building process was on the way or not yet finished.[41]

7.6 Conclusion

This study has investigated the role of economic equality in moderating the effect of democracy on institutional quality. Based on the arguments and predictions from an entire strand of the literature, the hypothesis motivating the analysis was that the quality of economic institutions is affected by an interaction between political institutions and equality. Based on a panel of 96 countries over the period 1970–2010, dynamic panel estimates deliver evidence for a non-monotonic effect of democracy on institutional quality that is moderated by inequality. The results are robust and quantitatively similar across different specifications of the dynamic panel model. The results are therefore not driven by endogeneity issues that are specific to certain models. The results are also robust to different empirical specifications, choices of explanatory variables, controls, dependent variables, and standard error corrections for multiply imputed data. Furthermore, the interaction of equality and political institutions that measures the degree of heterogeneity is positive and also significant in the presence of interactions between democracy and income, as well as between democracy and human capital.

Taken together, the findings suggest that equality is a pivotal factor, which determines whether democratic institutions have a positive and lasting effect on institutional quality. In particular, the results are consistent with a negative effect of democracy on institutional quality in very unequal societies, thus providing supportive evidence for the existence of two equilibria with efficient institutions, one characterized by a very equal distribution of resources and democracy, and another characterized by high inequality and autocracy. These findings support the argument by Lipset (1959) that democracies only work and provide good institutions under specific preconditions. Likewise, the findings support arguments that raise concerns about the consequences of increasing inequality (see, for example, Piketty, 2014).

This paper also provides a first step toward opening the black box of institutional quality and its determinants, and of the hierarchy and interdependencies of different concepts of institutions. Thereby, the results add to an ongoing debate about what is

[41] Detailed results are contained in the Supplementary Material.

meant by institutions and how institutions should be measured. While there is a common perception that different measures capture similar underlying institutional features, our data as well as our analysis documents that there is substantial variation across measures, and that the correlation is lower than commonly thought. By opening this black box and testing the prediction of a heterogeneous effect of democracy depending on the level of equality using different measures of political and economic institutions, our analysis essentially follows the suggestion by Voigt (2013) and investigates the variation across different dimensions of institutions. Thereby, the results shed new light on the question regarding what these different measures actually stand for and how this affects empirical results. To address this question, the analysis exploited within-country variation in a dynamic panel setting, and used various synthetic constructs that extract variation in institutional quality from different measures typically used in the literature. Overall, the results are supportive of a framework that views inequality and democracy as the key state variables of the institutional quality that determines individual economic behavior.

At the same time, the interpretation of the results raises additional implications for future research. Although the data suggest that high institutional quality can be achieved both under democracy and autocracy, the highest scores of civil liberty are observed for countries with the highest democracy scores. Consequently, the equilibrium with democratic rights and low economic inequality appears preferable to the autocratic regime in terms of institutional quality. Moreover, even though economic institutions of high quality can emerge in regimes with autocratic political institutions and high inequality, this does not imply that autocratic regimes are able or even willing to implement such institutions. In fact, many autocracies perform very poorly in terms of institutional quality; that is, the lowest scores of the Economic Freedom index are realized in purely autocratic countries. Hence, the results indicate that, while democracies appear neither necessary nor sufficient for high institutional quality, the corresponding corollary for states with limited franchise implies that autocratic regimes are neither necessary nor sufficient for low institutional quality. In our view, more research is needed on uncovering the precise channels underlying the empirical findings presented in this paper.

Chapter 8

Concluding Remarks

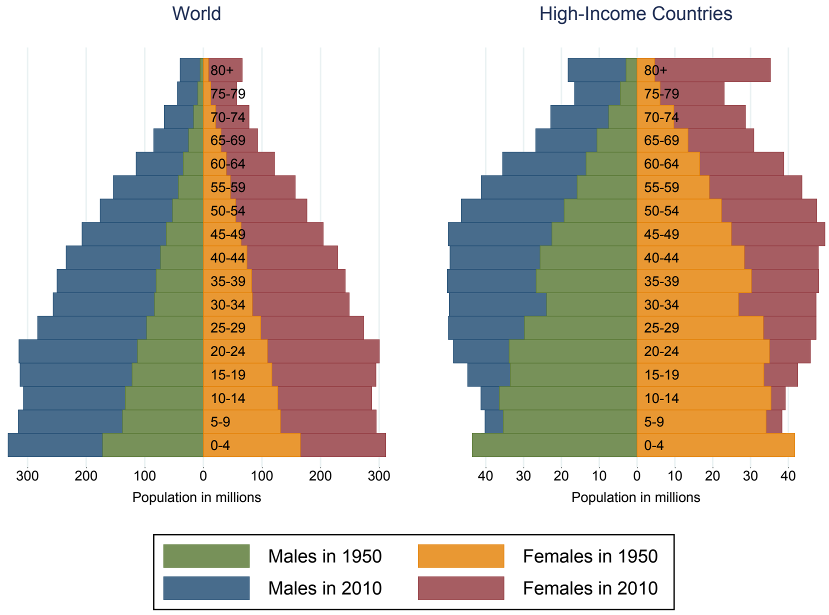
Working on this dissertation, I investigated long-run determinants and processes of socioeconomic development from two different perspectives: population economics and political economy. During this process, my interest for causes and consequences of population aging and demographic change on macroeconomic performance steadily increased. In particular, while researchers and informed laypeople generally agree about the important implications of aging processes on socioeconomic development, our scientific understanding of these processes seems surprisingly limited. For example, the rule of changing preferences over the life-cycle received attention only relatively recently (see, for example, Sunde and Dohmen, 2016). Moreover, changes and disparities in life-cycle productivity will likely play a significant role in the context of aging societies and skill-biased technical change. A sound understanding of these transformations requires knowledge about the role of individual health, educational attainment and occupational choice for worker productivity. Finally, population aging may also have major implications for social cohesiveness of societies through health and income inequality (OECD, 2017). Therefore, I am convinced that further research on the role of population aging and demographic change on socioeconomic development will yield valuable insights for our society.

In this setting, I also expect institutions to play a prominent role in shaping and moderating consequences of socioeconomic transformations. For example, differentials in longevity between different societal groups may unfold (unintended) redistributive effects through the pension system (see, for example, Sanchez-Romero and Prskawetz, 2017). Furthermore, population aging will shift up the age of the median voter in democracies. Correspondingly, policies will more likely target the needs of the elderly than of the young. Finally, technological progress and subsequent transformations of labor markets require policies that foster educational attainment and alleviate the matching process of workers to suitable jobs. In this context, I deem economic research necessary to inform the public and policymakers in order to devise policies that minimize economic inefficiencies and maximize social welfare.

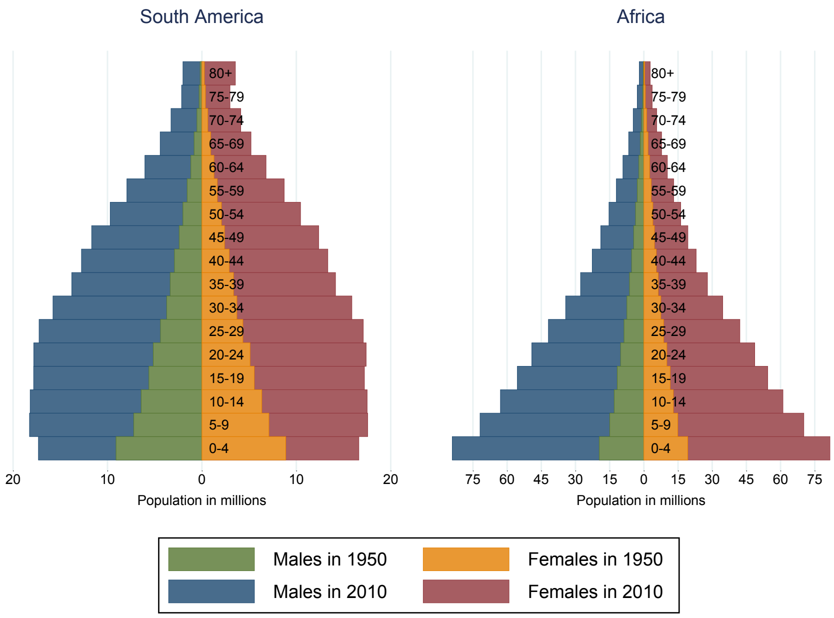
Appendix A

Appendix to Chapter 2

A.1 Additional Figures



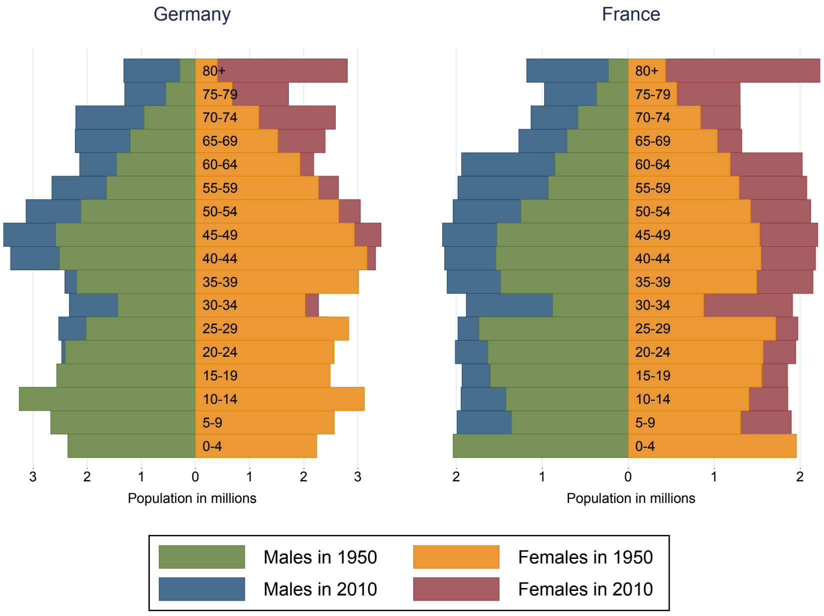
(a) World and High-Income Countries



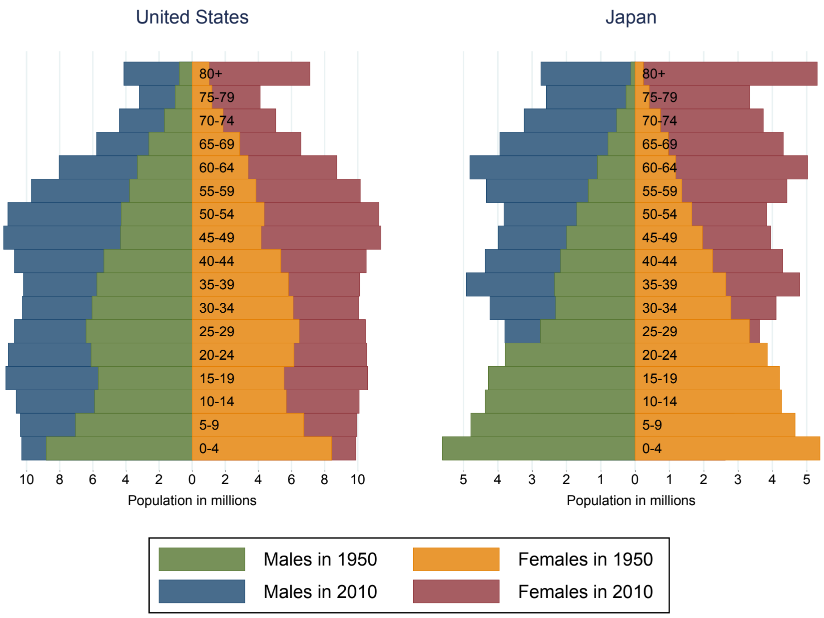
(b) South America and Africa

Figure A.1: Population Dynamics – Selected Regions

Data source: United Nations, Department of Economic and Social Affairs (2015). World Population Prospects: The 2015 Revision.



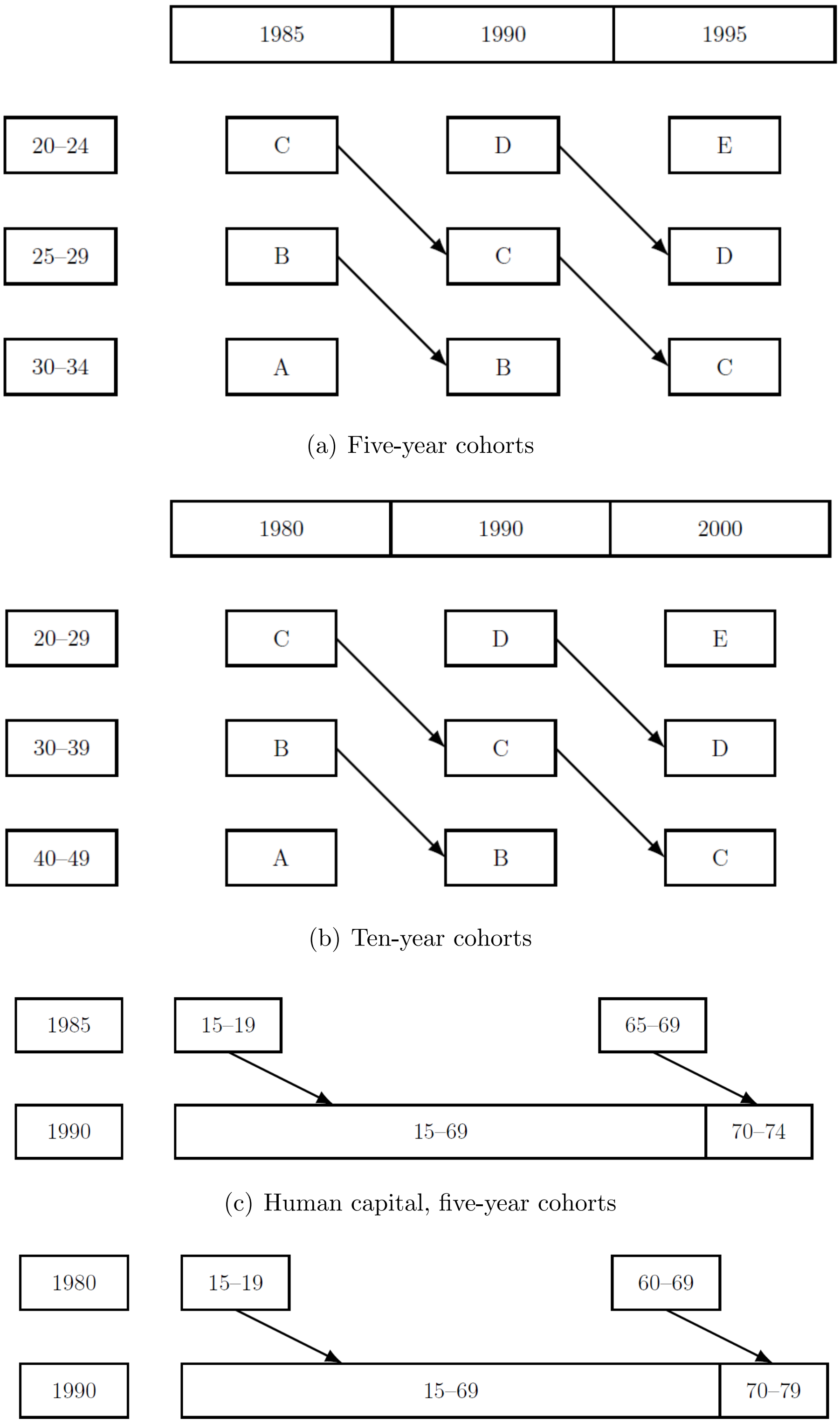
(a) Germany and France



(b) United States and Japan

Figure A.2: Population Dynamics – Selected Countries

Data source: United Nations, Department of Economic and Social Affairs (2015). World Population Prospects: The 2015 Revision.



(a) Five-year cohorts

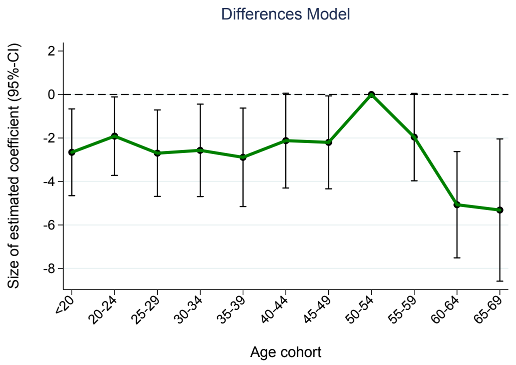
(b) Ten-year cohorts

(c) Human capital, five-year cohorts

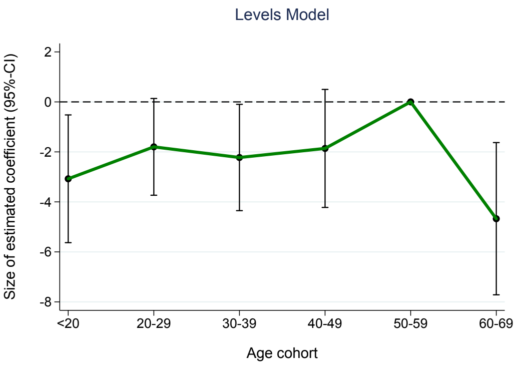
(d) Human capital, ten-year cohorts

Figure A.3: Illustration of Demographic Dynamics as Instrumental Variable

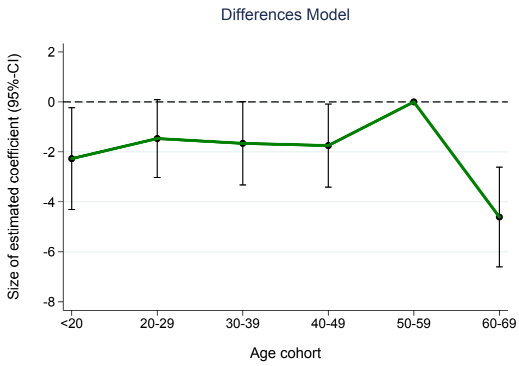
(a) Five-year cohorts (Table A.5, Column 3)



(b) Five-year cohorts (Table A.6, Column 3)

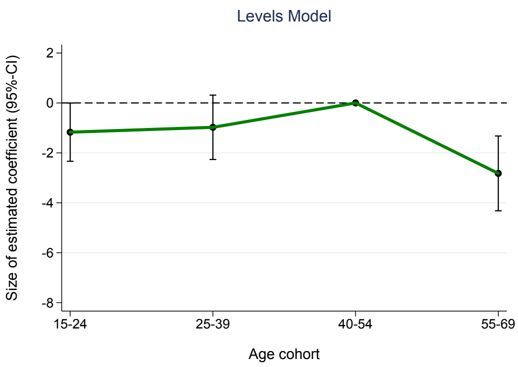


(c) Ten-year cohorts (Table A.7, Column 3)

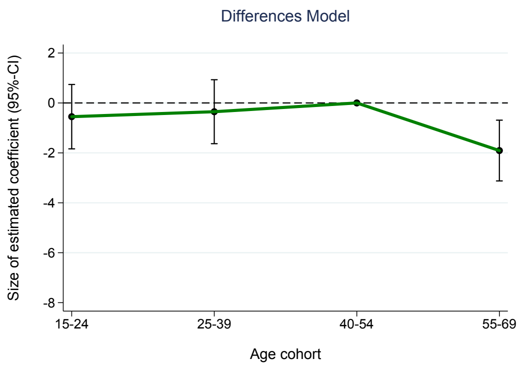


(d) Ten-year cohorts (Table A.8, Column 3)

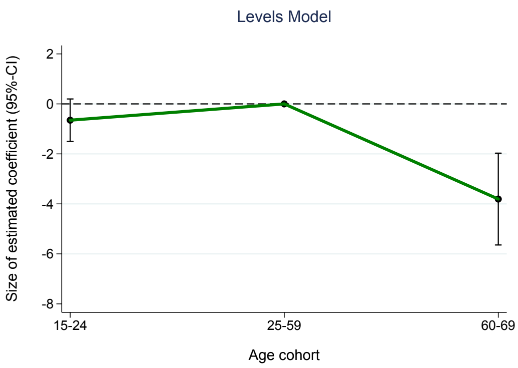
Figure A.4: Macro Productivity Profiles: Barro-Lee Data



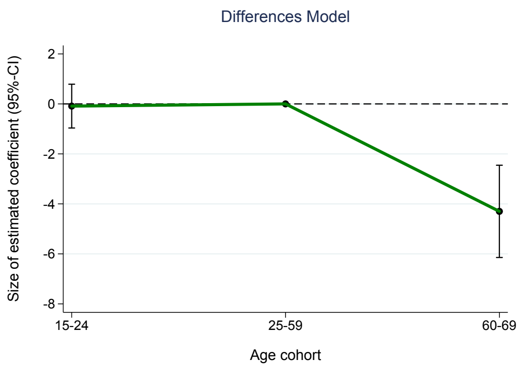
(a) Fifteen-year cohorts (estimates unreported)



(b) Fifteen-year cohorts (estimates unreported)



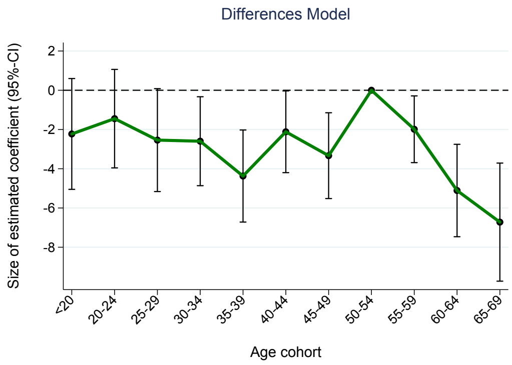
(c) Prime-age group (estimates unreported)



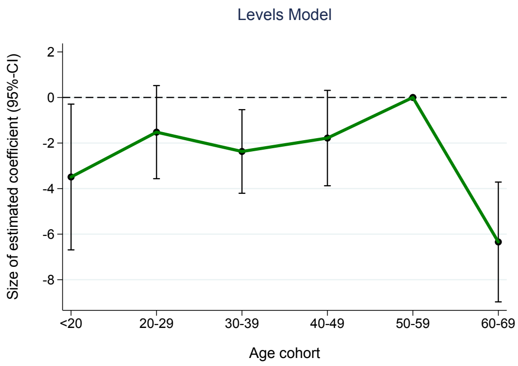
(d) Prime-age group (estimates unreported)

Figure A.5: Macro Productivity Profiles (Alternative Cohort Structure)

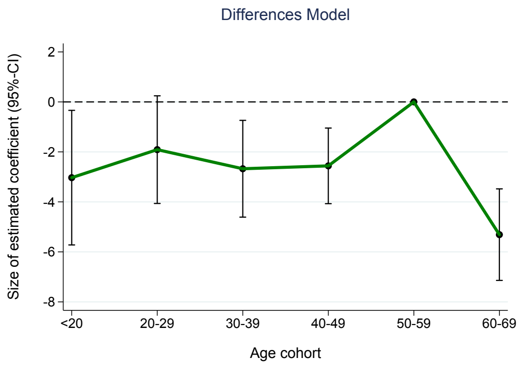
(a) Five-year cohorts (Table A.10, Column 3)



(b) Five-year cohorts (estimates unreported)



(c) Ten-year cohorts (estimates unreported)



(d) Ten-year cohorts (estimates unreported)

Figure A.6: Macro Productivity Profiles: Income Per Capita

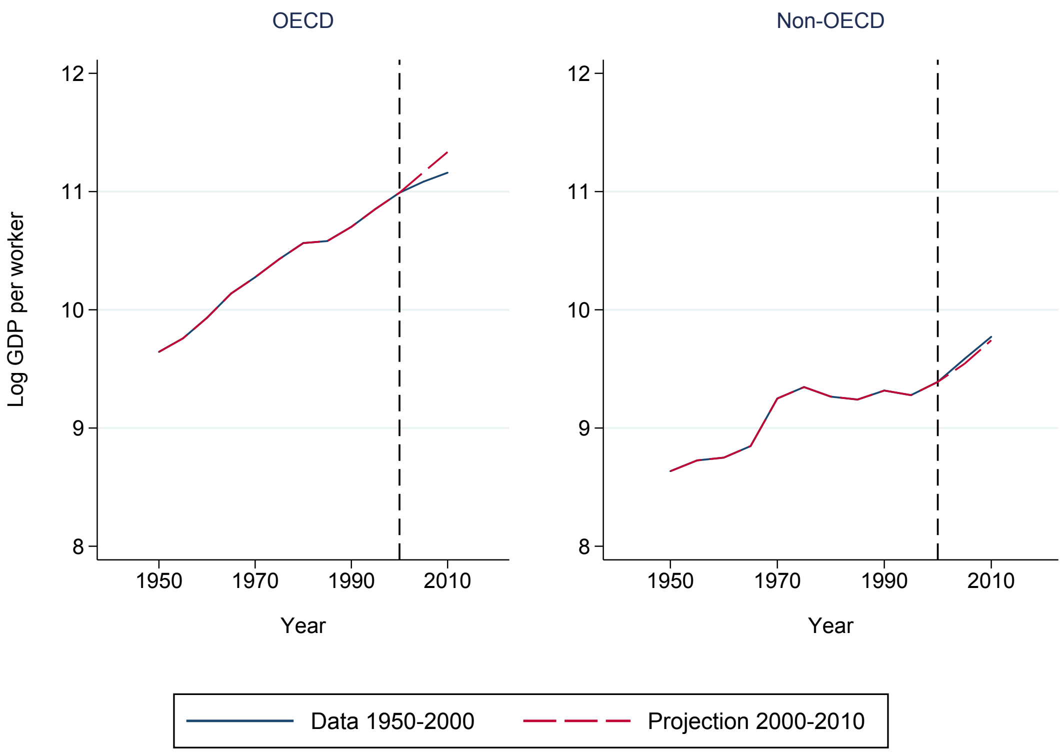
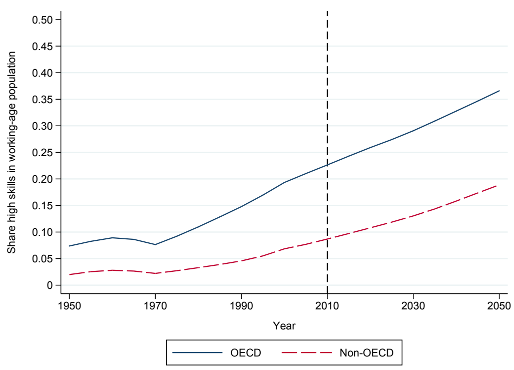
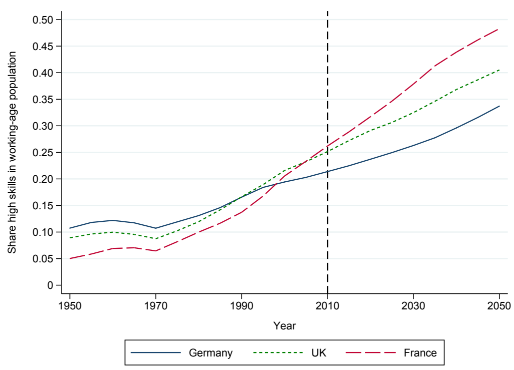


Figure A.7: Within-Sample Projection (2000–2010)

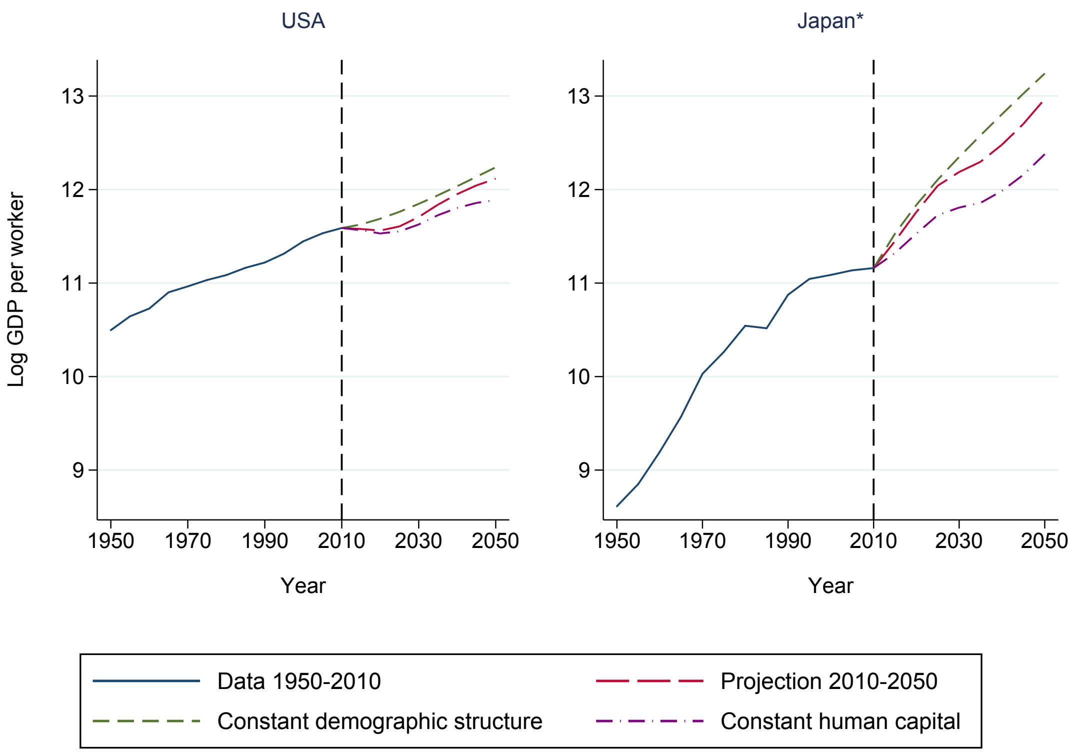


(a) OECD and Non-OECD Countries

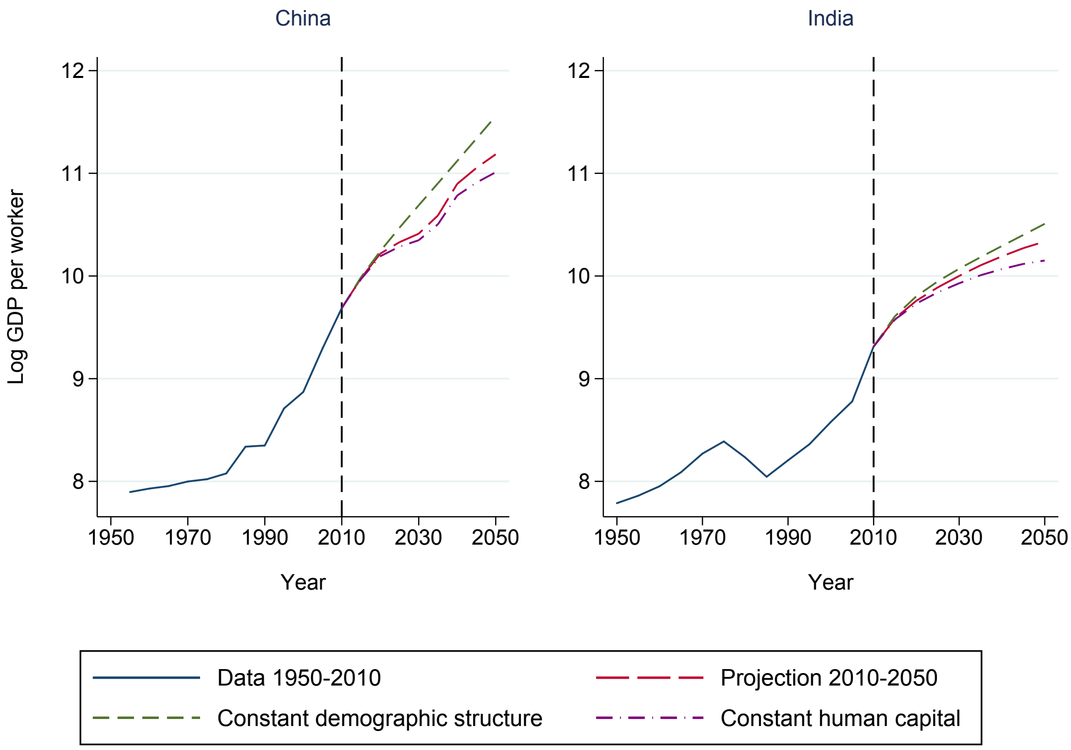


(b) Germany, United Kingdom and France

Figure A.8: Actual and Projected Educational Attainment

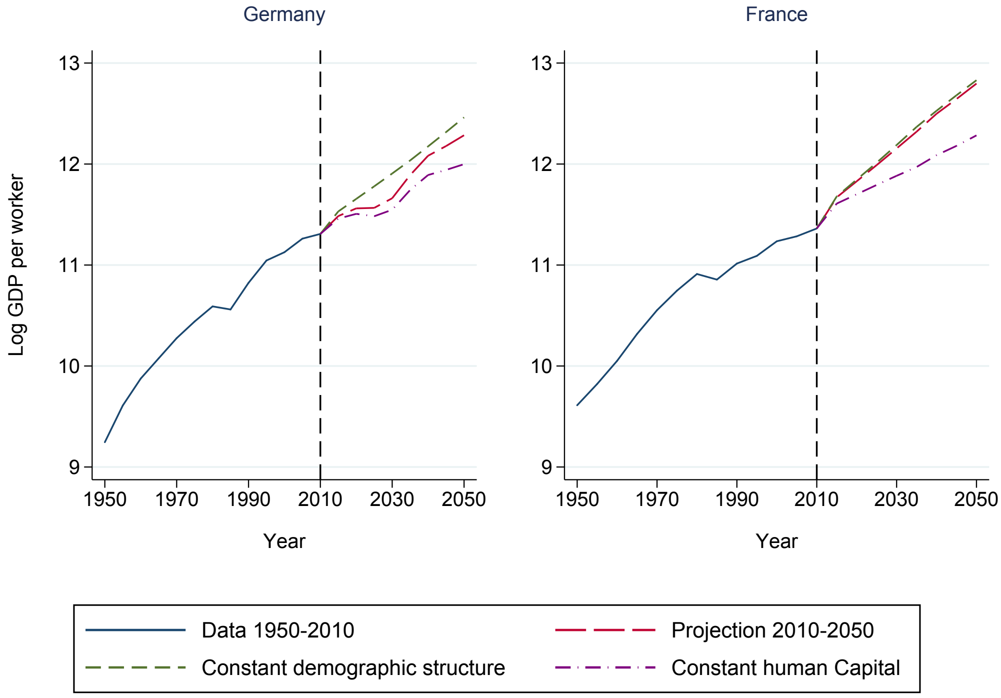


(a) Selected Countries: USA and Japan

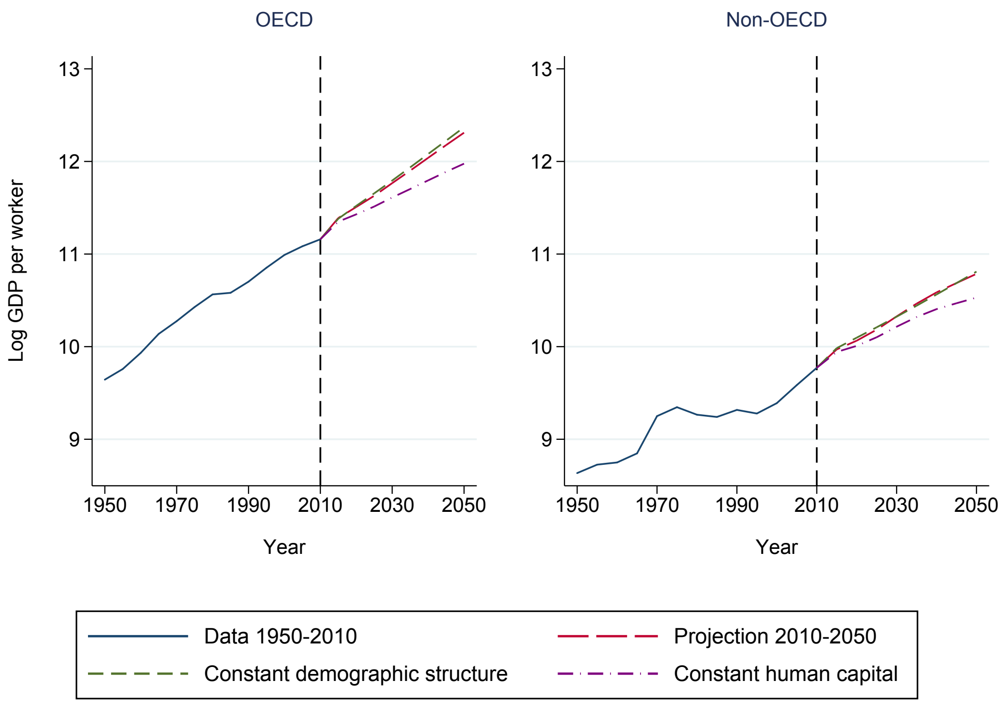


(b) Selected Countries: China and India

Figure A.9: Projections Under Different Scenarios

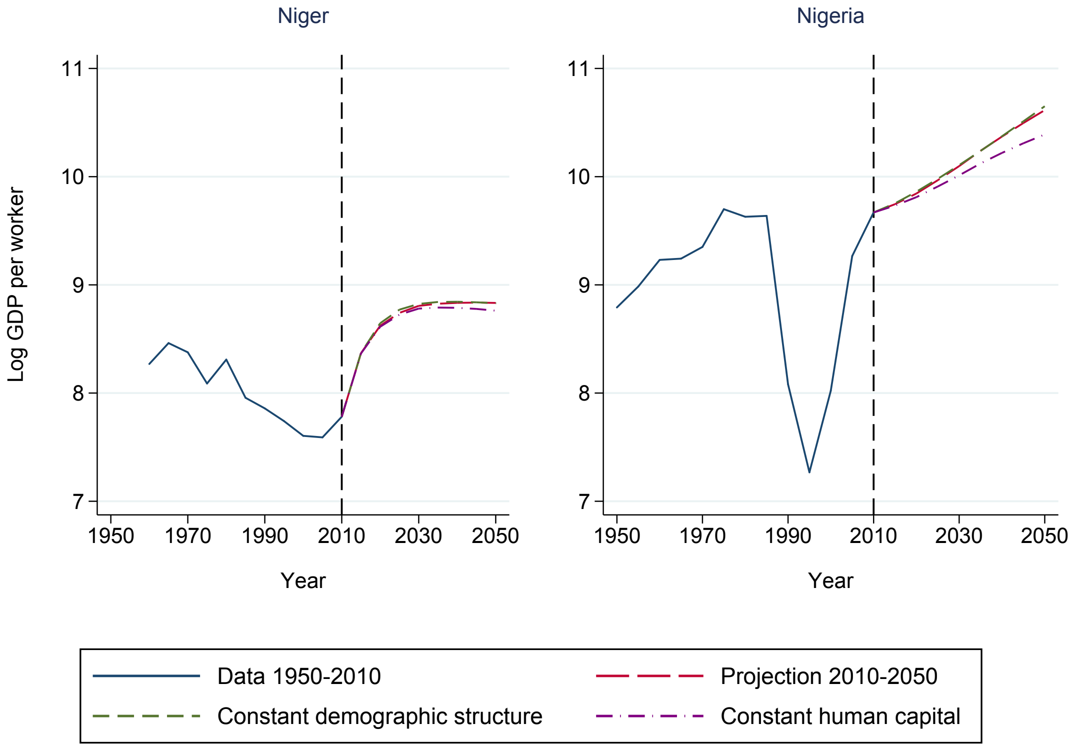


(a) Selected Countries: Germany and France

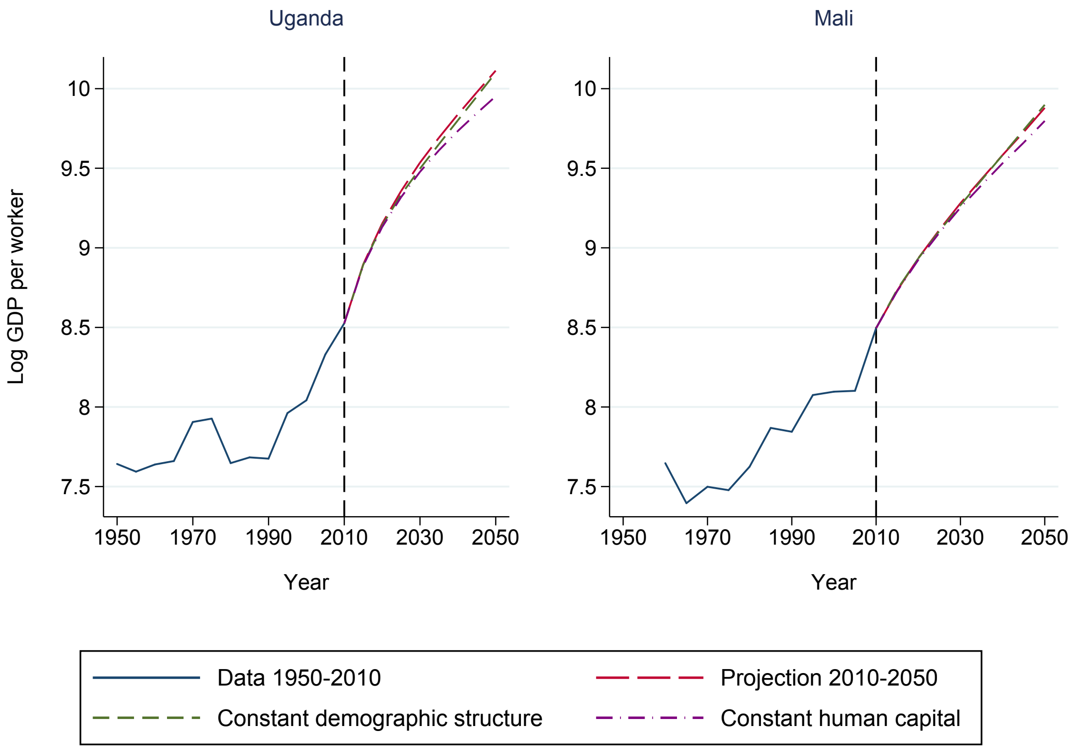


(b) Selected Regions: OECD and Non-OECD Countries

Figure A.10: Projections for Model Without Lagged Dependent Variable

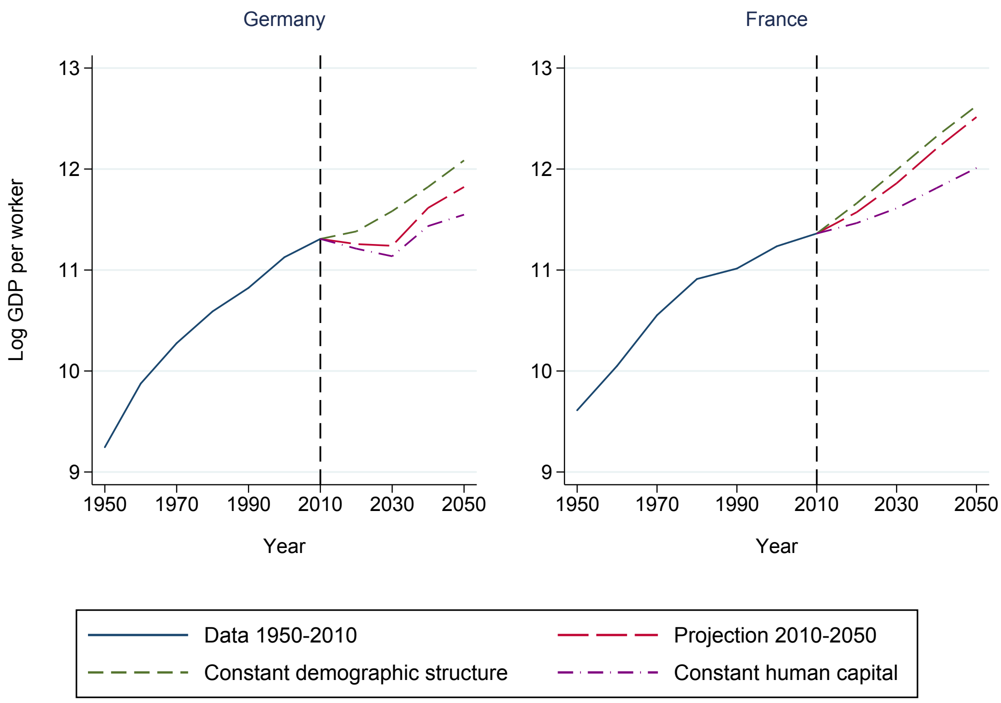


(a) Selected Countries: Niger and Nigeria

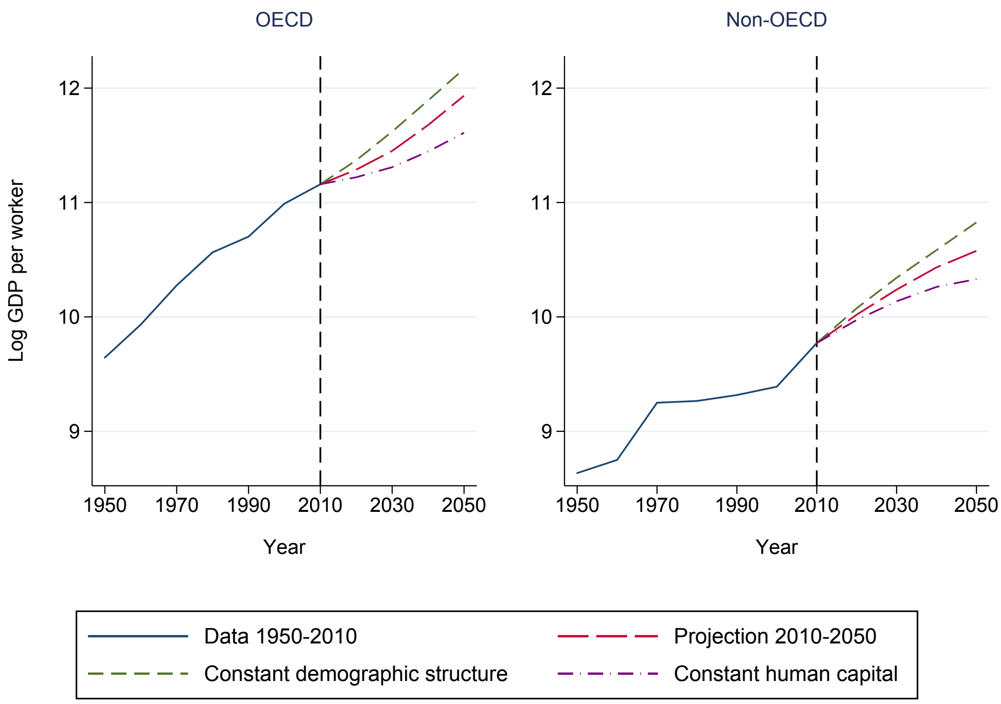


(b) Selected Countries: Uganda and Mali

Figure A.11: Projections Under Different Scenarios

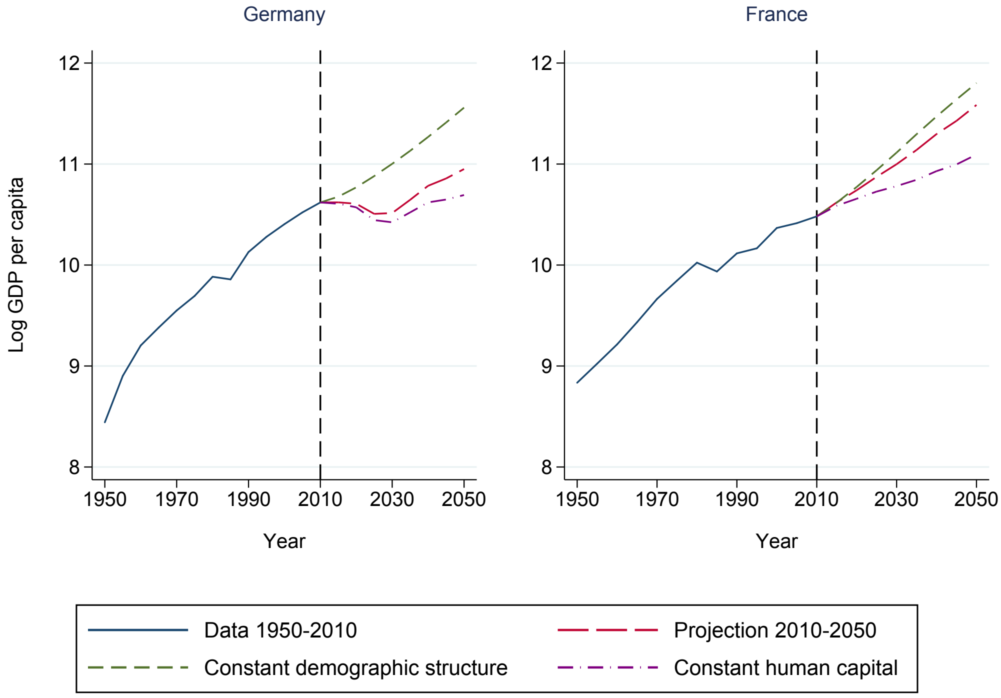


(a) Selected Countries: Germany and France

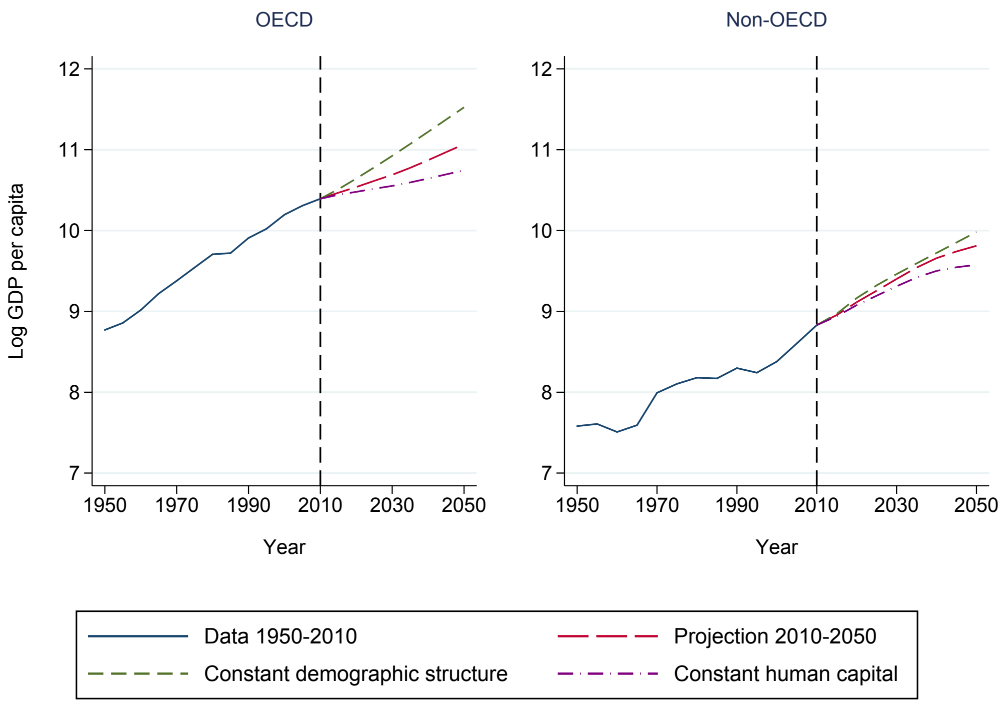


(b) Selected Regions: OECD and Non-OECD Countries

Figure A.12: Projections for Model with Prime-Age Group (See Figure A.5, Panel b)

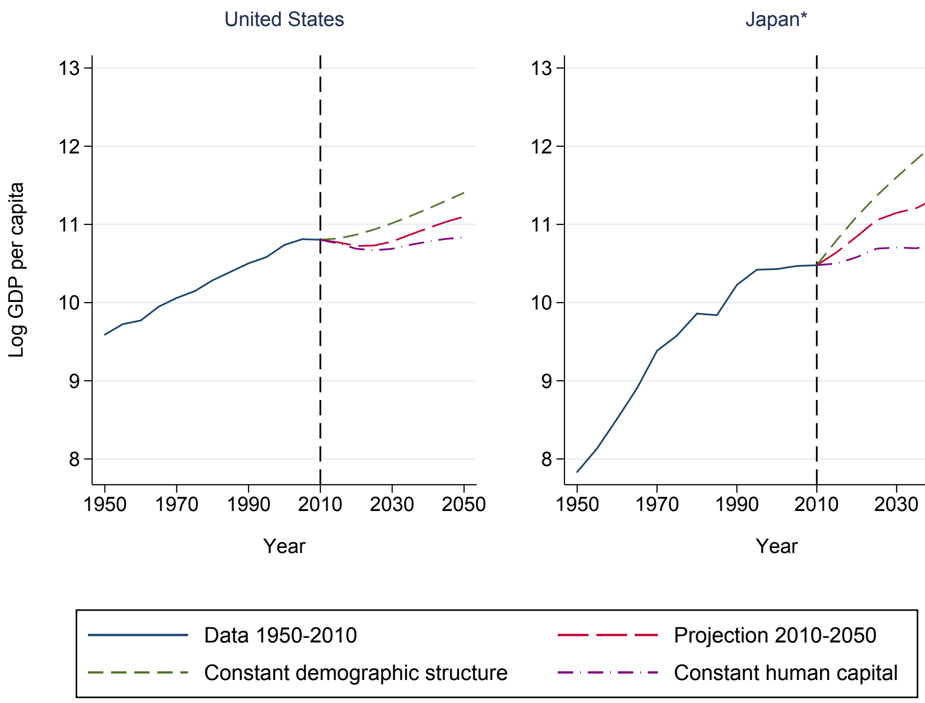


(a) Selected Countries: Germany and France

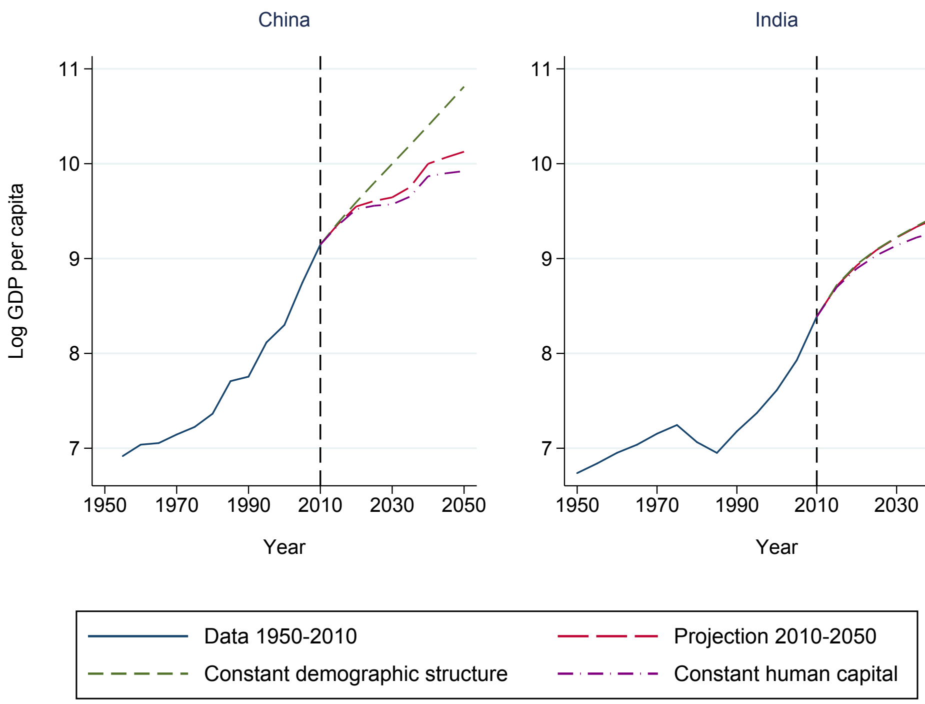


(b) Selected Regions: OECD and Non-OECD Countries

Figure A.13: Projections for Income Per Capita

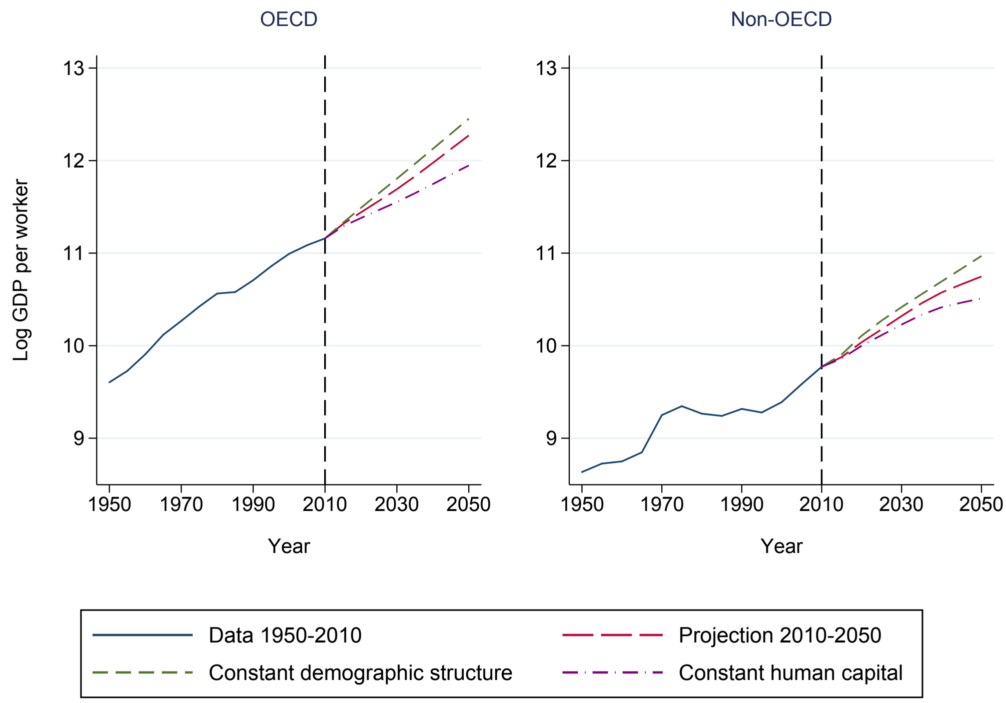


(a) Selected Countries: USA and Japan

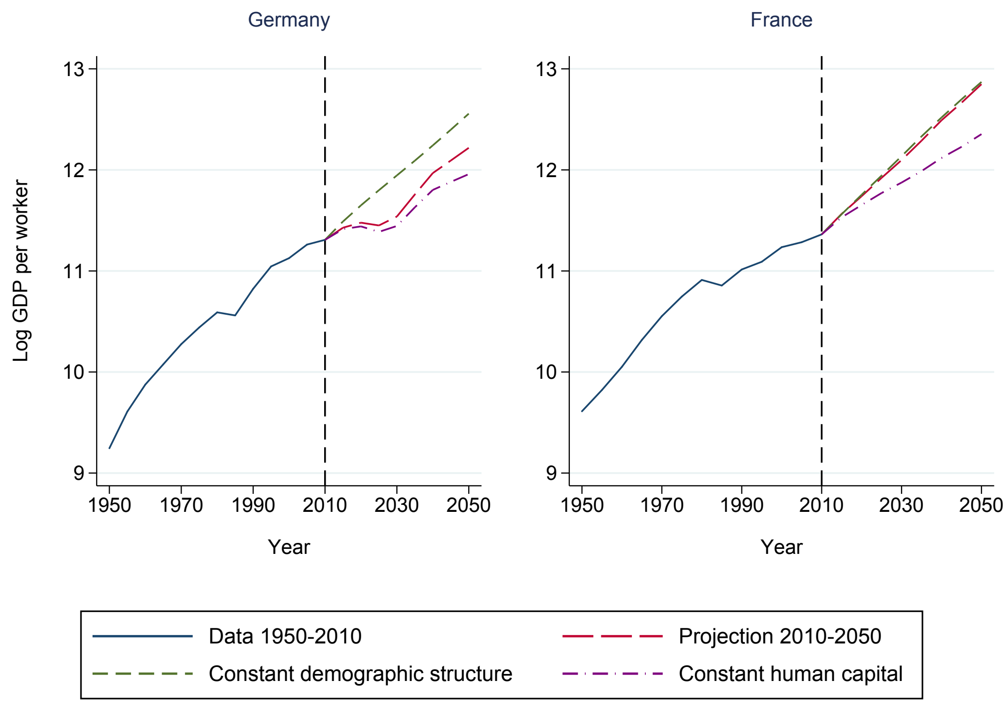


(b) Selected Countries: China and India

Figure A.14: Projections for Income Per Capita

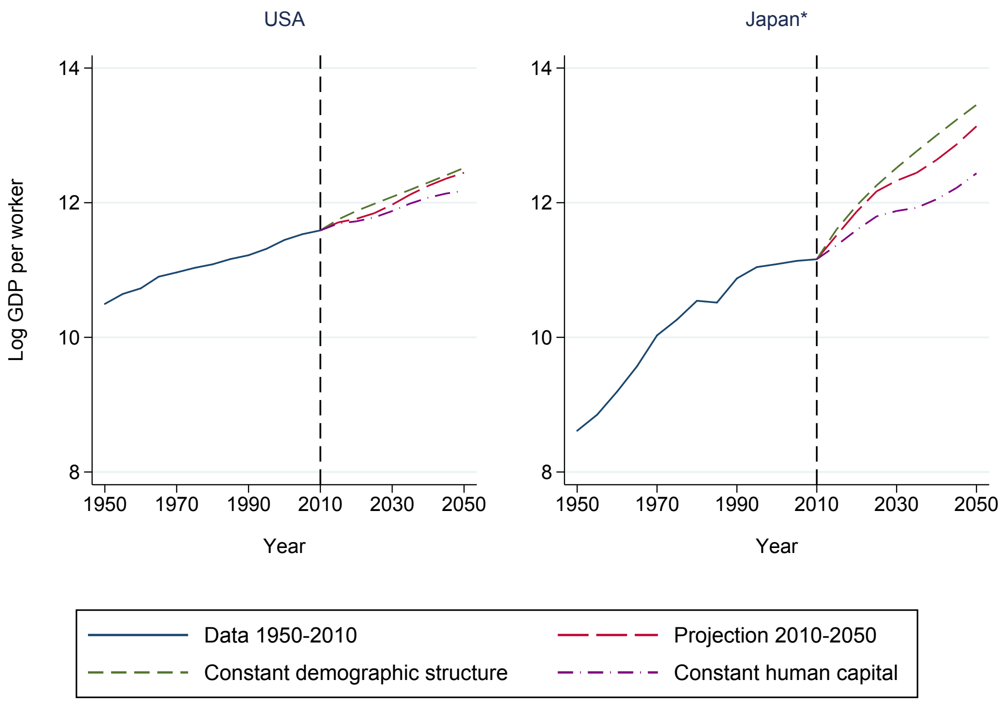


(a) Selected Regions: OECD and Non-OECD Countries

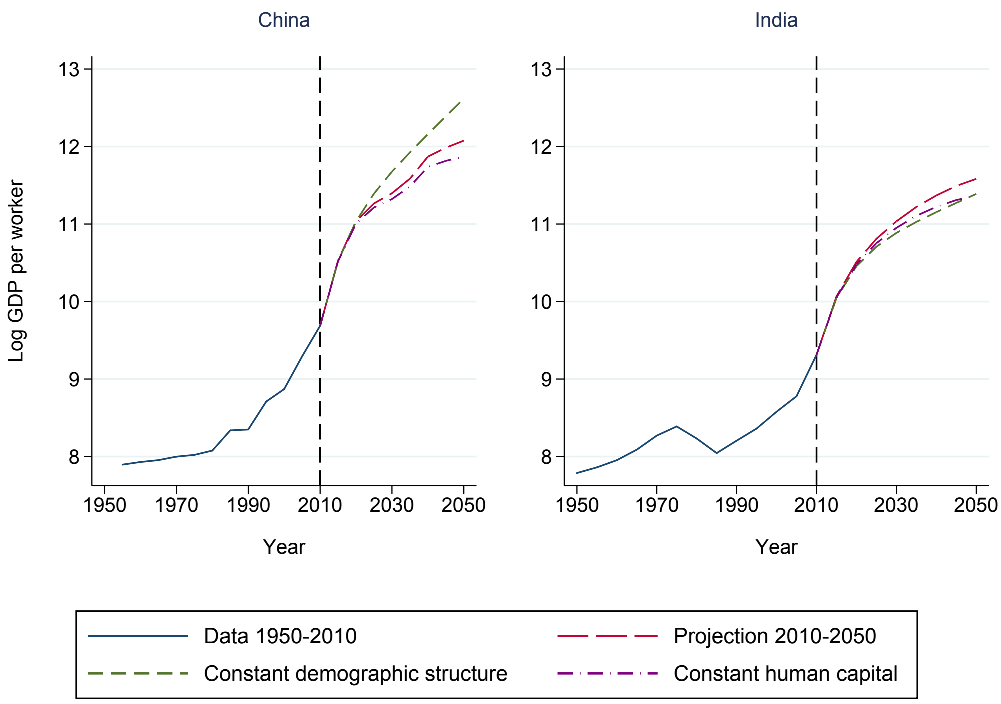


(b) Selected Countries: Germany and France

Figure A.15: Projections when Controlling for the Size of the Working-Age Population

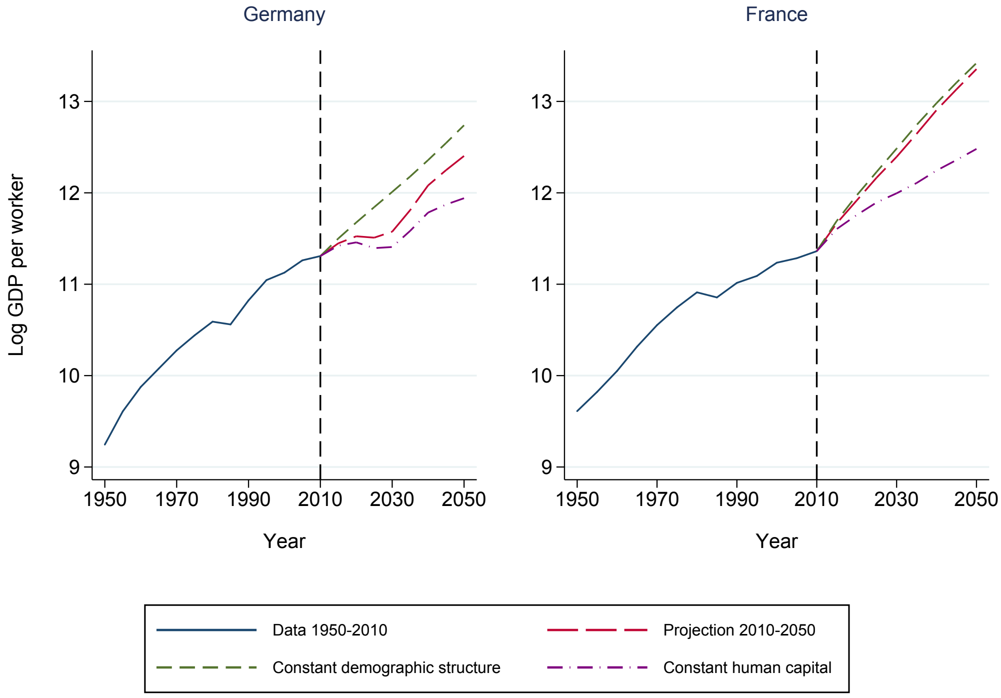


(a) Selected Countries: USA and Japan

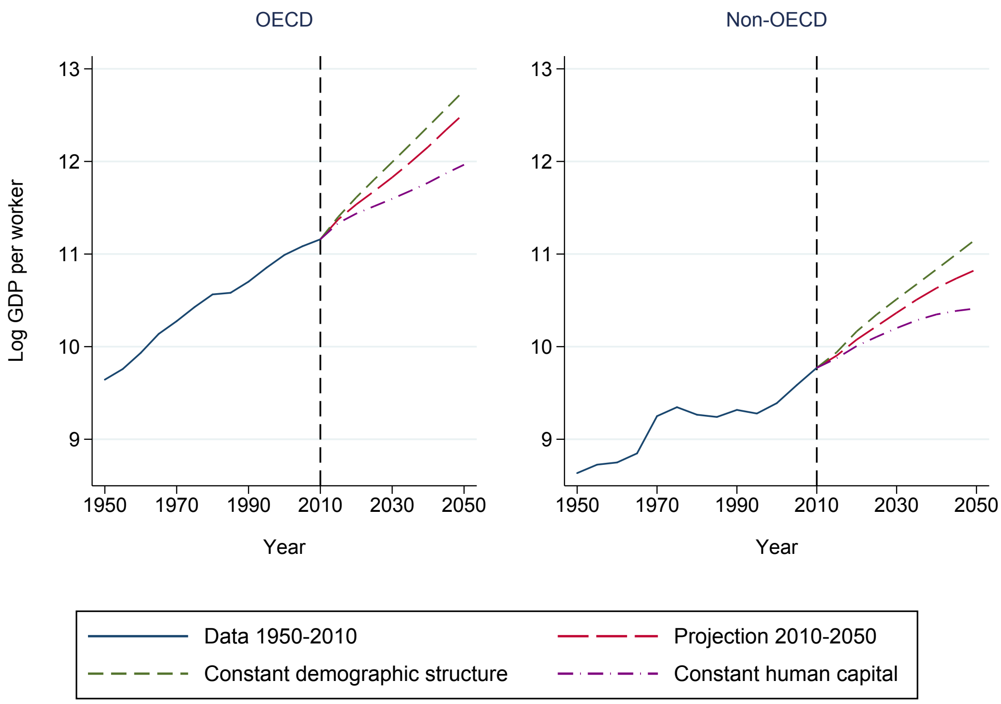


(b) Selected Countries: China and India

Figure A.16: Projections when Controlling for the Size of the Working-Age Population

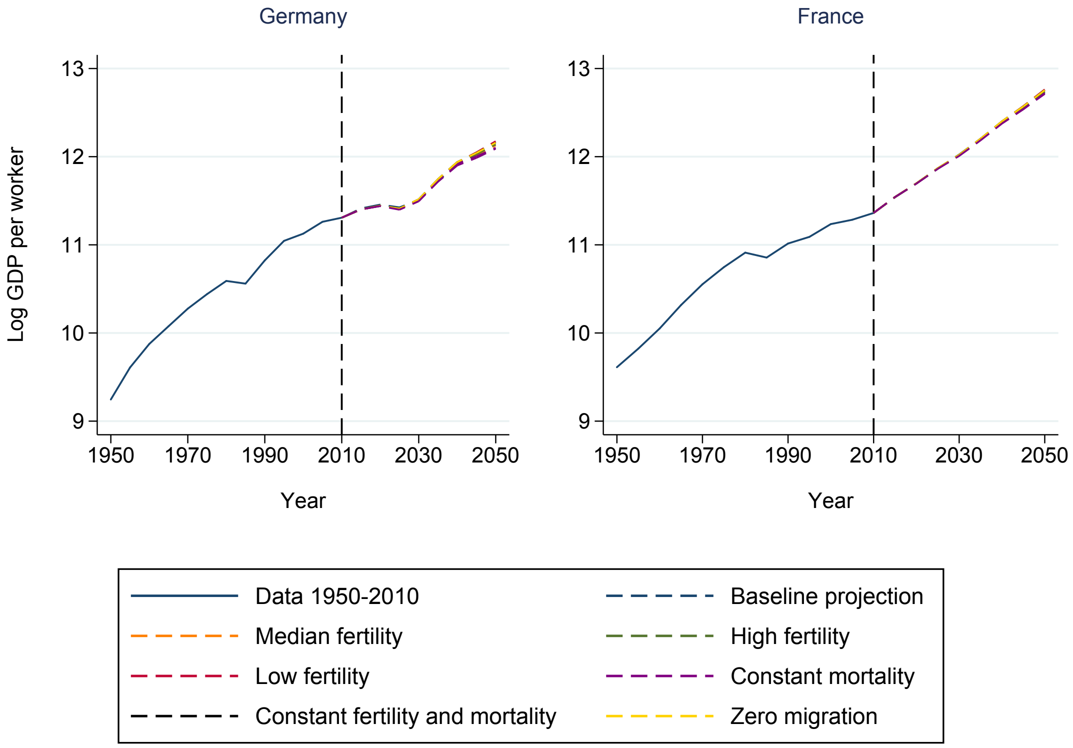


(a) Selected Countries: Germany and France

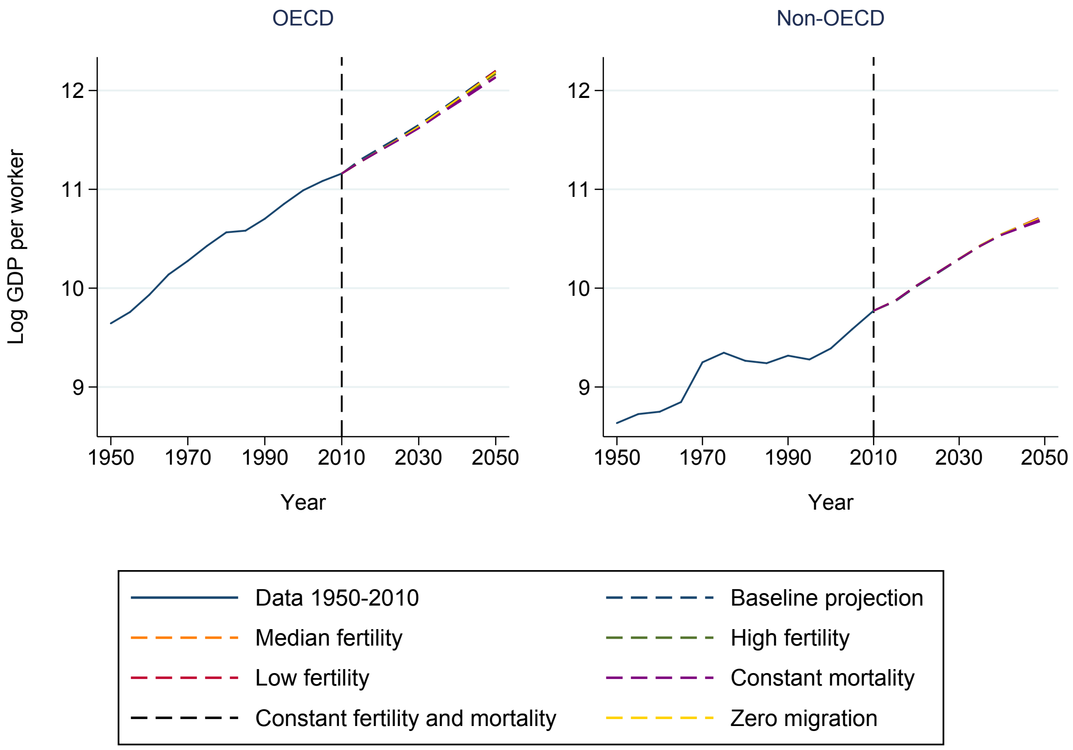


(b) Developed vs. Developing Economies

Figure A.17: Projections for Instrumental Variables Model

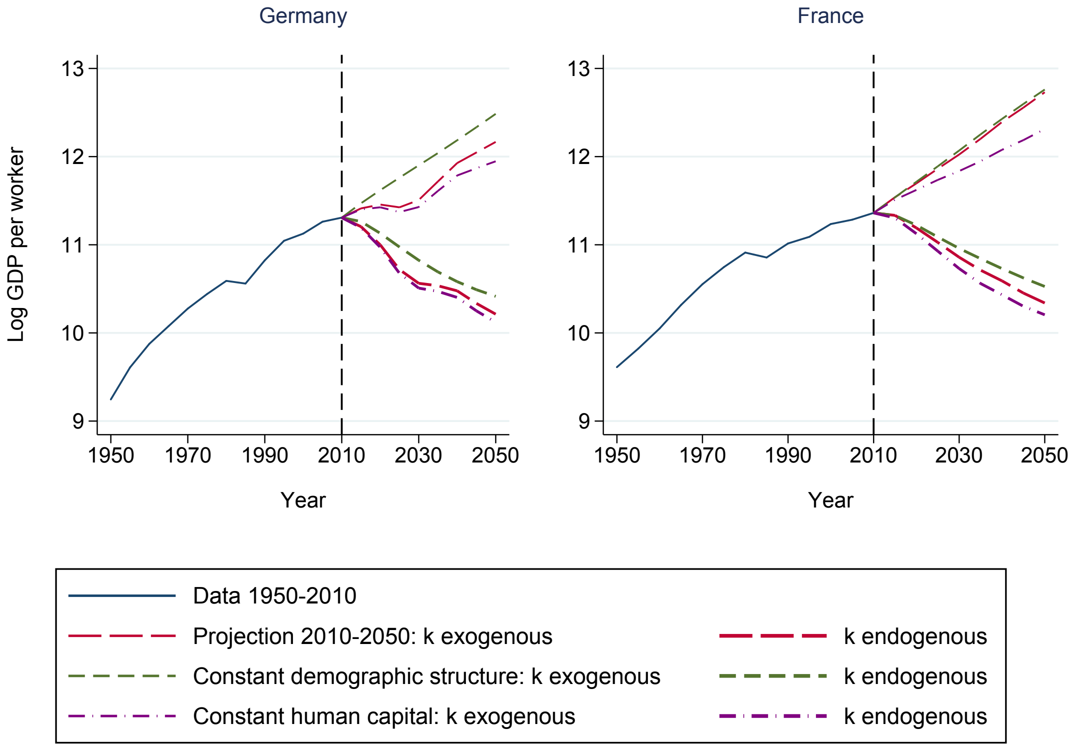


(a) Germany vs. France

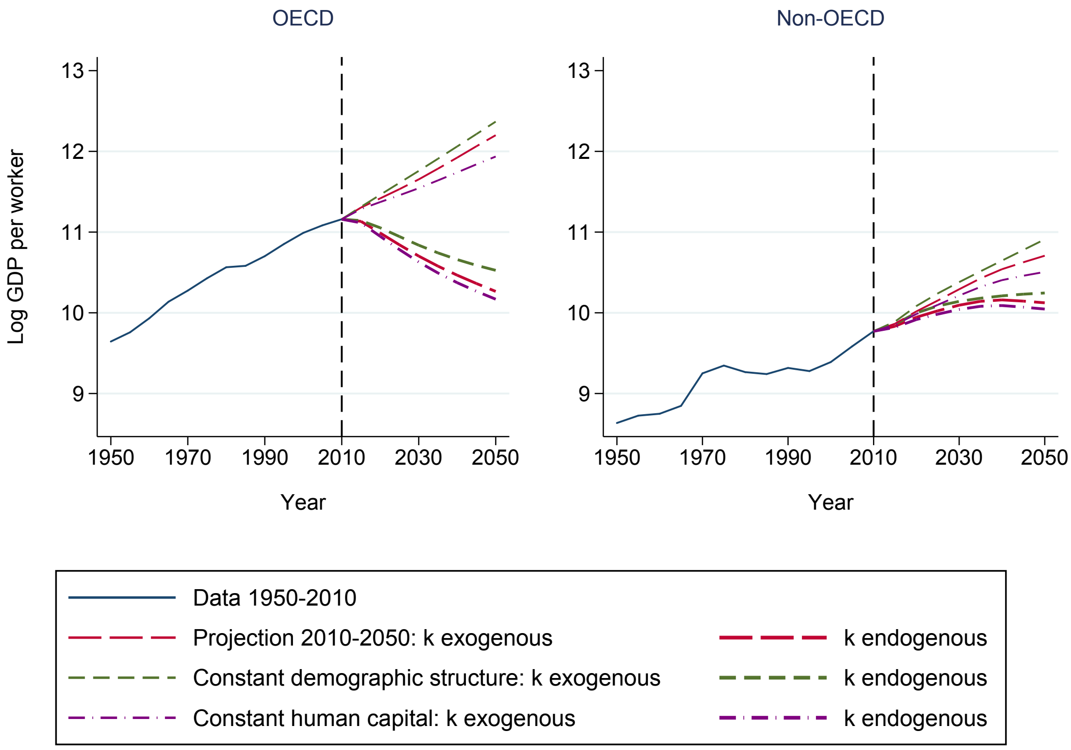


(b) OECD vs. Non-OECD Countries

Figure A.18: Projections with Alternative Assumptions About Fertility, Mortality, and Migration

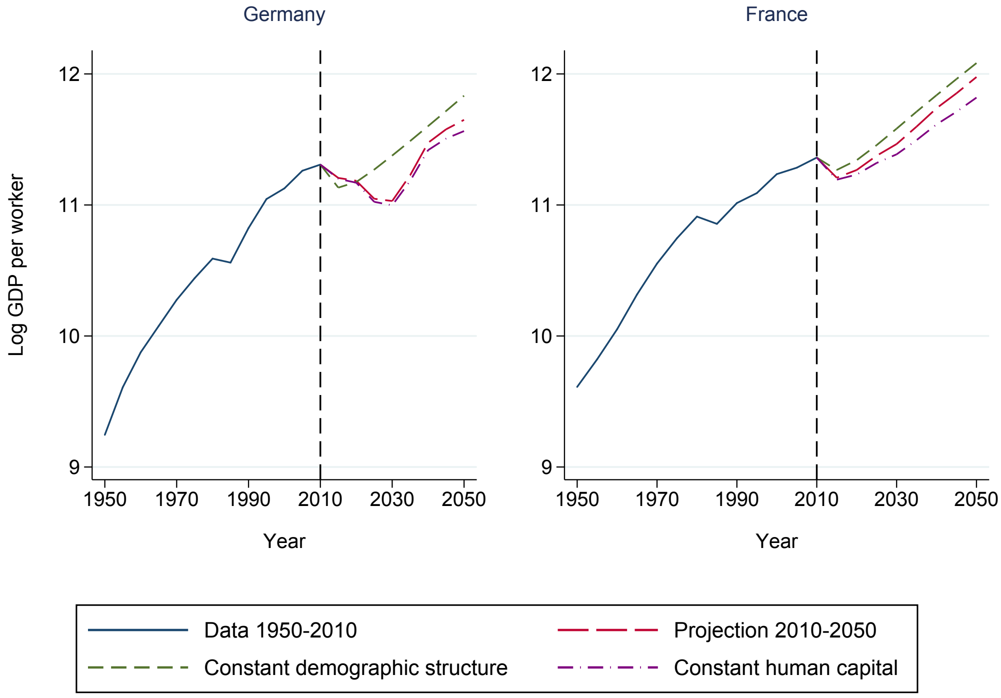


(a) Germany vs. France

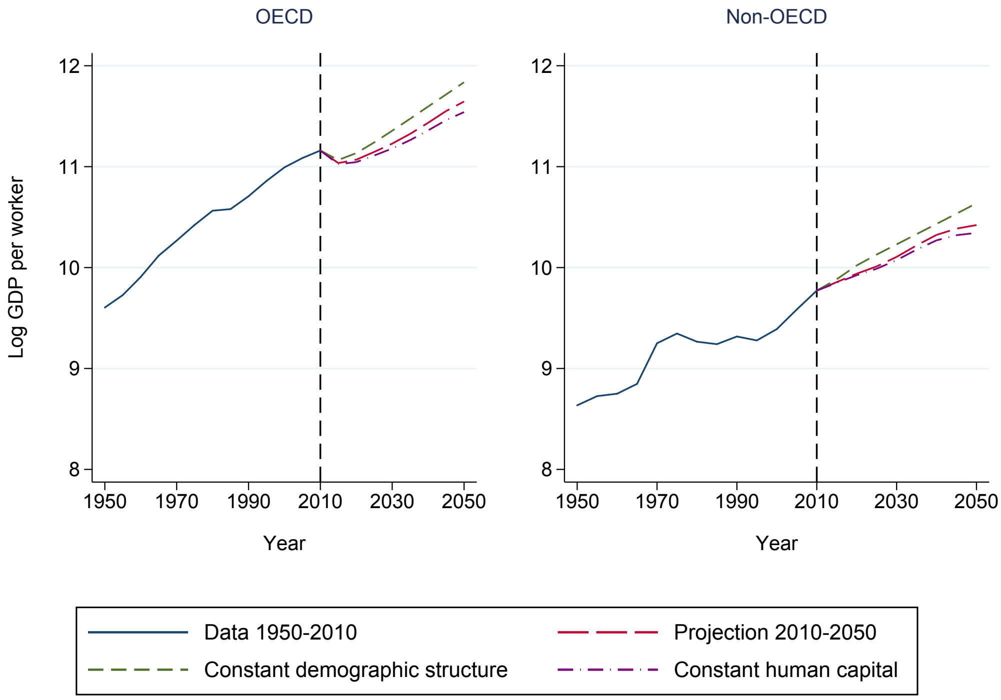


(b) OECD vs. Non-OECD Countries

Figure A.19: Projections with Endogenous Capital

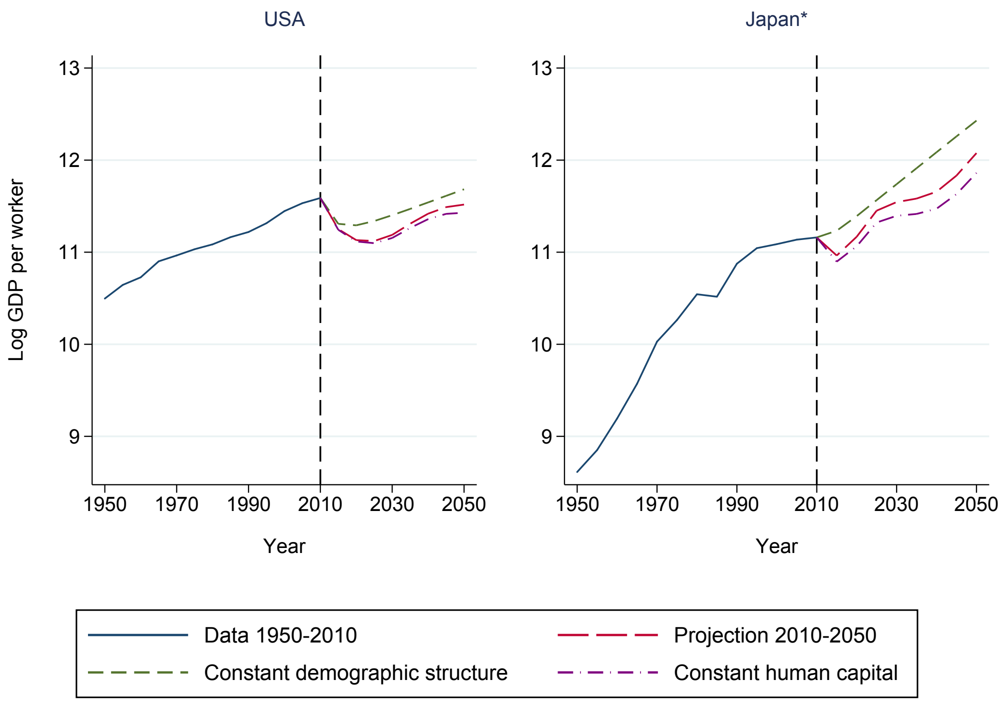


(a) Selected Countries: Germany and France

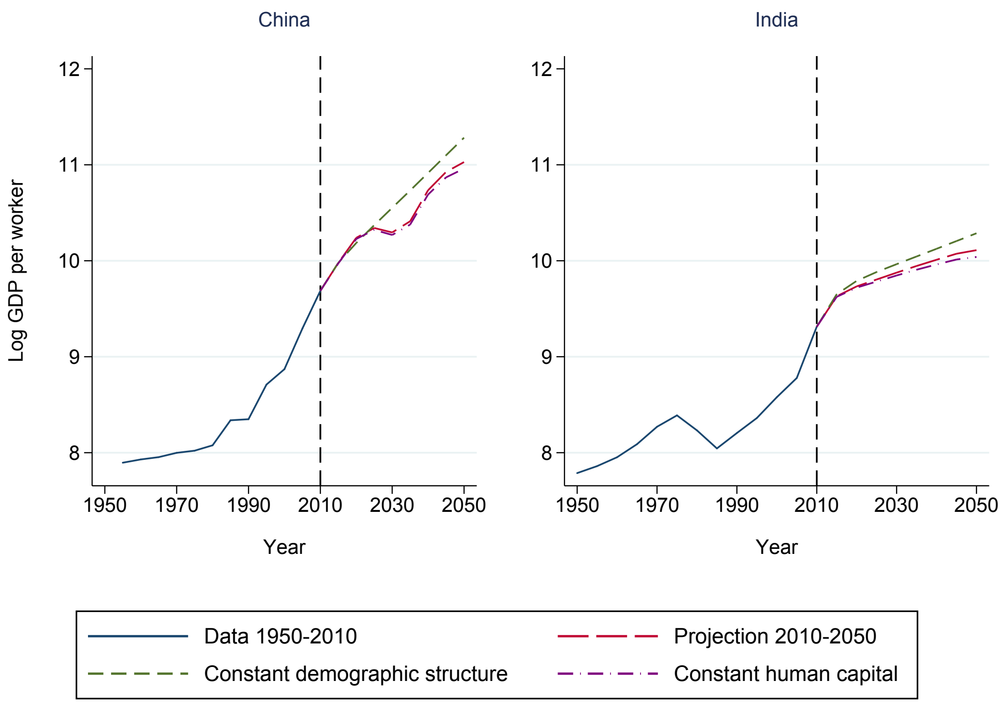


(b) Selected Regions: OECD and Non-OECD Countries

Figure A.20: Projections for 1990–2010 Sample

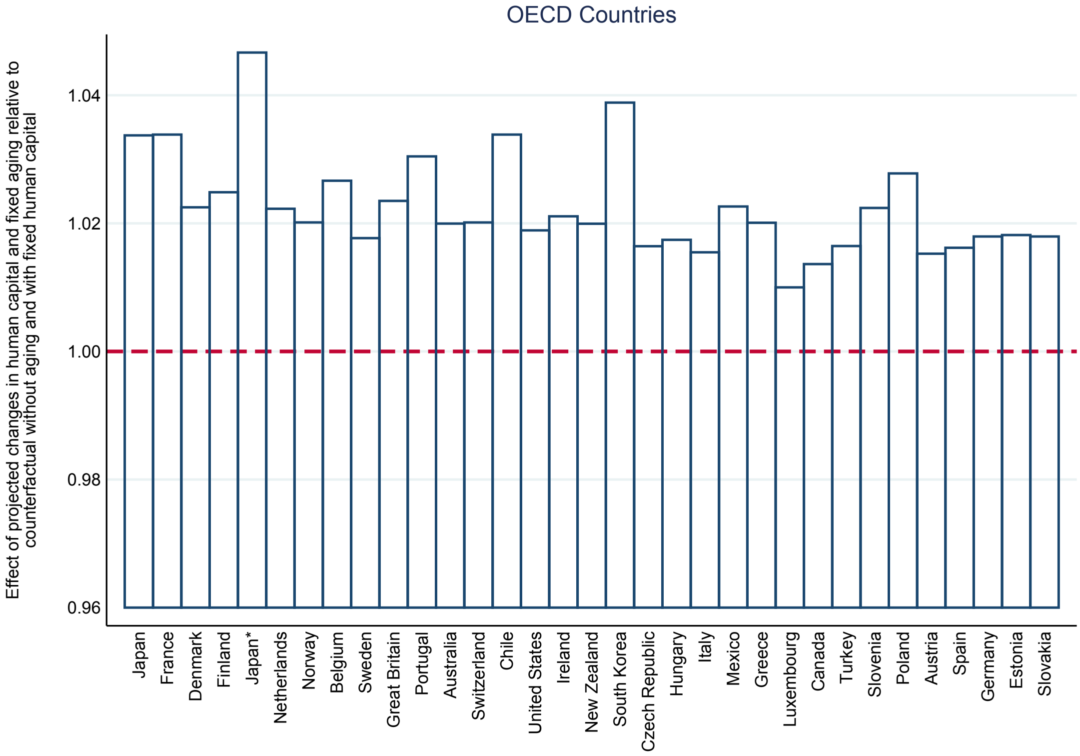


(a) Selected Countries: USA and Japan

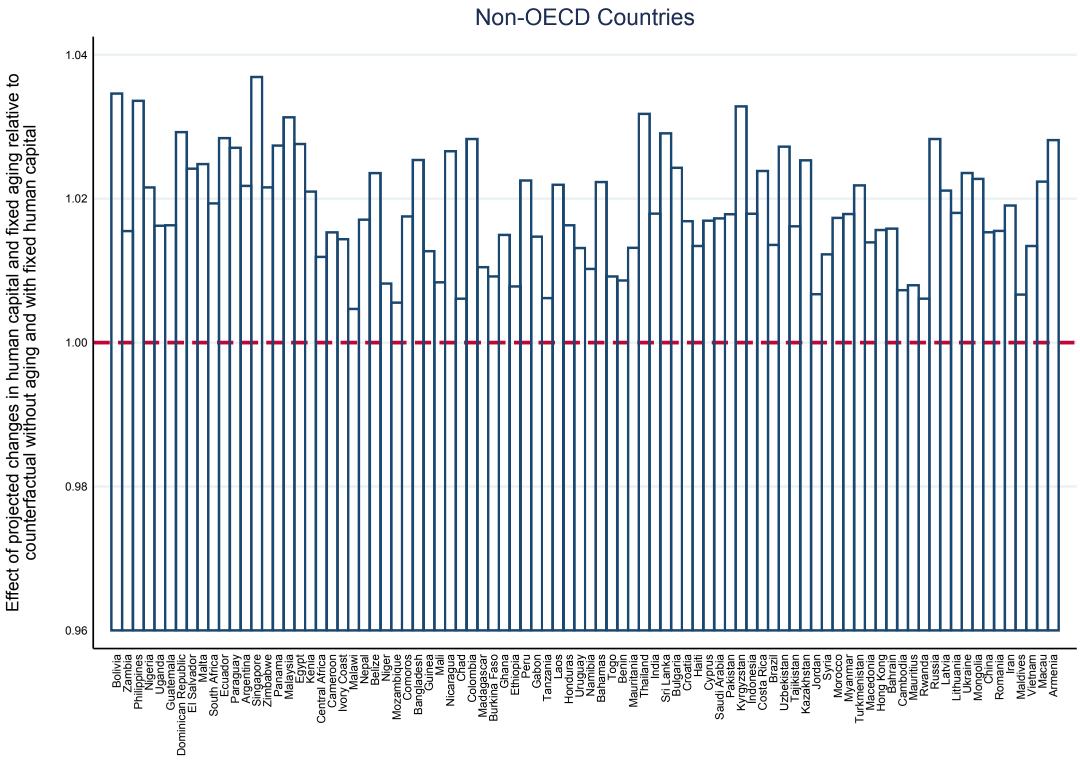


(b) Selected Countries: China and India

Figure A.21: Projections for 1990–2010 Sample

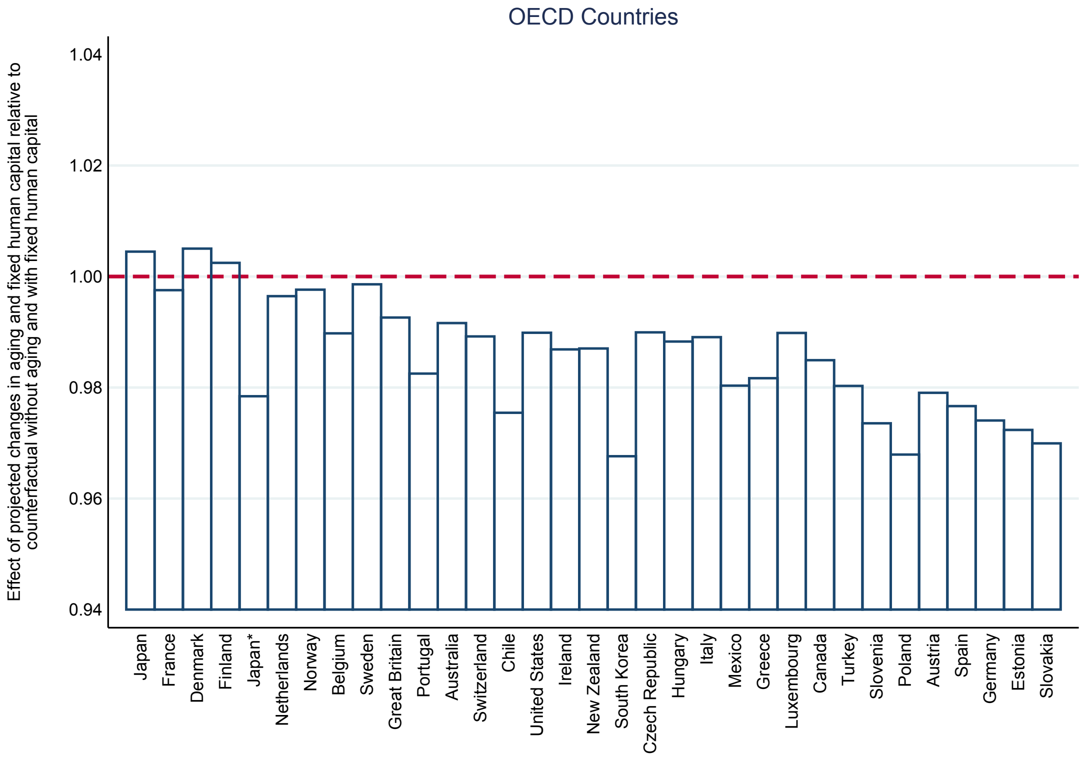


(a) OECD Countries

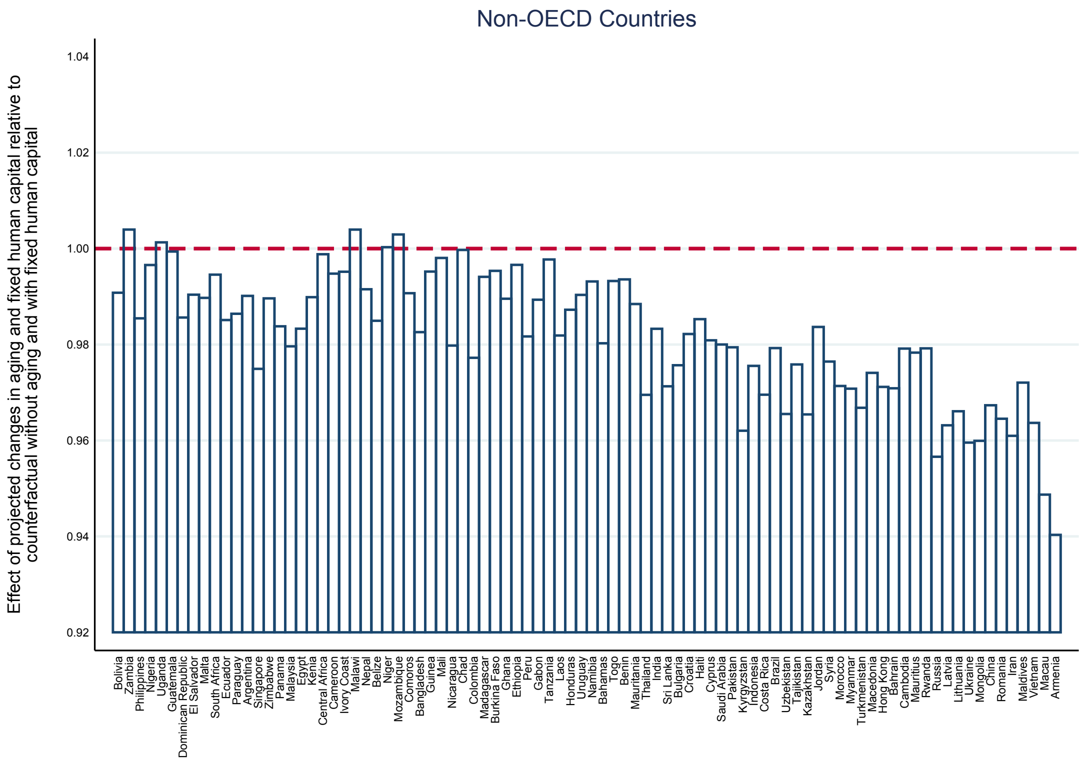


(b) Non-OECD Countries

Figure A.22: Economic Performance for Constant Relative to Changing Demographic Structure

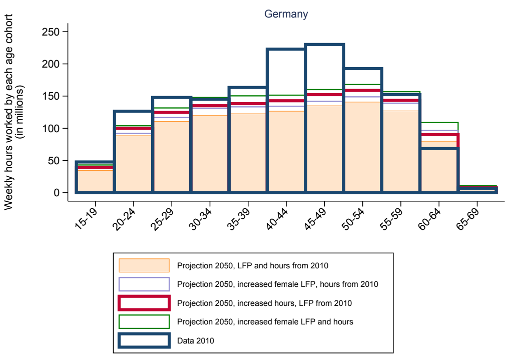


(a) OECD Countries

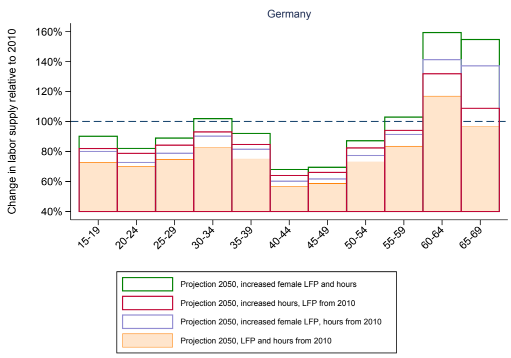


(b) Non-OECD Countries

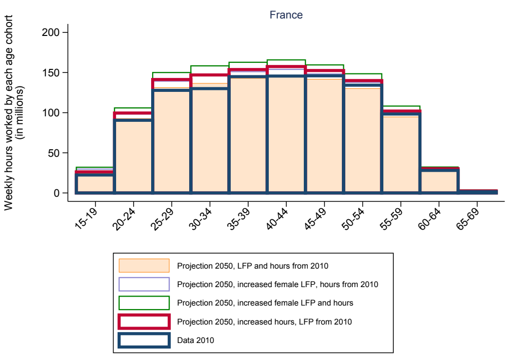
Figure A.23: Economic Performance for Constant Relative to Changing Human Capital



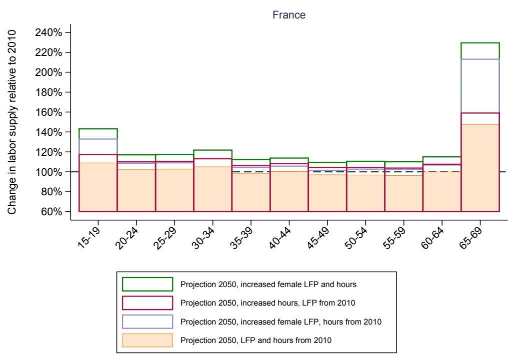
(a) Germany (in hours)



(b) Germany (in percent)

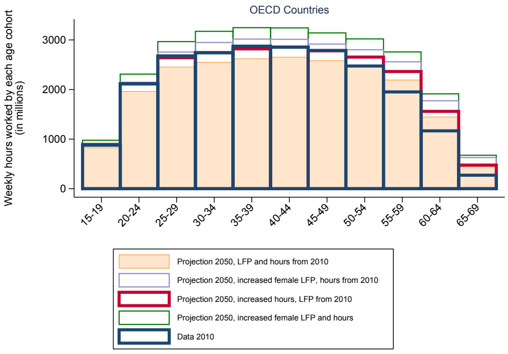


(c) France (in hours)

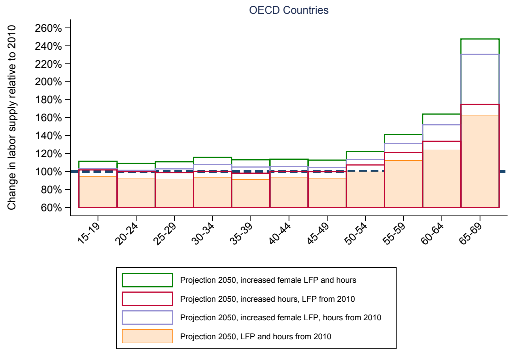


(d) France (in percent)

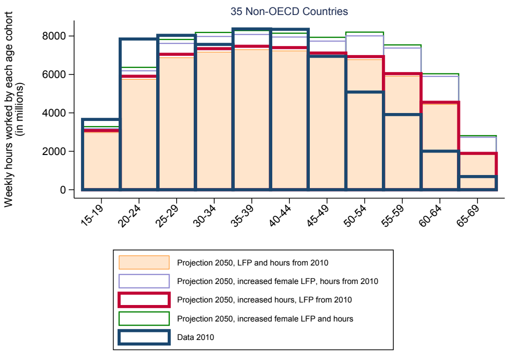
Figure A.24: Projected Change in Cohort Labor Supply Between 2010 and 2050



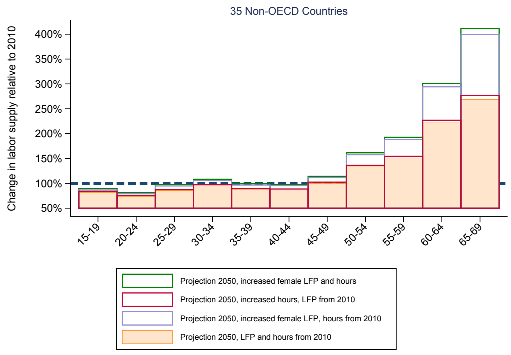
(a) OECD Countries (in hours)



(b) OECD Countries (in percent)

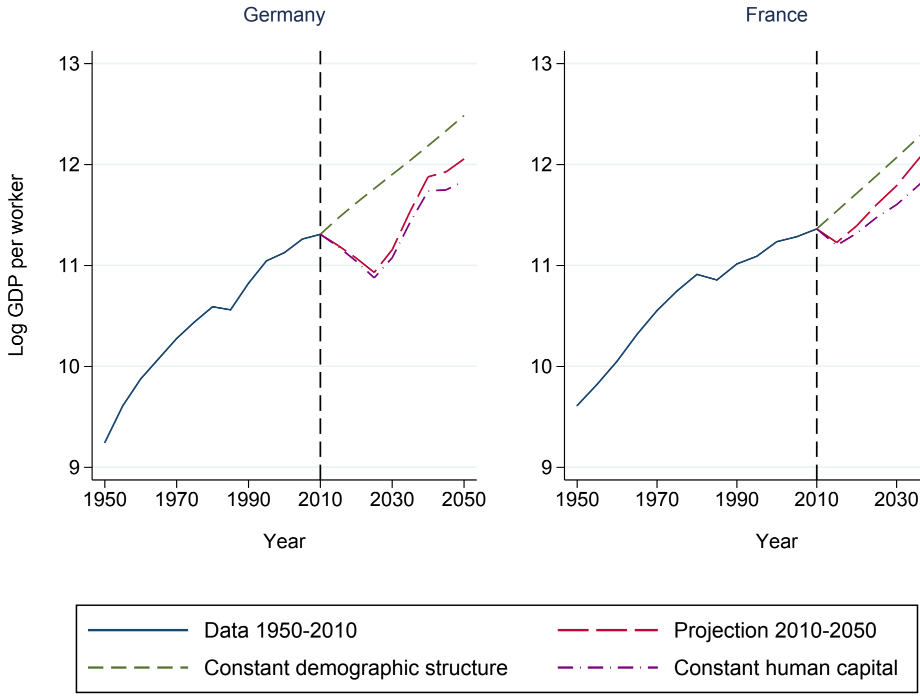


(c) 35 Non-OECD Countries (in hours)

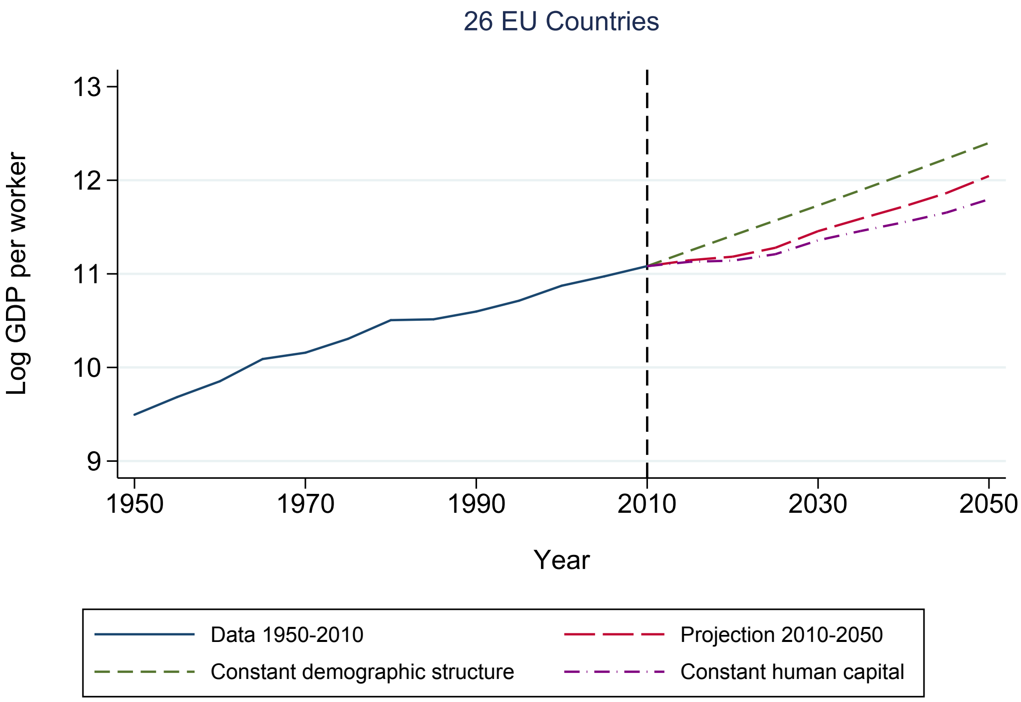


(d) 35 Non-OECD Countries (in percent)

Figure A.25: Projected Change in Cohort Labor Supply Between 2010 and 2050

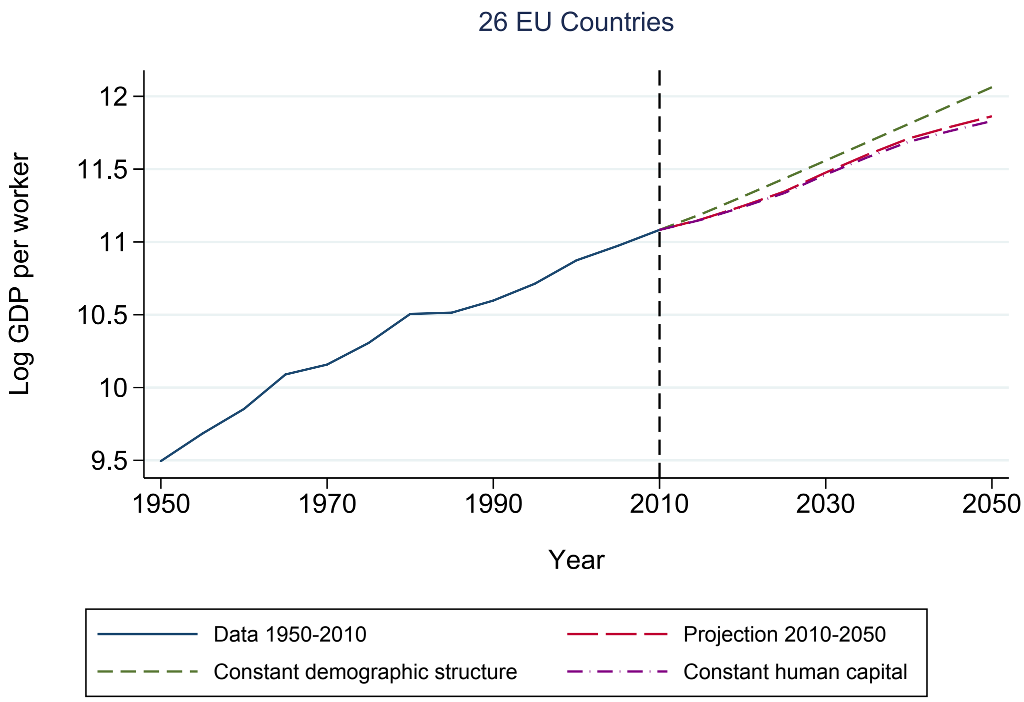


(a) Germany and France

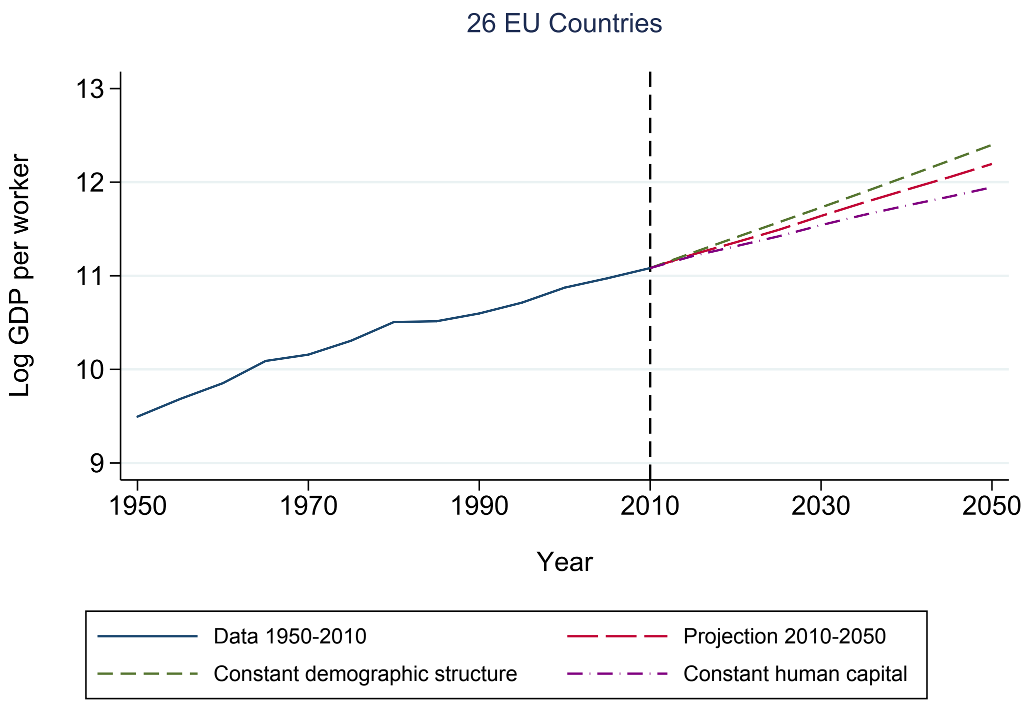


(b) 26 EU Countries

Figure A.26: Reduced-Form Projections: Changing Labor Force Participation Based on Estimates from Working-Age Population



(a) Estimates based on Labor Force



(b) Estimates based on Working-Age Population (Baseline)

Figure A.27: Projections: Labor Force Participation Versus Working-Age Population

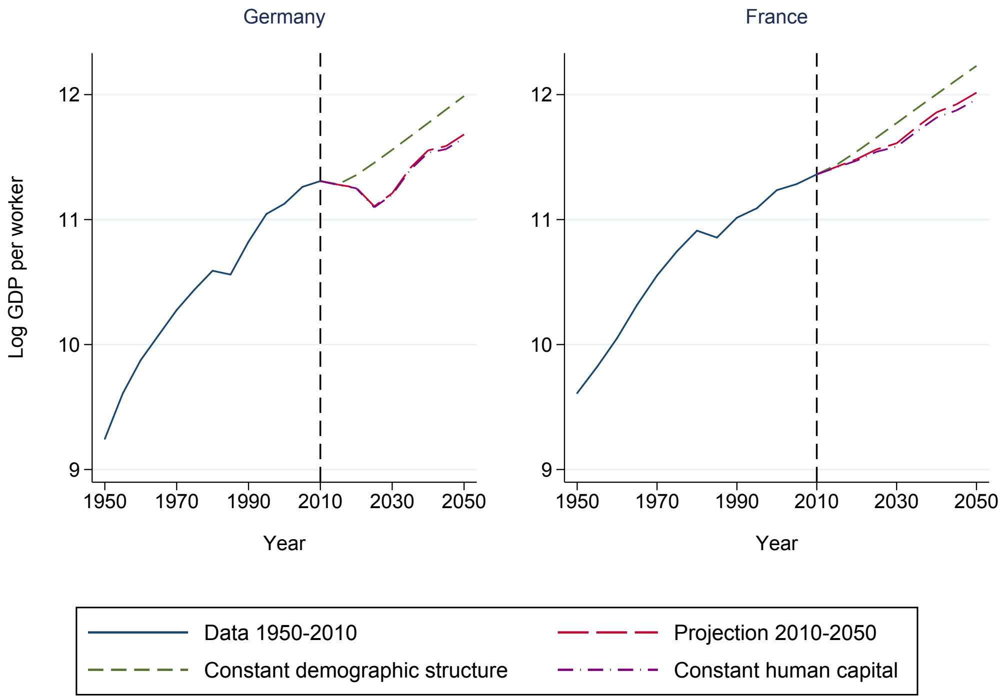
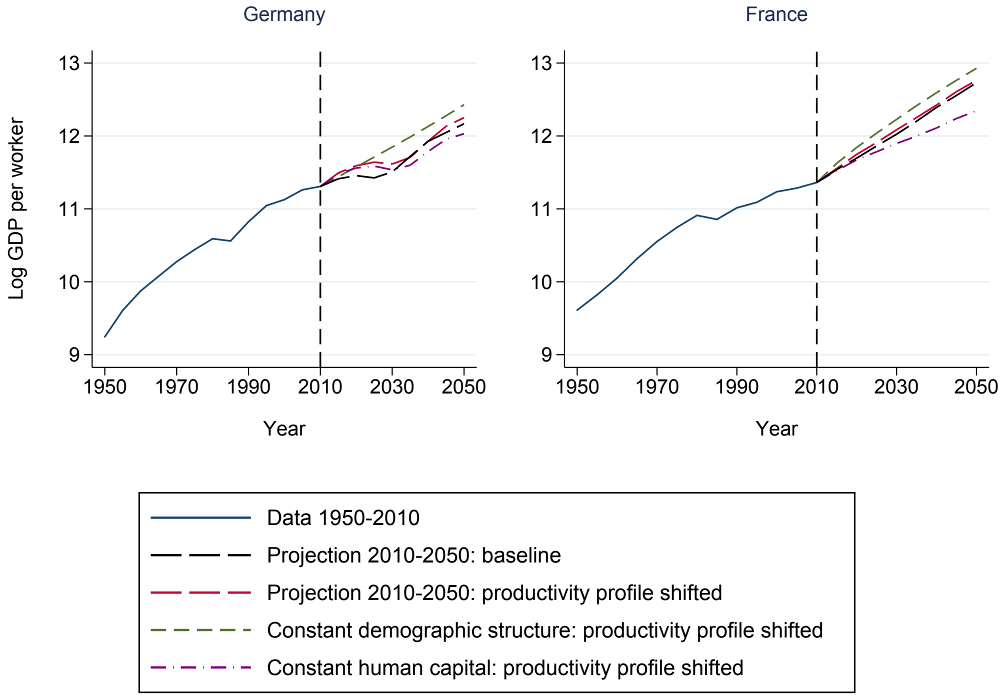
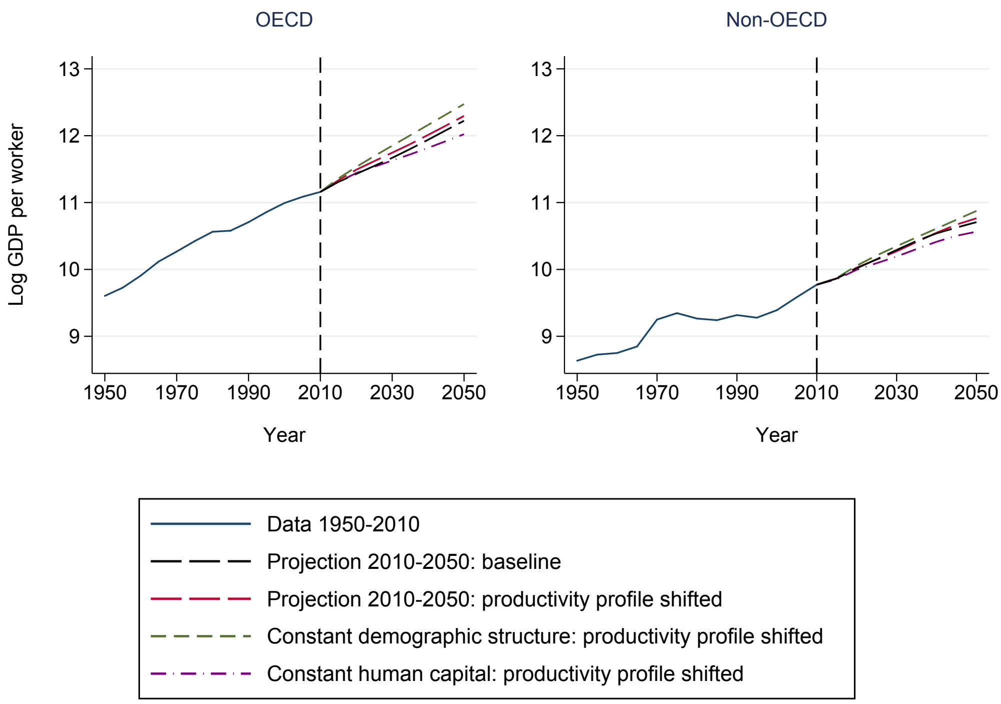


Figure A.28: Projections for Changing Labor Force Participation (Germany and France)

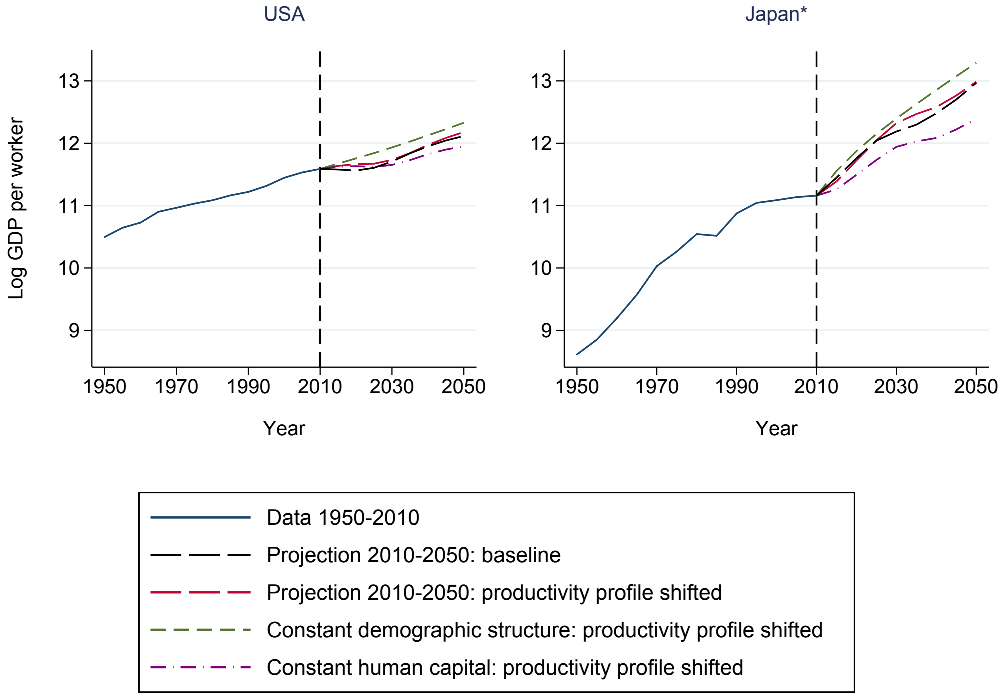


(a) Selected Countries: Germany and France

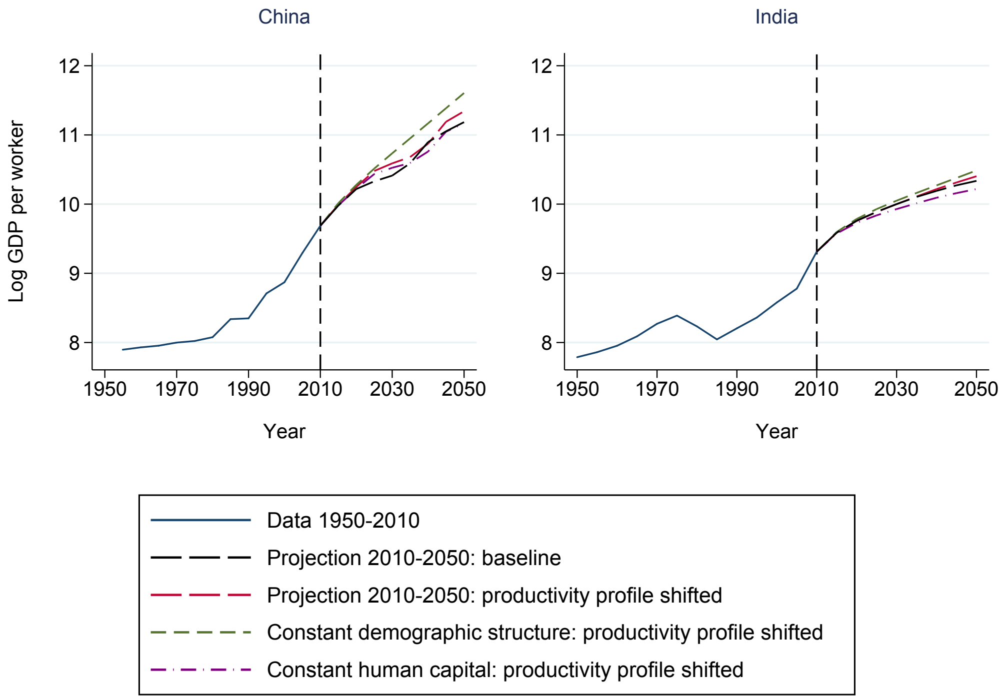


(b) Selected Regions: OECD and Non-OECD Countries

Figure A.29: Projections for Shifted Productivity Profile



(a) Selected Countries: USA and Japan



(b) Selected Countries: China and India

Figure A.30: Projections for Shifted Productivity Profile

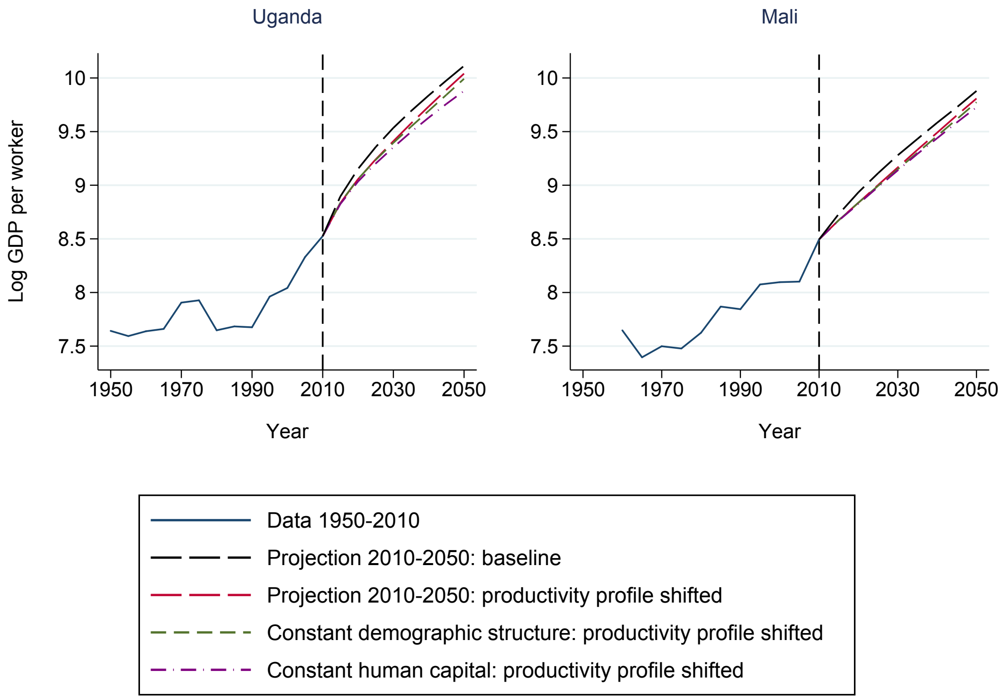
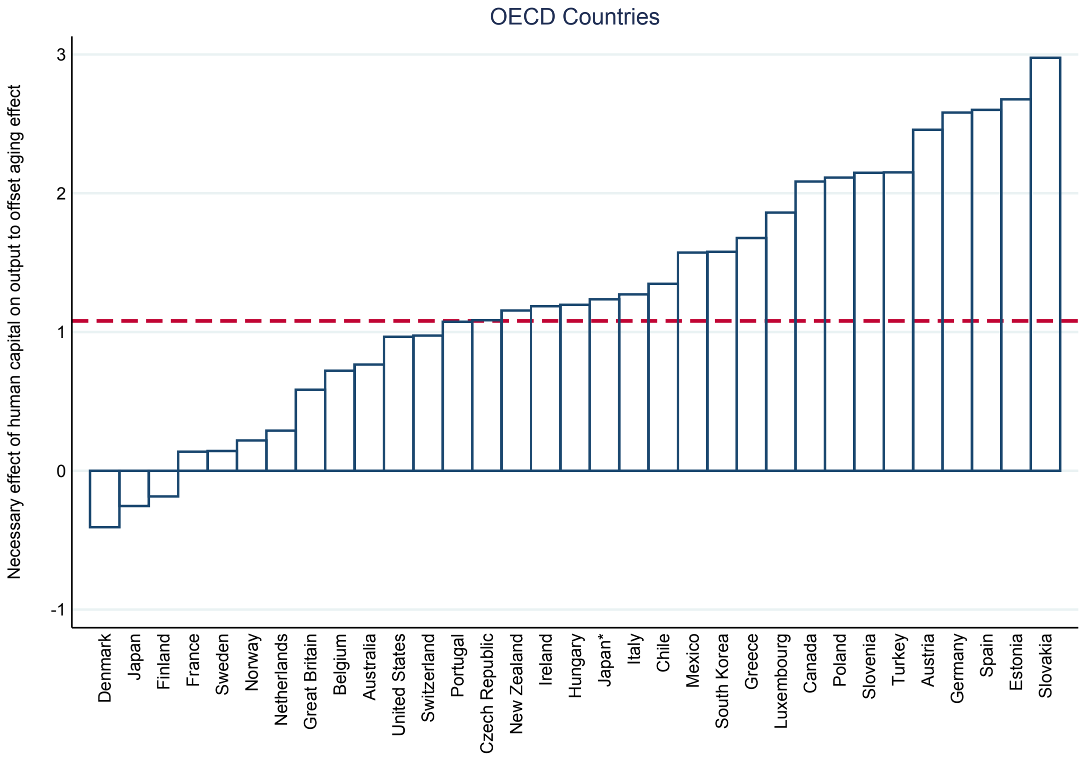
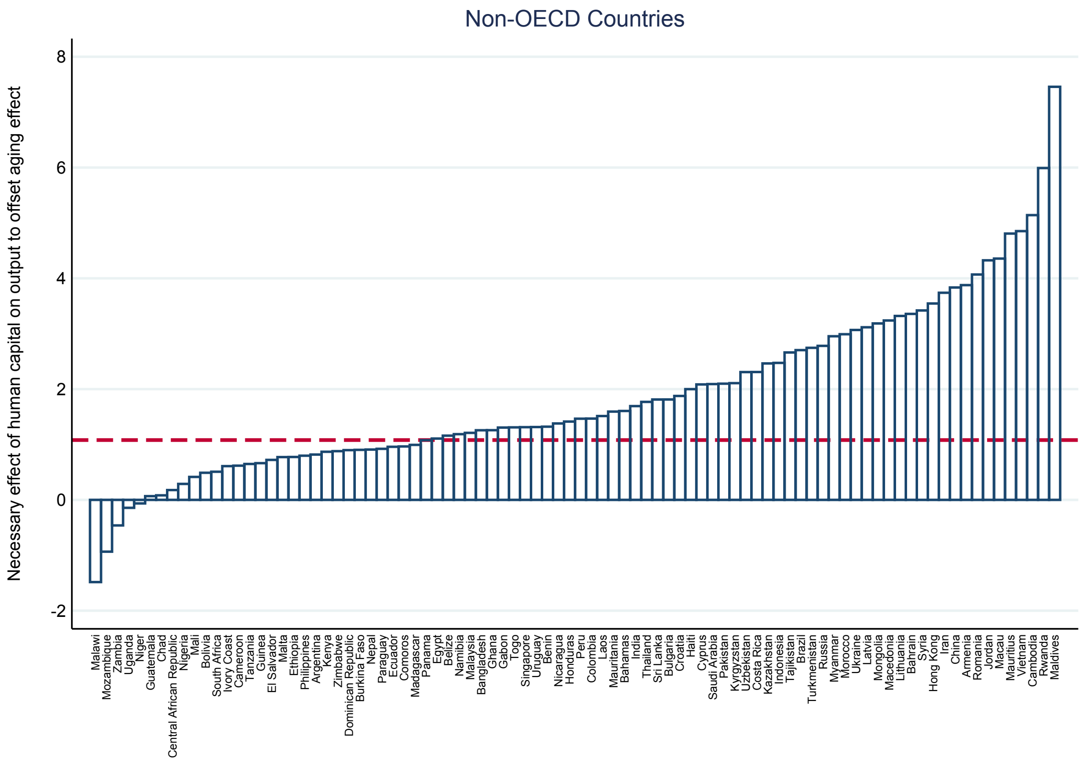


Figure A.31: Projections for Shifted Productivity Profile



(a) OECD Countries



(b) Non-OECD Countries

Figure A.32: How Large Must λ^h Be to Offset Aging Effect?

A.2 Additional Tables

Table A.1: Descriptive Statistics

	IIASA-VID sample ($n = 120$)					Barro-Lee sample ($n = 139$)				
	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max	Obs.
					GDP and physical capital					
Log output p.w.	9.59	1.12	6.92	12.12	1098	9.72	1.12	6.18	13.07	1259
Growth of output p.w.	0.11	0.22	-1.56	1.24	1098	0.10	0.23	-1.19	1.24	1259
Log output p.c.	8.61	1.18	5.95	11.45	1098	8.71	1.19	5.05	12.43	1259
Growth of output p.c.	0.12	0.23	-1.57	1.24	1098	0.11	0.23	-1.28	1.20	1259
Log capital p.w.	10.31	1.50	5.69	13.02	1098	10.45	1.45	5.69	13.87	1259
Growth of capital p.w.	0.17	0.25	-2.05	2.12	1098	0.17	0.25	-1.31	2.16	1259
					Share of age cohort in working-age population					
Total (in millions)	28.34	90.41	0.08	1013.06	1098	25.01	85.21	0.09	1018.27	1259
Change (in millions)	2.77	9.04	-4.18	93.45	1098	2.16	7.81	-4.13	96.55	1259
Share < 20	0.16	0.04	0.07	0.25	1098	0.16	0.04	0.06	0.25	1259
Share 20–24	0.14	0.03	0.07	0.21	1098	0.14	0.03	0.07	0.24	1259
Share 25–29	0.13	0.02	0.08	0.19	1098	0.13	0.02	0.08	0.25	1259
Share 30–34	0.11	0.01	0.07	0.19	1098	0.11	0.02	0.07	0.20	1259
Share 35–39	0.10	0.01	0.07	0.15	1098	0.10	0.01	0.05	0.18	1259
Share 40–44	0.09	0.01	0.05	0.15	1098	0.09	0.01	0.04	0.15	1259
Share 45–49	0.08	0.02	0.05	0.13	1098	0.08	0.02	0.04	0.13	1259
Share 50–54	0.07	0.02	0.03	0.12	1098	0.07	0.02	0.03	0.12	1259
Share 55–59	0.06	0.02	0.02	0.11	1098	0.06	0.02	0.01	0.11	1259
Share 60–64	0.05	0.02	0.01	0.11	1098	0.05	0.02	0.01	0.11	1259
Share 65+	0.04	0.02	0.01	0.09	1098	0.04	0.02	0.01	0.09	1259
					Share high-skills in working-age population					
Share high-skill	0.08	0.07	0.00	0.37	1098	0.08	0.09	0.00	0.58	1259
Change in share high-skill	0.01	0.01	-0.02	0.06	1098	0.01	0.02	-0.08	0.15	1259
Share < 20	0.01	0.03	0.00	0.25	1098	0.03	0.05	0.00	0.46	1259
Share 20–24	0.07	0.08	0.00	0.54	1098	0.12	0.13	0.00	0.93	1259
Share 25–29	0.12	0.10	0.00	0.56	1098	0.11	0.12	0.00	0.83	1259
Share 30–34	0.12	0.10	0.00	0.53	1098	0.10	0.11	0.00	0.66	1259
Share 35–39	0.11	0.09	0.00	0.50	1098	0.10	0.11	0.00	0.62	1259
Share 40–44	0.10	0.09	0.00	0.46	1098	0.09	0.10	0.00	0.62	1259
Share 45–49	0.08	0.08	0.00	0.44	1098	0.08	0.10	0.00	0.62	1259
Share 50–54	0.07	0.08	0.00	0.42	1098	0.07	0.09	0.00	0.60	1259
Share 55–59	0.06	0.07	0.00	0.39	1098	0.06	0.08	0.00	0.58	1259
Share 60–64	0.05	0.06	0.00	0.37	1098	0.05	0.07	0.00	0.57	1259
Share 65+	0.04	0.05	0.00	0.32	1098	0.04	0.06	0.00	0.54	1259
					Dependency ratio and life expectancy					
Dependency ratio	0.64	0.21	0.21	1.07	1086	0.64	0.21	0.16	1.09	1236
Life expectancy	64.79	11.20	23.73	82.98	1053	65.25	10.86	23.73	82.98	1198
					Share of age cohort in total labor force					
Total (in millions)	22.45	78.59	0.06	801.59	582	19.36	73.45	0.06	801.59	672
Change (in millions)	1.66	5.31	-6.22	58.49	479	1.44	4.96	-6.22	58.49	551
Share < 20	0.08	0.05	0.01	0.20	645	0.08	0.05	0.01	0.19	742
Share 20–24	0.13	0.03	0.05	0.22	645	0.13	0.03	0.05	0.22	742
Share 25–29	0.14	0.02	0.08	0.22	645	0.15	0.02	0.07	0.23	742
Share 30–34	0.13	0.02	0.09	0.23	645	0.14	0.02	0.09	0.23	742
Share 35–39	0.12	0.02	0.07	0.20	645	0.12	0.02	0.07	0.20	742
Share 40–44	0.11	0.02	0.05	0.17	645	0.11	0.02	0.05	0.17	742
Share 45–49	0.09	0.02	0.04	0.16	645	0.09	0.02	0.04	0.17	742
Share 50–54	0.07	0.02	0.03	0.14	645	0.07	0.02	0.03	0.14	742
Share 55–59	0.05	0.02	0.02	0.12	645	0.05	0.02	0.02	0.12	742
Share 60–64	0.03	0.01	0.00	0.09	645	0.03	0.01	0.00	0.09	742
Share 65+	0.03	0.02	0.00	0.10	645	0.03	0.02	0.00	0.10	742

Table A.2: Robustness: Differenced and Lagged Share of High-Skills

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) IIASA-VID sample						
Share < 20	-3.84***		-2.97**	-2.96**	-3.54***	-2.01	-2.73**
	(1.22)		(1.21)	(1.40)	(1.21)	(1.27)	(1.26)
Share 20–24	-2.37**		-1.87*	-1.86	-3.27**	-0.96	-2.40*
	(1.10)		(1.10)	(1.39)	(1.36)	(1.13)	(1.39)
Share 25–29	-3.56**		-3.25**	-3.28*	-2.88**	-2.22*	-1.80
	(1.42)		(1.39)	(1.90)	(1.47)	(1.34)	(1.41)
Share 30–34	-3.06**		-2.80**	-2.79*	-4.03***	-1.84	-3.35**
	(1.27)		(1.25)	(1.54)	(1.43)	(1.22)	(1.49)
Share 35–39	-4.01***		-3.72**	-3.64**	-2.99*	-3.84**	-2.58
	(1.44)		(1.43)	(1.75)	(1.57)	(1.55)	(1.68)
Share 40–44	-1.53		-1.32	-1.34	-1.81	-0.79	-1.82
	(1.32)		(1.30)	(1.62)	(1.42)	(1.30)	(1.44)
Share 45–49	-3.19**		-3.01**	-3.24*	-3.93**	-2.81**	-3.80**
	(1.43)		(1.40)	(1.94)	(1.55)	(1.40)	(1.53)
Share 55–59	-4.66**		-4.28**	-4.68**	-4.30**	-3.94**	-4.15**
	(1.85)		(1.81)	(2.17)	(1.83)	(1.81)	(1.83)
Share 60–64	-5.48***		-5.45***	-5.90***	-6.01***	-5.64***	-6.46***
	(1.37)		(1.35)	(1.76)	(1.51)	(1.30)	(1.48)
Share 65+	-3.06*		-3.28**	-3.39*	-4.26***	-3.31*	-3.97**
	(1.58)		(1.54)	(1.80)	(1.60)	(1.73)	(1.75)
Δ Share high-skill		1.89*	2.23*	1.80	1.82*	0.68	-0.12
		(1.07)	(1.16)	(1.23)	(1.10)	(2.31)	(2.29)
Share high-skill (t–1)		0.87**	0.94**	0.70**	0.72	2.80***	2.65***
		(0.35)	(0.43)	(0.33)	(0.45)	(0.84)	(0.86)
Cohort shares (p–value)	0.01		0.01	0.02	0.01	0.00	0.00
Skill shares (p–value)		0.01	0.02	0.03	0.06	0.00	0.01
First-stage F–statistic					13.5	37.1	8.7
Hansen test (p–value)					—	0.39	0.50
	(b) Barro-Lee sample						
Share < 20	-3.64***		-3.18***	-2.88**	-3.03*	-3.26**	-3.01*
	(1.17)		(1.20)	(1.36)	(1.55)	(1.30)	(1.61)
Share 20–24	-1.55		-1.24	-1.29	-1.72	-1.08	-1.58
	(1.14)		(1.13)	(1.61)	(1.63)	(1.17)	(1.63)
Share 25–29	-3.61**		-3.39**	-3.38**	-2.88	-3.16**	-2.51
	(1.41)		(1.41)	(1.57)	(1.84)	(1.41)	(1.81)
Share 30–34	-2.39*		-2.23*	-1.95	-1.93	-1.99	-1.55
	(1.29)		(1.29)	(1.65)	(1.81)	(1.30)	(1.83)
Share 35–39	-2.53*		-2.32*	-2.01	-2.71	-2.90*	-3.15
	(1.39)		(1.40)	(1.65)	(2.01)	(1.60)	(2.24)
Share 40–44	-2.24		-2.03	-2.13	-1.08	-1.58	-0.74
	(1.44)		(1.41)	(1.63)	(1.76)	(1.41)	(1.74)
Share 45–49	-1.83		-1.82	-1.72	-2.39	-1.66	-1.97
	(1.47)		(1.46)	(2.14)	(2.15)	(1.46)	(2.09)
Share 55–59	-3.75**		-3.51*	-3.88	-2.71	-3.36*	-2.52
	(1.85)		(1.84)	(2.49)	(2.40)	(1.83)	(2.38)
Share 60–64	-4.95***		-4.92***	-5.50***	-5.40***	-5.05***	-5.40***
	(1.40)		(1.39)	(1.90)	(1.67)	(1.36)	(1.62)
Share 65+	-1.44		-1.48	-1.48	-2.28	-1.43	-2.33
	(1.65)		(1.61)	(2.06)	(2.08)	(1.67)	(2.20)
Δ Share high-skill		0.69	0.51	0.37	0.56	-1.35	-0.93
		(0.50)	(0.51)	(0.40)	(0.50)	(2.08)	(2.29)
Share high-skill (t–1)		0.68***	0.55**	0.34	0.60**	0.78*	0.83*
		(0.22)	(0.25)	(0.23)	(0.25)	(0.40)	(0.42)
Cohort shares (p–value)	0.00		0.02	0.05	0.01	0.01	0.00
Skill shares (p–value)		0.01	0.07	0.23	0.05	0.02	0.03
First-stage F–statistic					12.5	4.0	1.0
Hansen test (p–value)					—	0.78	0.88

Notes: Panel (a) reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007), Panel (b) for data from Barro and Lee (2013). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms (coefficients unreported). Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.3: Robustness: Levels Model Without Lagged Dependent Variable

	Demography	Skills	Demography & Skills	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)
Share < 20	-5.56***		-3.77***	-4.46***	-2.80**	-3.32**
	(1.14)		(1.22)	(1.43)	(1.32)	(1.56)
Share 20–24	-1.65		-0.56	-2.35*	0.03	-1.75
	(1.11)		(1.14)	(1.35)	(1.19)	(1.35)
Share 25–29	-4.41***		-3.41***	-2.91**	-2.87**	-2.42*
	(1.19)		(1.19)	(1.41)	(1.19)	(1.43)
Share 30–34	-3.79***		-3.04***	-4.56***	-2.63**	-4.18***
	(1.20)		(1.15)	(1.34)	(1.16)	(1.35)
Share 35–39	-3.63***		-3.02**	-2.15	-2.69**	-1.72
	(1.28)		(1.30)	(1.36)	(1.30)	(1.34)
Share 40–44	-1.87		-1.45	-1.71	-1.22	-1.66
	(1.26)		(1.23)	(1.38)	(1.21)	(1.36)
Share 45–49	-2.12*		-1.94*	-2.49*	-1.85*	-2.47*
	(1.14)		(1.12)	(1.34)	(1.10)	(1.32)
Share 55–59	-2.02		-1.60	-0.89	-1.37	-0.83
	(1.30)		(1.29)	(1.37)	(1.27)	(1.34)
Share 60–64	-5.25***		-5.32***	-5.66***	-5.35***	-5.95***
	(1.32)		(1.28)	(1.39)	(1.26)	(1.39)
Share 65+	-2.40		-2.84*	-3.97**	-3.08*	-3.95**
	(1.72)		(1.68)	(1.69)	(1.69)	(1.75)
Share high-skill		2.64***	2.32***	1.80**	3.59***	3.24***
		(0.52)	(0.70)	(0.80)	(1.09)	(1.19)
Capital p.w.	0.56***	0.57***	0.56***	0.56***	0.57***	0.56***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Cohort shares (p–value)	0.00		0.00	0.00	0.00	0.00
Skill share (p–value)		0.00	0.00	0.02	0.00	0.01
First-stage F–statistic				13.2	27.4	4.5
Hansen test (p–value)				—	0.37	0.59
Countries	120	120	120	120	120	120
Observations	1,098	1,098	1,098	1,098	1,098	1,098
R^2	0.80	0.80	0.81	0.80	0.80	0.80

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Capital p.w., measured in logarithms, is included as control in all specifications. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5) (see Figure A.3), the lagged shares of high skills of cohorts at the edge of the working-age population in Column (6), and a combination of both in Column (7). First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.4: Robustness: Alternative Instrumentation of Human Capital (Levels)

	Inflow and Outflow (Baseline)		Inflow and Outflow GMM		Outflow Only	
	(1)	(2)	(3)	(4)	(5)	(6)
Share < 20	-2.09*	-2.53**	-1.79	-2.31**	-1.87	-2.31**
	(1.20)	(1.18)	(1.18)	(1.16)	(1.18)	(1.17)
Share 20–24	-1.16	-2.58*	-0.93	-2.25*	-1.01	-2.45*
	(1.11)	(1.36)	(1.09)	(1.32)	(1.09)	(1.33)
Share 25–29	-2.58*	-2.30	-2.27*	-2.05	-2.45*	-2.16
	(1.32)	(1.41)	(1.30)	(1.39)	(1.30)	(1.39)
Share 30–34	-2.33*	-3.61**	-2.37*	-3.72***	-2.24*	-3.56**
	(1.21)	(1.41)	(1.21)	(1.40)	(1.21)	(1.41)
Share 35–39	-3.33**	-2.49*	-2.95**	-1.93	-3.24**	-2.39
	(1.39)	(1.51)	(1.35)	(1.42)	(1.37)	(1.47)
Share 40–44	-1.12	-1.81	-0.66	-1.46	-1.07	-1.80
	(1.26)	(1.39)	(1.19)	(1.35)	(1.24)	(1.39)
Share 45–49	-2.92**	-3.89**	-2.71**	-3.64**	-2.89**	-3.89**
	(1.37)	(1.51)	(1.35)	(1.50)	(1.36)	(1.51)
Share 55–59	-4.00**	-4.16**	-3.59**	-3.79**	-3.91**	-4.11**
	(1.74)	(1.77)	(1.70)	(1.73)	(1.72)	(1.75)
Share 60–64	-5.53***	-6.33***	-5.29***	-6.10***	-5.53***	-6.38***
	(1.30)	(1.48)	(1.29)	(1.47)	(1.30)	(1.48)
Share 65+	-3.47**	-4.22**	-3.28**	-3.88**	-3.52**	-4.02**
	(1.56)	(1.68)	(1.55)	(1.65)	(1.58)	(1.68)
Share high-skill	2.45***	2.35***	2.56***	2.44***	2.77***	2.64***
	(0.76)	(0.77)	(0.75)	(0.77)	(0.80)	(0.82)
Output p.w. (t–1)	0.46***	0.46***	0.46***	0.45***	0.46***	0.45***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Capital p.w.	0.35***	0.34***	0.34***	0.34***	0.35***	0.35***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Cohort shares (p–value)	0.00	0.00	0.00	0.01	0.00	0.01
Skill share (p–value)	0.00	0.00	0.00	0.00	0.00	0.00
First-stage F–statistic	27.9	4.5	27.9	4.53	59.7	5.4
Hansen test (p–value)	0.25	0.28	0.25	0.28	—	—
Countries	120	120	120	120	120	120
Observations	1,098	1,098	1,098	1,098	1,098	1,098
R^2	0.86	0.86	0.86	0.86	0.86	0.86

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w., measured in logarithms, are included as controls in all specifications. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are the shifted shares of high-skills for the 15- to 19-year-olds (inflow) and the 65- to 69-year-olds (outflow) in all columns as well as the shifted age cohorts as in even columns. See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.5: Robustness: Barro-Lee Data (Levels)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share < 20	-3.64***		-3.18***	-2.88**	-3.03*	-2.86**	-2.60*
	(1.17)		(1.20)	(1.35)	(1.56)	(1.15)	(1.47)
Share 20–24	-1.55		-1.24	-1.29	-1.72	-1.02	-1.50
	(1.14)		(1.13)	(1.60)	(1.63)	(1.09)	(1.58)
Share 25–29	-3.61**		-3.39**	-3.37**	-2.89	-3.24**	-2.68
	(1.41)		(1.42)	(1.57)	(1.88)	(1.38)	(1.82)
Share 30–34	-2.39*		-2.23*	-1.96	-1.93	-2.11*	-1.65
	(1.29)		(1.28)	(1.64)	(1.81)	(1.25)	(1.79)
Share 35–39	-2.53*		-2.31*	-2.02	-2.71	-2.15	-2.49
	(1.39)		(1.38)	(1.64)	(2.01)	(1.35)	(1.95)
Share 40–44	-2.24		-2.04	-2.13	-1.08	-1.89	-0.92
	(1.44)		(1.42)	(1.62)	(1.78)	(1.36)	(1.75)
Share 45–49	-1.83		-1.82	-1.72	-2.40	-1.82	-2.21
	(1.47)		(1.46)	(2.13)	(2.19)	(1.43)	(2.14)
Share 55–59	-3.75**		-3.51*	-3.89	-2.71	-3.33*	-2.46
	(1.85)		(1.83)	(2.47)	(2.40)	(1.77)	(2.34)
Share 60–64	-4.95***		-4.91***	-5.49***	-5.41***	-4.89***	-5.41***
	(1.40)		(1.39)	(1.90)	(1.68)	(1.37)	(1.68)
Share 65–69	-1.44		-1.47	-1.47	-2.28	-1.50	-2.10
	(1.65)		(1.60)	(2.05)	(2.07)	(1.57)	(2.04)
Share high-skill		0.68***	0.55**	0.36*	0.60**	0.94**	0.96**
		(0.22)	(0.24)	(0.21)	(0.24)	(0.37)	(0.38)
Output p.w. (t–1)	0.53***	0.52***	0.51***	0.64***	0.51***	0.50***	0.50***
	(0.05)	(0.05)	(0.05)	(0.03)	(0.05)	(0.05)	(0.05)
Capital p.w.	0.27***	0.28***	0.28***	0.23***	0.29***	0.29***	0.29***
	(0.04)	(0.04)	(0.04)	(0.02)	(0.04)	(0.04)	(0.04)
Cohort shares (p–value)	0.00		0.02	0.05	0.00	0.02	0.00
Skill share (p–value)		0.00	0.02	0.09	0.02	0.01	0.01
First-stage F–statistic					12.6	27.8	5.1
Hansen test (p–value)					—	0.30	0.45
Countries	139	139	139	139	138	139	138
Observations	1,259	1,259	1,259	1,211	1,248	1,259	1,248
R^2	0.84	0.84	0.84		0.84	0.81	0.84

Notes: This table reports results for demographic and human capital data by Barro and Lee (2013). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged shares of high skills of cohorts at the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.6: Robustness: Barro-Lee Data (Differences)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ Share < 20	-3.16***		-2.66***	-1.66*	-4.13***	-2.84**	-4.66***
	(1.01)		(1.01)	(0.98)	(1.39)	(1.12)	(1.70)
Δ Share 20–24	-2.52***		-1.92**	-1.60*	-3.14**	-2.06*	-3.63**
	(0.93)		(0.91)	(0.92)	(1.44)	(1.10)	(1.68)
Δ Share 25–29	-3.16***		-2.70***	-1.89*	-4.08***	-2.82**	-4.58**
	(1.03)		(1.01)	(1.13)	(1.57)	(1.11)	(1.87)
Δ Share 30–34	-2.99***		-2.57**	-2.44**	-3.21**	-2.40**	-3.58**
	(1.08)		(1.08)	(1.20)	(1.56)	(1.10)	(1.78)
Δ Share 35–39	-3.38***		-2.89**	-2.43**	-3.80**	-3.23**	-4.27**
	(1.14)		(1.15)	(1.20)	(1.78)	(1.32)	(2.01)
Δ Share 40–44	-2.53**		-2.12*	-1.16	-2.18	-2.17*	-2.55*
	(1.11)		(1.10)	(0.97)	(1.36)	(1.15)	(1.49)
Δ Share 45–49	-2.41**		-2.20**	-0.88	-3.12**	-2.14**	-3.26**
	(1.09)		(1.08)	(0.92)	(1.56)	(1.08)	(1.62)
Δ Share 55–59	-2.08**		-1.96*	-0.97	-1.75	-1.88*	-1.86
	(1.02)		(1.02)	(1.00)	(1.17)	(1.03)	(1.24)
Δ Share 60–64	-5.06***		-5.07***	-3.10***	-4.83***	-5.20***	-4.85***
	(1.22)		(1.24)	(1.07)	(1.44)	(1.23)	(1.46)
Δ Share 65+	-5.56***		-5.31***	-1.89	-5.19***	-5.50***	-5.38**
	(1.66)		(1.65)	(1.47)	(1.97)	(1.72)	(2.13)
Δ Share high-skill		1.07**	0.94**	0.65*	0.91*	0.34	-0.30
		(0.48)	(0.47)	(0.33)	(0.47)	(1.89)	(2.18)
Share high-skill (t–1)		0.63***	0.61***	0.37***	0.61***	0.63***	0.52**
		(0.20)	(0.20)	(0.10)	(0.23)	(0.20)	(0.24)
Output p.w. (t–1)	-0.24***	-0.26***	-0.25***	-0.03***	-0.27***	-0.26***	-0.27***
	(0.03)	(0.03)	(0.03)	(0.01)	(0.03)	(0.03)	(0.03)
Δ Capital p.w.	0.36***	0.36***	0.37***	0.39***	0.36***	0.36***	0.36***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Cohort shares (p–value)	0.01		0.02	0.14	0.00	0.00	0.01
Skills shares (p–value)		0.00	0.00	0.00	0.01	0.00	0.04
First-stage F–statistic					6.8	6.5	0.8
AR(2) test (p–value)				0.35			
Hansen test (p–value)				0.25	—	0.07	0.09
Countries	139	139	139	139	138	139	138
Observations	1,259	1,259	1,259	1,120	1,200	1,211	1,200
R^2	0.37	0.36	0.38		0.39	0.39	0.38

Notes: This table reports results for demographic and human capital data by Barro and Lee (2013). The dependent variable is the log difference in output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Column (4) corrects for the dynamic-panel bias using the system GMM estimator by Arellano and Bover (1995) and Blundell and Bond (1998). The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged shares of high skills of cohorts at the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. For system GMM, also the p–values of the AR(2) test are reported. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.7: Robustness: Ten-Year Cohorts (Levels)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(a) IIASA-VID sample			
Share < 20	-4.19***		-2.89***	-2.23	-3.30***	-1.79	-2.13*
	(1.03)		(1.02)	(1.39)	(1.13)	(1.14)	(1.28)
Share 20–29	-2.26**		-1.56	-0.93	-1.90*	-0.97	-1.20
	(0.92)		(0.95)	(1.34)	(1.08)	(1.00)	(1.12)
Share 30–39	-2.82***		-2.43***	-1.87*	-2.57**	-2.11**	-2.31**
	(0.92)		(0.92)	(1.08)	(1.01)	(0.92)	(1.01)
Share 40–49	-2.17*		-1.93*	-1.29	-2.35*	-1.73	-2.07*
	(1.11)		(1.09)	(1.68)	(1.23)	(1.08)	(1.21)
Share 60+	-5.68***		-5.83***	-6.09***	-5.84***	-5.95***	-5.89***
	(1.38)		(1.34)	(1.94)	(1.34)	(1.34)	(1.35)
Share high-skill		1.81***	1.78***	1.35**	1.65***	3.26***	3.17***
		(0.47)	(0.52)	(0.58)	(0.53)	(0.94)	(0.97)
Output p.w. (t–1)	0.22***	0.17***	0.19***	0.38***	0.19***	0.16***	0.16***
	(0.04)	(0.04)	(0.04)	(0.08)	(0.04)	(0.05)	(0.04)
Capital p.w.	0.47***	0.49***	0.48***	0.42***	0.48***	0.50***	0.50***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Cohort shares (p–value)	0.00		0.00	0.02	0.00	0.00	0.00
Skill share (p–value)		0.00	0.00	0.02	0.00	0.00	0.00
First-stage F–statistic					121.8	30.0	8.0
Hansen test (p–value)					—	0.32	0.33
Countries	120	120	120	120	119	119	119
Observations	541	541	541	496	540	540	540
R^2	0.82	0.82	0.83		0.83	0.82	0.82
				(b) Barro-Lee sample			
Share < 20	-3.95***		-3.08**	-2.52*	-3.14***	-2.36*	-2.43**
	(1.24)		(1.29)	(1.51)	(1.20)	(1.33)	(1.20)
Share 20–29	-2.19**		-1.80*	-1.55	-2.01*	-1.47	-1.68
	(0.98)		(0.98)	(1.35)	(1.15)	(0.98)	(1.17)
Share 30–39	-2.51**		-2.22**	-1.75	-2.21**	-1.99*	-1.96*
	(1.07)		(1.08)	(1.28)	(1.08)	(1.09)	(1.11)
Share 40–49	-2.09*		-1.86	-1.51	-2.20	-1.68	-1.95
	(1.20)		(1.19)	(1.87)	(1.43)	(1.17)	(1.43)
Share 60+	-4.77***		-4.67***	-4.99**	-5.59***	-4.59***	-5.51***
	(1.58)		(1.54)	(2.33)	(1.46)	(1.52)	(1.44)
Share high-skill		1.20***	0.97***	0.73*	1.09***	1.76***	1.85***
		(0.32)	(0.36)	(0.42)	(0.37)	(0.65)	(0.68)
Output p.w. (t–1)	0.25***	0.21***	0.22***	0.36***	0.23***	0.19***	0.20***
	(0.05)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.05)
Capital p.w.	0.41***	0.44***	0.43***	0.38***	0.43***	0.44***	0.45***
	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
Cohort shares (p–value)	0.01		0.03	0.18	0.00	0.04	0.00
Skill share (p–value)		0.00	0.01	0.08	0.00	0.01	0.01
First-stage F–statistic					23.8	20.0	5.8
Hansen test (p–value)					—	0.24	0.39
Countries	139	139	139	139	137	138	137
Observations	621	621	621	573	615	620	615
R^2	0.78	0.78	0.78		0.78	0.78	0.78

Notes: Panel (a) reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007), Panel (b) for data from Barro and Lee (2013). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.8: Robustness: Ten-Year Cohorts (Differences)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(a) IIASA-VID sample			
Δ Share < 20	-3.87***		-2.67***	-3.01**	-2.31**	-4.42***	-3.60***
	(0.97)		(0.99)	(1.20)	(1.08)	(1.16)	(1.26)
Δ Share 20–29	-2.68***		-1.77**	-1.92**	-2.34**	-2.70***	-2.89***
	(0.80)		(0.84)	(0.94)	(1.06)	(0.88)	(1.05)
Δ Share 30–39	-3.00***		-2.53***	-2.95***	-2.82**	-2.49**	-2.19*
	(0.82)		(0.87)	(1.04)	(1.11)	(0.97)	(1.24)
Δ Share 40–49	-2.50***		-2.30***	-2.53***	-2.95***	-2.54***	-2.37**
	(0.78)		(0.75)	(0.81)	(0.91)	(0.84)	(0.98)
Δ Share 60+	-5.28***		-5.03***	-4.47***	-4.26***	-5.26***	-5.05***
	(0.94)		(0.93)	(0.93)	(1.05)	(0.96)	(1.18)
Δ Share high-skill		4.05***	3.44***	2.35***	3.95***	-2.04	-2.46
		(0.96)	(1.07)	(0.79)	(1.06)	(3.16)	(3.30)
Share high-skill (t–1)		1.10**	1.08**	0.51*	0.90	1.58***	1.51**
		(0.46)	(0.49)	(0.28)	(0.62)	(0.61)	(0.73)
Output p.w. (t–1)	-0.41***	-0.47***	-0.45***	-0.03***	-0.49***	-0.46***	-0.46***
	(0.05)	(0.05)	(0.05)	(0.01)	(0.05)	(0.05)	(0.05)
Δ Capital p.w.	0.40***	0.40***	0.41***	0.52***	0.41***	0.40***	0.40***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Cohort shares (p–value)	0.00		0.00	0.00	0.00	0.00	0.00
Skills shares (p–value)		0.00	0.00	0.00	0.00	0.03	0.11
First-stage F–statistic					49.1	31.8	8.1
AR(2) test (p–value)				0.13			
Hansen test (p–value)				0.75	—	0.14	0.11
Countries	120	120	120	119	119	119	119
Observations	541	541	541	421	495	495	495
R^2	0.58	0.56	0.60		0.60	0.60	0.60
				(b) Barro-Lee sample			
Δ Share < 20	-2.93***		-2.27**	-2.62**	-2.96**	-3.38***	-3.77**
	(1.03)		(1.03)	(1.09)	(1.38)	(1.30)	(1.67)
Δ Share 20–29	-2.02**		-1.47*	-1.31*	-2.22**	-2.15**	-2.88**
	(0.79)		(0.79)	(0.78)	(1.05)	(1.03)	(1.38)
Δ Share 30–39	-2.15**		-1.66*	-2.22**	-1.69	-2.17**	-2.39*
	(0.83)		(0.84)	(0.99)	(1.13)	(1.04)	(1.44)
Δ Share 40–49	-2.06**		-1.75**	-1.32	-1.43	-2.33**	-1.94
	(0.84)		(0.84)	(0.90)	(1.03)	(0.93)	(1.18)
Δ Share 60+	-4.74***		-4.61***	-3.64***	-4.72***	-4.72***	-4.89***
	(1.00)		(1.01)	(1.10)	(1.14)	(1.04)	(1.17)
Δ Share high-skill		1.28**	0.99*	0.88*	1.08**	-0.58	-1.05
		(0.56)	(0.53)	(0.50)	(0.54)	(1.75)	(2.07)
Share high-skill (t–1)		1.10***	1.06***	0.84***	1.15***	0.84	0.62
		(0.39)	(0.39)	(0.22)	(0.44)	(0.53)	(0.65)
Output p.w. (t–1)	-0.43***	-0.48***	-0.46***	-0.06***	-0.51***	-0.48***	-0.48***
	(0.05)	(0.05)	(0.04)	(0.01)	(0.05)	(0.05)	(0.05)
Δ Capital p.w.	0.32***	0.34***	0.34***	0.40***	0.31***	0.32***	0.31***
	(0.05)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.05)
Cohort shares (p–value)	0.00		0.00	0.01	0.00	0.00	0.00
Skills shares (p–value)		0.00	0.01	0.00	0.01	0.04	0.12
First-stage F–statistic					27.5	6.1	1.5
AR(2) test (p–value)				0.13			
Hansen test (p–value)				0.15	—	0.09	0.06
Countries	139	139	139	138	137	138	137
Observations	621	621	621	482	567	572	567
R^2	0.52	0.51	0.54		0.56	0.55	0.54

Notes: Panel (a) reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007), Panel (b) for data from Barro and Lee (2013). The dependent variable is the log difference in output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Column (4) corrects for the dynamic-panel bias using the system GMM estimator by Arellano and Bover (1995) and Blundell and Bond (1998). The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. For system GMM, also the p–values of the AR(2) test are reported. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.9: Robustness: Human Capital Granularity (Levels)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share < 20	-3.64***		-3.51***	-3.10***	-3.53**	-3.50***	-3.50**
	(1.17)		(1.21)	(1.20)	(1.54)	(1.22)	(1.52)
Share 20–24	-1.55		-1.36	-1.43	-1.93	-1.47	-2.19
	(1.14)		(1.10)	(1.40)	(1.61)	(1.11)	(1.68)
Share 25–29	-3.61**		-3.55**	-3.45**	-3.09	-3.57**	-3.14*
	(1.41)		(1.43)	(1.66)	(1.89)	(1.46)	(1.90)
Share 30–34	-2.39*		-2.42*	-2.11	-2.35	-2.66**	-2.59
	(1.29)		(1.29)	(1.39)	(1.77)	(1.28)	(1.85)
Share 35–39	-2.53*		-2.38*	-2.08	-2.75	-2.44*	-2.82
	(1.39)		(1.38)	(1.55)	(2.00)	(1.38)	(2.02)
Share 40–44	-2.24		-2.21	-2.25	-1.39	-2.32*	-1.68
	(1.44)		(1.42)	(1.51)	(1.75)	(1.39)	(1.75)
Share 45–49	-1.83		-2.01	-1.84	-2.76	-2.29	-3.03
	(1.47)		(1.45)	(2.12)	(2.20)	(1.43)	(2.27)
Share 55–59	-3.75**		-3.64**	-3.97*	-3.09	-3.76**	-3.27
	(1.85)		(1.83)	(2.11)	(2.40)	(1.82)	(2.43)
Share 60–64	-4.95***		-5.15***	-5.65***	-5.59***	-5.03***	-5.61***
	(1.40)		(1.43)	(1.74)	(1.71)	(1.48)	(1.73)
Share 65–69	-1.44		-1.77	-1.72	-2.76	-2.17	-3.08
	(1.65)		(1.60)	(1.68)	(2.09)	(1.58)	(2.15)
Years of schooling < 4		0.01	0.01	0.02	0.01	0.00	-0.00
		(0.05)	(0.05)	(0.04)	(0.05)	(0.08)	(0.08)
4–6 years of schooling		-0.00	0.01	0.02	0.01	0.01	0.00
		(0.03)	(0.04)	(0.03)	(0.03)	(0.05)	(0.05)
6–7 years of schooling		0.01	0.02	0.03	0.02	0.09	0.08
		(0.02)	(0.02)	(0.03)	(0.02)	(0.09)	(0.09)
8–10 years of schooling		0.05**	0.04*	0.03	0.04*	0.09	0.08
		(0.02)	(0.02)	(0.03)	(0.02)	(0.06)	(0.06)
Years of schooling > 10		0.13***	0.11***	0.08	0.12***	0.20**	0.20**
		(0.04)	(0.04)	(0.05)	(0.04)	(0.09)	(0.10)
Output p.w. (t–1)	0.53***	0.52***	0.51***	0.65***	0.51***	0.51***	0.51***
	(0.05)	(0.05)	(0.05)	(0.03)	(0.05)	(0.05)	(0.05)
Capital p.w.	0.27***	0.28***	0.28***	0.23***	0.28***	0.28***	0.28***
	(0.04)	(0.04)	(0.04)	(0.02)	(0.04)	(0.04)	(0.04)
Cohort shares (p–value)	0.00		0.01	0.00	0.00	0.08	0.05
Skill share (p–value)		0.00	0.01	0.49	0.01	0.07	0.07
First-stage F–statistic					12.7	10.5	3.4
Countries	139	139	139	139	138	139	138
Observations	1,259	1,259	1,259	1,211	1,248	1,259	1,248
R^2	0.84	0.84	0.84		0.84	0.84	0.84

Notes: This table reports results for demographic and human capital data by Barro and Lee (2013). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged years of schooling (6), and a combination of both in Column (7). See Panel (a) of Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.10: Robustness: Income Per Capita

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share < 20	-3.41**		-2.60*	-2.45	-4.56**	-1.84	-3.40*
	(1.41)		(1.40)	(1.51)	(1.89)	(1.42)	(1.84)
Share 20–24	-3.06***		-2.43**	-2.52*	-3.50**	-1.83*	-2.99**
	(1.11)		(1.09)	(1.43)	(1.49)	(1.10)	(1.47)
Share 25–29	-4.41***		-3.89***	-4.03**	-3.39**	-3.41**	-2.98*
	(1.42)		(1.39)	(1.95)	(1.59)	(1.35)	(1.54)
Share 30–34	-2.96**		-2.58**	-2.66*	-3.66***	-2.23*	-3.35**
	(1.22)		(1.19)	(1.59)	(1.41)	(1.18)	(1.40)
Share 35–39	-5.05***		-4.67***	-4.78***	-3.83**	-4.31***	-3.38**
	(1.38)		(1.38)	(1.79)	(1.58)	(1.35)	(1.53)
Share 40–44	-1.80		-1.56	-1.73	-1.91	-1.33	-1.82
	(1.31)		(1.28)	(1.67)	(1.42)	(1.25)	(1.40)
Share 45–49	-4.31***		-4.13***	-4.56**	-4.72***	-3.96***	-4.62***
	(1.50)		(1.47)	(2.01)	(1.78)	(1.43)	(1.75)
Share 55–59	-5.29***		-4.92***	-5.50**	-4.75***	-4.57***	-4.50**
	(1.70)		(1.65)	(2.25)	(1.83)	(1.59)	(1.77)
Share 60–64	-5.95***		-5.94***	-6.67***	-6.35***	-5.93***	-6.54***
	(1.41)		(1.37)	(1.83)	(1.54)	(1.33)	(1.52)
Share 65+	-3.61**		-3.80**	-4.20**	-4.47**	-3.97**	-4.54**
	(1.60)		(1.58)	(1.87)	(1.74)	(1.58)	(1.77)
Dependency Ratio	-0.79**	-1.19***	-0.80**	-0.71*	-1.32**	-0.81**	-1.25**
	(0.39)	(0.36)	(0.38)	(0.38)	(0.56)	(0.38)	(0.54)
Dependency Ratio (t–1)	-0.27	0.22	-0.25	-0.20	0.68	-0.24	0.53
	(0.51)	(0.36)	(0.51)	(0.49)	(0.81)	(0.51)	(0.77)
Share high-skill		1.03***	1.22***	0.91***	0.95**	2.37***	2.20***
		(0.39)	(0.40)	(0.34)	(0.43)	(0.74)	(0.75)
Output p.c. (t–1)	0.54***	0.50***	0.51***	0.64***	0.50***	0.49***	0.48***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
Capital p.w.	0.28***	0.28***	0.29***	0.24***	0.29***	0.30***	0.30***
	(0.04)	(0.04)	(0.04)	(0.02)	(0.04)	(0.05)	(0.05)
Cohort shares (p–value)	0.00		0.00	0.00	0.00	0.00	0.00
Skill share (p–value)		0.01	0.00	0.01	0.03	0.00	0.00
First-stage F–statistic					13.2	28.7	6.7
Hansen test (p–value)					—	0.19	0.22
Countries	120	120	120	120	120	120	120
Observations	1,098	1,098	1,098	1,053	1,098	1,098	1,098
R^2	0.88	0.88	0.89		0.88	0.88	0.88

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The dependent variable is log output per capita. All regressions include country-specific fixed and time effects. Lagged output p.c., capital p.w. and the (lagged) dependency ratio, measured in logarithms, are included as controls in all specifications. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged shares of high skills of cohorts at the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.11: Heterogeneity: Accounting for Human Capital Differences Between Cohorts

	Demography & Skills		Bias Correction		Skills Instrumented	
	IIASA-VID	Barro-Lee	IIASA-VID	Barro-Lee	IIASA-VID	Barro-Lee
	(1)	(2)	(3)	(4)	(5)	(6)
Share < 20	-2.59**	-3.16***	-2.55*	-2.64**	-1.91	-2.63
	(1.20)	(1.16)	(1.33)	(1.32)	(1.75)	(1.92)
Share 20–24	-1.04	-1.07	-1.00	-0.96	-1.99	-0.50
	(1.13)	(1.14)	(1.31)	(1.53)	(1.75)	(2.18)
Share 25–29	-2.07	-2.97**	-2.15	-2.73*	-1.53	-2.56
	(1.41)	(1.43)	(1.84)	(1.60)	(2.14)	(1.86)
Share 30–34	-2.02	-2.11*	-2.07	-1.67	-1.86	-1.80
	(1.24)	(1.26)	(1.53)	(1.59)	(1.62)	(1.46)
Share 35–39	-2.77*	-2.21	-2.72	-1.69	-2.30	-1.69
	(1.47)	(1.39)	(1.66)	(1.64)	(1.96)	(2.16)
Share 40–44	-0.58	-1.86	-0.62	-1.77	-0.33	-1.32
	(1.27)	(1.35)	(1.58)	(1.60)	(1.69)	(2.17)
Share 45–49	-2.36*	-1.77	-2.65	-1.51	-2.35	-1.74
	(1.40)	(1.50)	(1.88)	(2.09)	(2.13)	(1.65)
Share 55–59	-3.39*	-3.02*	-3.64*	-2.97	-3.07	-2.46
	(1.78)	(1.74)	(1.95)	(2.36)	(2.24)	(2.67)
Share 60–64	-4.78***	-4.49***	-5.24***	-4.85**	-4.99***	-3.94*
	(1.28)	(1.46)	(1.70)	(1.90)	(1.70)	(2.10)
Share 65+	-2.19	-1.18	-2.30	-0.97	-0.95	-0.72
	(1.56)	(1.57)	(1.80)	(2.11)	(3.18)	(1.88)
Share high-skill < 20	0.53	-0.27	0.39	-0.21	-1.76	-0.70
	(0.37)	(0.23)	(0.59)	(0.24)	(2.22)	(2.11)
Share high-skill 20–24	-0.52***	0.22*	-0.43*	0.22	-1.11	1.16
	(0.19)	(0.11)	(0.23)	(0.14)	(1.05)	(3.78)
Share high-skill 25–29	1.16***	0.07	1.01***	0.04	6.19	-0.58
	(0.35)	(0.12)	(0.37)	(0.17)	(6.92)	(2.69)
Share high-skill 30–34	0.46	0.03	0.31	-0.03	-9.38	-0.03
	(0.43)	(0.18)	(0.56)	(0.25)	(12.76)	(0.41)
Share high-skill 35–39	-1.04*	-0.31	-1.09	-0.33	6.13	-0.44
	(0.62)	(0.27)	(0.70)	(0.39)	(9.39)	(1.06)
Share high-skill 40–44	0.12	-0.17	0.25	-0.25	-1.51	-0.05
	(0.54)	(0.31)	(0.79)	(0.49)	(2.39)	(0.78)
Share high-skill 45–49	0.44	-0.39	0.40	-0.55	0.21	-0.35
	(0.59)	(0.34)	(0.82)	(0.52)	(0.74)	(0.76)
Share high-skill 50–54	-0.31	0.70	-0.39	0.77	-0.51	0.36
	(0.71)	(0.47)	(0.80)	(0.52)	(0.78)	(0.74)
Share high-skill 55–59	-0.61	0.55	-0.52	0.71	-1.64	0.74
	(0.62)	(0.56)	(0.75)	(0.61)	(1.68)	(0.97)
Share high-skill 60–64	0.86	-0.07	0.89	-0.13	1.07	0.97
	(0.59)	(0.44)	(0.84)	(0.57)	(0.84)	(3.59)
Share high-skill 65+	0.78	0.13	0.76	0.13	0.64	-0.76
	(0.55)	(0.41)	(0.75)	(0.53)	(0.86)	(2.76)
Output p.w. (t–1)	0.46***	0.52***	0.57***	0.66***	0.47***	0.52***
	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.05)
Capital p.w.	0.35***	0.29***	0.30***	0.23***	0.33***	0.28***
	(0.05)	(0.04)	(0.02)	(0.02)	(0.05)	(0.04)
Cohort shares (p–value)	0.01	0.04	0.02	0.24	0.04	0.22
Skill share (p–value)	0.00	0.11	0.00	0.14	0.02	0.35
First-stage F–statistic					0.1	0.0
Countries	120	139	120	139	120	139
Observations	1,098	1,259	1,053	1,211	1,098	1,259
R^2	0.87	0.84			0.82	0.84

Notes: Columns (1), (3), and (5) report results for demographic and human capital data from IIASA-VID (Lutz et al., 2007), Columns (2), (4), and (6) for data by Barro and Lee (2013). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Columns (3) and (4) correct for the dynamic-panel bias using the Bruno (2005) estimator. Instruments are the lagged shares of high skills of cohorts at the edge of the working-age population in Columns (5) and (6); see Panel (c) of Figure A.3 for an illustration. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) and the first-stage Kleibergen-Paap rk Wald F–statistic are reported. The IV specification is just-identified. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.12: Heterogeneity: Accounting for Human Capital Differences Between Cohorts

	Demography & Skills		Bias Correction		Skills Instrumented	
	IIASA-VID	Barro-Lee	IIASA-VID	Barro-Lee	IIASA-VID	Barro-Lee
	(1)	(2)	(3)	(4)	(5)	(6)
Share < 20	-3.33***	-3.41***	-3.08**	-2.86**	-3.34***	-3.39***
	(1.21)	(1.18)	(1.36)	(1.30)	(1.19)	(1.16)
Share 20–24	-1.83	-1.13	-1.65	-1.05	-1.83*	-1.11
	(1.13)	(1.15)	(1.37)	(1.55)	(1.11)	(1.12)
Share 25–29	-3.37**	-3.43**	-3.21*	-3.23**	-3.37**	-3.42**
	(1.40)	(1.43)	(1.85)	(1.53)	(1.37)	(1.40)
Share 30–34	-3.06**	-2.26*	-2.90*	-1.81	-3.06**	-2.26*
	(1.25)	(1.29)	(1.54)	(1.62)	(1.22)	(1.27)
Share 35–39	-3.68**	-2.56*	-3.47**	-2.06	-3.68***	-2.56*
	(1.44)	(1.35)	(1.71)	(1.59)	(1.41)	(1.32)
Share 40–44	-1.67	-2.00	-1.54	-1.93	-1.67	-1.99
	(1.30)	(1.44)	(1.56)	(1.60)	(1.28)	(1.41)
Share 45–49	-3.30**	-1.94	-3.41*	-1.71	-3.30**	-1.95
	(1.40)	(1.47)	(2.01)	(2.07)	(1.37)	(1.45)
Share 55–59	-4.50**	-3.26*	-4.59**	-3.42	-4.50**	-3.23*
	(1.80)	(1.82)	(2.05)	(2.26)	(1.77)	(1.78)
Share 60–64	-5.76***	-5.10***	-6.12***	-5.52***	-5.76***	-5.11***
	(1.38)	(1.41)	(1.74)	(1.83)	(1.36)	(1.38)
Share 65+	-3.74**	-1.69	-3.75**	-1.58	-3.74**	-1.70
	(1.55)	(1.63)	(1.85)	(2.07)	(1.52)	(1.60)
Share high-skill 50–54	0.78**	0.52***	0.69***	0.40**	0.78***	0.56***
	(0.31)	(0.19)	(0.23)	(0.20)	(0.30)	(0.21)
Rel. sh. high-skill < 20	-0.0001***	0.00	-0.00	0.00	-0.0001***	0.00
	(0.0000)	(0.00)	(0.00)	(0.01)	(0.0000)	(0.00)
Rel. sh. high-skill 20–24	-0.00	0.00	-0.00	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Rel. sh. high-skill 25–29	0.00	0.00	0.00	-0.00	0.00	0.00
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Rel. sh. high-skill 30–34	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Rel. sh. high-skill 35–39	0.00	-0.01	0.00	-0.01	0.00	-0.01
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
Rel. sh. high-skill 40–44	0.00	-0.01	-0.00	-0.01	0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Rel. sh. high-skill 45–49	0.02	0.01	0.02	0.01	0.02	0.01
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Rel. sh. high-skill 55–59	0.01	0.04	0.01	0.05	0.01	0.04
	(0.01)	(0.04)	(0.04)	(0.04)	(0.01)	(0.04)
Rel. sh. high-skill 60–64	0.01	-0.06	0.01	-0.06	0.01	-0.06
	(0.01)	(0.04)	(0.03)	(0.05)	(0.01)	(0.04)
Rel. sh. high-skill 65+	0.03**	0.03	0.03	0.04	0.03**	0.03
	(0.02)	(0.04)	(0.03)	(0.05)	(0.02)	(0.04)
Output p.w. (t–1)	0.48***	0.51***	0.59***	0.64***	0.48***	0.51***
	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)
Capital p.w.	0.35***	0.29***	0.31***	0.23***	0.35***	0.29***
	(0.05)	(0.04)	(0.02)	(0.03)	(0.05)	(0.04)
Cohort shares (p–value)	0.01	0.01	0.01	0.05	0.00	0.00
Skill share (p–value)	0.00	0.00	0.37	0.81	0.00	0.00
First-stage F–statistic					7.3e+5	2505.3
Countries	120	139	120	139	120	139
Observations	1,098	1,255	1,053	1,208	1,098	1,255
R^2	0.87	0.85			0.87	0.85

Notes: Columns (1), (3), and (5) report results for demographic and human capital data from IIASA-VID (Lutz et al., 2007), Columns (2), (4), and (6) for data by Barro and Lee (2013). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Columns (3) and (4) correct for the dynamic-panel bias using the Bruno (2005) estimator. Instruments are the lagged shares of high skills of cohorts at the edge of the working-age population in Columns (5) and (6); see Panel (c) of Figure A.3 for an illustration. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) and the first-stage Kleibergen-Paap rk Wald F–statistic are reported. The IV specification is just-identified. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.13: Robustness: Controlling for the Dependency Ratio (Levels)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(a) IIASA-VID sample			
Share < 20	-3.45***		-2.70**	-2.61*	-3.37***	-1.71	-2.21*
	(1.26)		(1.25)	(1.43)	(1.27)	(1.25)	(1.24)
Share 20–24	-2.65**		-2.10*	-2.21	-3.46**	-1.38	-2.79**
	(1.08)		(1.08)	(1.38)	(1.38)	(1.08)	(1.36)
Share 25–29	-3.56**		-3.13**	-3.22*	-2.82*	-2.56*	-2.34*
	(1.42)		(1.39)	(1.89)	(1.46)	(1.32)	(1.40)
Share 30–34	-3.03**		-2.71**	-2.74*	-3.96***	-2.29*	-3.59**
	(1.28)		(1.25)	(1.53)	(1.43)	(1.22)	(1.41)
Share 35–39	-4.07***		-3.77***	-3.72**	-3.03*	-3.37**	-2.54*
	(1.44)		(1.43)	(1.74)	(1.58)	(1.39)	(1.51)
Share 40–44	-1.66		-1.47	-1.55	-2.00	-1.22	-1.94
	(1.30)		(1.28)	(1.62)	(1.39)	(1.23)	(1.37)
Share 45–49	-3.26**		-3.14**	-3.37*	-4.02***	-2.98**	-3.93***
	(1.43)		(1.41)	(1.95)	(1.55)	(1.37)	(1.51)
Share 55–59	-4.63**		-4.35**	-4.75**	-4.39**	-3.96**	-4.14**
	(1.85)		(1.81)	(2.18)	(1.83)	(1.74)	(1.77)
Share 60–64	-5.44***		-5.46***	-5.93***	-6.04***	-5.49***	-6.32***
	(1.36)		(1.33)	(1.76)	(1.50)	(1.29)	(1.48)
Share 65+	-2.90*		-3.09*	-3.25*	-4.16**	-3.34**	-4.20**
	(1.60)		(1.57)	(1.79)	(1.64)	(1.57)	(1.69)
Dependency ratio	-0.13	-0.16	-0.12	-0.11	-0.08	-0.11	-0.09
	(0.13)	(0.11)	(0.13)	(0.10)	(0.13)	(0.13)	(0.12)
Share high-skill		0.89**	1.08***	0.85***	0.84**	2.51***	2.40***
		(0.34)	(0.40)	(0.32)	(0.41)	(0.76)	(0.77)
Cohort shares (p–value)	0.01		0.01	0.04	0.01	0.00	0.00
Skill share (p–value)		0.01	0.01	0.01	0.04	0.00	0.00
First-stage F–statistic					12.6	28.6	4.7
Hansen test (p–value)					—	0.20	0.25
				(b) Barro-Lee sample			
Share < 20	-3.63***		-3.15**	-2.71**	-3.08*	-2.87**	-2.65*
	(1.22)		(1.24)	(1.36)	(1.58)	(1.20)	(1.50)
Share 20–24	-1.64		-1.32	-1.56	-1.67	-1.13	-1.46
	(1.15)		(1.12)	(1.58)	(1.67)	(1.08)	(1.63)
Share 25–29	-3.61**		-3.38**	-3.39**	-2.88	-3.24**	-2.68
	(1.42)		(1.42)	(1.59)	(1.88)	(1.38)	(1.83)
Share 30–34	-2.46*		-2.31*	-2.08	-1.93	-2.23*	-1.65
	(1.30)		(1.30)	(1.76)	(1.81)	(1.27)	(1.79)
Share 35–39	-2.63*		-2.41*	-2.18	-2.69	-2.28*	-2.49
	(1.38)		(1.38)	(1.75)	(2.00)	(1.34)	(1.95)
Share 40–44	-2.24		-2.02	-2.19	-1.06	-1.90	-0.90
	(1.43)		(1.40)	(1.63)	(1.78)	(1.34)	(1.75)
Share 45–49	-1.76		-1.77	-1.69	-2.39	-1.78	-2.21
	(1.48)		(1.46)	(2.31)	(2.19)	(1.44)	(2.15)
Share 55–59	-3.70**		-3.43*	-3.82*	-2.71	-3.27*	-2.46
	(1.84)		(1.82)	(2.23)	(2.40)	(1.76)	(2.33)
Share 60–64	-5.16***		-5.18***	-5.79***	-5.41***	-5.20***	-5.41***
	(1.39)		(1.37)	(1.58)	(1.68)	(1.34)	(1.68)
Share 65+	-1.47		-1.48	-1.53	-2.29	-1.49	-2.12
	(1.66)		(1.61)	(2.13)	(2.07)	(1.58)	(2.04)
Dependency ratio	-0.01	-0.09	-0.01	-0.01	0.02	-0.01	0.01
	(0.14)	(0.12)	(0.14)	(0.10)	(0.14)	(0.13)	(0.14)
Share high-skill		0.71***	0.61**	0.43*	0.60**	0.97**	0.96**
		(0.22)	(0.24)	(0.24)	(0.25)	(0.38)	(0.39)
Cohort shares (p–value)	0.00		0.01	0.00	0.00	0.00	0.00
Skill share (p–value)		0.00	0.01	0.07	0.02	0.01	0.01
First-stage F–statistic					10.5	29.5	5.5
Hansen test (p–value)					—	0.47	0.47

Notes: Panel (a) reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007), Panel (b) for data from Barro and Lee (2013). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms (coefficients unreported). Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.14: Robustness: Population Scale Effects (in Logarithms)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share < 20	-3.54***		-3.06***	-3.04**	-3.62***	-2.61**	-3.06***
	(1.13)		(1.16)	(1.40)	(1.14)	(1.21)	(1.15)
Share 20–24	-1.65		-1.39	-1.45	-2.71**	-1.14	-2.47*
	(1.05)		(1.06)	(1.39)	(1.33)	(1.08)	(1.34)
Share 25–29	-3.08**		-2.86**	-2.97	-2.64*	-2.64**	-2.46*
	(1.38)		(1.37)	(1.90)	(1.44)	(1.33)	(1.40)
Share 30–34	-2.58**		-2.43**	-2.43	-3.50**	-2.28*	-3.39**
	(1.18)		(1.18)	(1.54)	(1.38)	(1.17)	(1.37)
Share 35–39	-3.60**		-3.45**	-3.48**	-2.80*	-3.31**	-2.58*
	(1.38)		(1.38)	(1.75)	(1.52)	(1.36)	(1.50)
Share 40–44	-0.95		-0.90	-0.95	-1.48	-0.86	-1.52
	(1.23)		(1.22)	(1.63)	(1.34)	(1.19)	(1.33)
Share 45–49	-2.67*		-2.66*	-2.95	-3.76**	-2.65**	-3.76**
	(1.36)		(1.35)	(1.96)	(1.48)	(1.32)	(1.47)
Share 55–59	-4.00**		-3.90**	-4.36**	-4.11**	-3.79**	-4.05**
	(1.66)		(1.66)	(2.19)	(1.69)	(1.64)	(1.68)
Share 60–64	-5.35***		-5.38***	-5.87***	-6.10***	-5.41***	-6.25***
	(1.33)		(1.32)	(1.77)	(1.47)	(1.29)	(1.47)
Share 65+	-4.52***		-4.45***	-4.41**	-4.62***	-4.37***	-4.67***
	(1.49)		(1.48)	(1.83)	(1.59)	(1.47)	(1.60)
Working-Age Population	-0.20***	-0.12**	-0.17***	-0.14***	-0.17***	-0.14***	-0.14***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Share high-skill		0.70**	0.72*	0.55	0.51	1.42*	1.34*
		(0.32)	(0.39)	(0.35)	(0.40)	(0.75)	(0.75)
Output p.w. (t–1)	0.46***	0.46***	0.46***	0.56***	0.46***	0.45***	0.45***
	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)
Capital p.w.	0.33***	0.33***	0.34***	0.29***	0.33***	0.34***	0.34***
	(0.05)	(0.05)	(0.05)	(0.02)	(0.05)	(0.05)	(0.05)
Cohort shares (p–value)	0.00		0.00	0.00	0.00	0.00	0.00
Skill share (p–value)		0.03	0.07	0.11	0.21	0.06	0.08
First-stage F–statistic					13.2	26.4	4.2
Hansen test (p–value)					—	0.18	0.16
Countries	120	120	120	120	120	120	120
Observations	1,098	1,098	1,098	1,053	1,098	1,098	1,098
R^2	0.87	0.86	0.87		0.87	0.87	0.87

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w., capital p.w. and the working-age population are measured in logarithms. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.15: Robustness: Population Scale Effects (in Absolute Values)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share < 20	-3.75***		-2.80**	-2.75**	-3.41***	-1.69	-2.16*
	(1.23)		(1.22)	(1.38)	(1.22)	(1.22)	(1.20)
Share 20–24	-2.41**		-1.80	-1.78	-3.28**	-1.09	-2.61*
	(1.10)		(1.09)	(1.39)	(1.37)	(1.10)	(1.35)
Share 25–29	-3.48**		-2.94**	-2.97	-2.60*	-2.31*	-2.03
	(1.42)		(1.39)	(1.89)	(1.46)	(1.32)	(1.39)
Share 30–34	-3.04**		-2.66**	-2.64*	-3.90***	-2.21*	-3.51**
	(1.28)		(1.25)	(1.54)	(1.44)	(1.21)	(1.42)
Share 35–39	-4.01***		-3.65**	-3.58**	-2.96*	-3.24**	-2.46
	(1.45)		(1.44)	(1.75)	(1.58)	(1.40)	(1.52)
Share 40–44	-1.54		-1.33	-1.33	-1.86	-1.08	-1.80
	(1.32)		(1.29)	(1.62)	(1.41)	(1.26)	(1.40)
Share 45–49	-3.15**		-3.00**	-3.23*	-3.88**	-2.81**	-3.75**
	(1.43)		(1.40)	(1.96)	(1.55)	(1.36)	(1.51)
Share 55–59	-4.71**		-4.39**	-4.79**	-4.38**	-4.01**	-4.13**
	(1.85)		(1.81)	(2.20)	(1.83)	(1.74)	(1.77)
Share 60–64	-5.49***		-5.52***	-6.00***	-6.01***	-5.56***	-6.28***
	(1.37)		(1.34)	(1.77)	(1.50)	(1.30)	(1.48)
Share 65+	-2.95*		-3.12**	-3.22*	-4.11**	-3.32**	-4.06**
	(1.59)		(1.57)	(1.81)	(1.63)	(1.59)	(1.70)
Working-Age Population	0.04*	0.05**	0.06***	0.06***	0.05**	0.07***	0.07***
$(\widehat{\beta}, se(\widehat{\beta}) \times 100)$	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Share high-skill		1.08***	1.28***	1.06***	0.99**	2.77***	2.62***
		(0.34)	(0.41)	(0.34)	(0.43)	(0.79)	(0.81)
Output p.w. (t–1)	0.51***	0.48***	0.49***	0.59***	0.48***	0.46***	0.46***
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)
Capital p.w.	0.32***	0.33***	0.33***	0.29***	0.33***	0.34***	0.34***
	(0.05)	(0.04)	(0.05)	(0.02)	(0.05)	(0.05)	(0.05)
Cohort shares (p–value)	0.01		0.01	0.01	0.01	0.00	0.00
Skill share (p–value)		0.00	0.00	0.00	0.02	0.00	0.00
First-stage F–statistic					13.2	26.6	4.4
Hansen test (p–value)					—	0.31	0.37
Countries	120	120	120	120	120	120	120
Observations	1,098	1,098	1,098	1,053	1,098	1,098	1,098
R^2	0.87	0.86	0.87		0.87	0.86	0.86

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. The working-age population is measured in millions. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.16: Robustness: Controlling for Life Expectancy (Levels)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(a) IIASA-VID sample			
Share < 20	-3.70***		-2.71**	-2.69*	-3.49***	-1.29	-1.91
	(1.23)		(1.22)	(1.38)	(1.22)	(1.26)	(1.23)
Share 20–24	-2.36**		-1.72	-1.83	-3.22**	-0.80	-2.33*
	(1.12)		(1.11)	(1.36)	(1.39)	(1.13)	(1.39)
Share 25–29	-3.51**		-2.95**	-3.06	-2.65*	-2.13	-1.82
	(1.45)		(1.41)	(1.87)	(1.49)	(1.35)	(1.42)
Share 30–34	-2.52**		-2.11*	-2.16	-3.61**	-1.52	-3.21**
	(1.25)		(1.22)	(1.53)	(1.46)	(1.19)	(1.43)
Share 35–39	-4.59***		-4.30***	-4.28**	-3.29*	-3.88**	-2.61
	(1.63)		(1.60)	(1.72)	(1.77)	(1.55)	(1.68)
Share 40–44	-1.17		-0.93	-1.03	-1.64	-0.59	-1.55
	(1.31)		(1.29)	(1.60)	(1.41)	(1.29)	(1.43)
Share 45–49	-2.95**		-2.80**	-3.07	-3.77**	-2.59*	-3.58**
	(1.43)		(1.40)	(1.93)	(1.52)	(1.37)	(1.48)
Share 55–59	-4.45**		-4.11**	-4.57**	-4.16**	-3.62**	-3.79**
	(1.87)		(1.84)	(2.16)	(1.82)	(1.78)	(1.77)
Share 60–64	-5.37***		-5.41***	-5.88***	-5.99***	-5.47***	-6.32***
	(1.35)		(1.32)	(1.75)	(1.48)	(1.27)	(1.46)
Share 65+	-2.97*		-3.04*	-3.18*	-4.22**	-3.14*	-4.11**
	(1.77)		(1.74)	(1.80)	(1.70)	(1.79)	(1.81)
Life expectancy	0.00	0.01	0.00	0.005**	0.00	0.01	0.01
	(0.00)	(0.00)	(0.00)	(0.002)	(0.00)	(0.00)	(0.00)
Share high-skill		1.14***	1.32***	1.13***	0.98**	3.21***	3.04***
		(0.35)	(0.42)	(0.32)	(0.44)	(0.86)	(0.89)
Cohort shares (p–value)	0.01		0.01	0.01	0.01	0.00	0.00
Skill share (p–value)		0.00	0.00	0.00	0.03	0.00	0.00
First-stage F–statistic					12.2	30.4	5.1
Hansen test (p–value)					—	0.47	0.57
				(b) Barro-Lee sample			
Share < 20	-3.21***		-2.61**	-2.35*	-2.41	-2.24*	-1.84
	(1.21)		(1.24)	(1.31)	(1.52)	(1.20)	(1.44)
Share 20–24	-1.12		-0.66	-0.90	-1.24	-0.38	-0.99
	(1.14)		(1.11)	(1.54)	(1.61)	(1.07)	(1.56)
Share 25–29	-3.10**		-2.78*	-2.87*	-2.13	-2.57*	-1.86
	(1.42)		(1.44)	(1.54)	(1.81)	(1.40)	(1.76)
Share 30–34	-1.78		-1.62	-1.48	-0.88	-1.52	-0.52
	(1.29)		(1.27)	(1.72)	(1.84)	(1.24)	(1.83)
Share 35–39	-2.87*		-2.54*	-2.42	-2.85	-2.34	-2.59
	(1.53)		(1.51)	(1.70)	(2.23)	(1.47)	(2.16)
Share 40–44	-1.34		-1.07	-1.28	-0.08	-0.90	0.12
	(1.39)		(1.36)	(1.57)	(1.69)	(1.31)	(1.68)
Share 45–49	-1.15		-1.12	-1.08	-1.37	-1.11	-1.15
	(1.51)		(1.49)	(2.24)	(2.13)	(1.47)	(2.09)
Share 55–59	-2.87		-2.49	-3.00	-1.91	-2.26	-1.62
	(1.83)		(1.82)	(2.19)	(2.35)	(1.76)	(2.30)
Share 60–64	-4.53***		-4.50***	-5.08***	-4.52***	-4.49***	-4.57***
	(1.35)		(1.33)	(1.55)	(1.61)	(1.30)	(1.61)
Share 65+	-0.74		-0.75	-0.77	-1.65	-0.76	-1.44
	(1.72)		(1.67)	(2.08)	(2.12)	(1.64)	(2.09)
Life expectancy	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Share high-skill		0.86***	0.74***	0.60***	0.69***	1.20***	1.17***
		(0.23)	(0.26)	(0.23)	(0.26)	(0.44)	(0.44)
Cohort shares (p–value)	0.00		0.01	0.00	0.00	0.01	0.00
Skill share (p–value)		0.00	0.00	0.01	0.01	0.01	0.01
First-stage F–statistic					15.5	28.1	4.9
Hansen test (p–value)					—	0.15	0.15

Notes: Panel (a) reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007), Panel (b) for data from Barro and Lee (2013). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms (coefficients unreported). Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.17: Robustness: Average Years of Schooling (Levels)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share < 20	-3.64***		-3.66***	-3.23**	-3.70**	-3.83***	-3.97***
	(1.17)		(1.18)	(1.37)	(1.55)	(1.16)	(1.53)
Share 20–24	-1.55		-1.54	-1.53	-2.04	-1.47	-1.83
	(1.14)		(1.14)	(1.60)	(1.65)	(1.11)	(1.68)
Share 25–29	-3.61**		-3.62**	-3.52**	-3.15*	-3.65***	-3.36*
	(1.41)		(1.41)	(1.58)	(1.89)	(1.36)	(1.92)
Share 30–34	-2.39*		-2.37*	-2.05	-2.30	-2.28*	-2.24
	(1.29)		(1.28)	(1.64)	(1.77)	(1.26)	(1.69)
Share 35–39	-2.53*		-2.54*	-2.21	-3.01	-2.55*	-2.86
	(1.39)		(1.39)	(1.64)	(2.03)	(1.37)	(2.03)
Share 40–44	-2.24		-2.24	-2.29	-1.25	-2.23	-1.44
	(1.44)		(1.44)	(1.62)	(1.79)	(1.42)	(1.81)
Share 45–49	-1.83		-1.81	-1.70	-2.60	-1.70	-2.53
	(1.47)		(1.47)	(2.13)	(2.20)	(1.48)	(2.20)
Share 55–59	-3.75**		-3.78**	-4.11*	-3.05	-3.91**	-3.39
	(1.85)		(1.85)	(2.47)	(2.42)	(1.82)	(2.43)
Share 60–64	-4.95***		-4.96***	-5.53***	-5.43***	-5.07***	-5.69***
	(1.40)		(1.40)	(1.90)	(1.70)	(1.36)	(1.68)
Share 65+	-1.44		-1.40	-1.39	-2.47	-1.20	-1.95
	(1.65)		(1.63)	(2.04)	(2.10)	(1.63)	(2.07)
Years of schooling		0.00	-0.01	-0.01	-0.01	-0.04**	-0.04**
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Output p.w. (t–1)	0.53***	0.54***	0.53***	0.66***	0.53***	0.53***	0.53***
	(0.05)	(0.05)	(0.05)	(0.03)	(0.05)	(0.05)	(0.05)
Capital p.w.	0.27***	0.27***	0.28***	0.22***	0.28***	0.28***	0.28***
	(0.04)	(0.04)	(0.04)	(0.02)	(0.04)	(0.04)	(0.04)
Cohort shares (p–value)	0.00		0.00	0.03	0.00	0.00	0.00
Skill share (p–value)		1.00	0.68	0.26	0.67	0.04	0.04
First-stage F–statistic					12.3	103.5	18.9
Hansen test (p–value)					—	0.00	0.00
Countries	139	139	139	139	138	139	138
Observations	1,259	1,259	1,259	1,211	1,248	1,259	1,248
R^2	0.84	0.84	0.84		0.84	0.84	0.83

Notes: This table reports results for demographic and human capital data by Barro and Lee (2013). Human capital is proxied by average years of schooling. The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.18: Heterogeneity: Sample Split OECD and Non-OECD Countries (Levels)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(a) OECD Countries			
Share < 20	0.35		0.60	0.44	0.72	1.28	1.43
	(0.99)		(1.07)	(1.23)	(1.05)	(1.09)	(1.16)
Share 20–24	-0.98		-0.89	-0.66	-1.89*	-0.63	-1.73*
	(0.89)		(0.89)	(1.35)	(0.97)	(0.82)	(0.98)
Share 25–29	-0.40		-0.26	-0.15	1.05	0.12	1.16
	(1.22)		(1.24)	(1.56)	(1.24)	(1.21)	(1.29)
Share 30–34	-0.41		-0.32	-0.04	-1.18	-0.07	-1.02
	(1.12)		(1.12)	(1.33)	(1.32)	(1.03)	(1.26)
Share 35–39	0.49		0.62	0.87	1.01	1.00	1.25
	(0.97)		(0.98)	(1.54)	(0.98)	(0.94)	(1.05)
Share 40–44	-0.77		-0.74	-0.70	-0.55	-0.66	-0.62
	(1.10)		(1.10)	(1.31)	(1.01)	(1.08)	(1.02)
Share 45–49	-0.54		-0.52	-0.47	-0.78	-0.47	-0.90
	(1.12)		(1.13)	(1.84)	(0.99)	(1.10)	(1.01)
Share 55–59	0.02		0.04	-0.04	-0.00	0.09	-0.08
	(1.59)		(1.62)	(1.69)	(1.47)	(1.59)	(1.52)
Share 60–64	-1.57		-1.65	-1.75	-1.79	-1.87*	-2.18**
	(1.14)		(1.12)	(1.52)	(1.11)	(1.02)	(1.06)
Share 65+	-1.76		-1.71	-1.87	-2.22	-1.60	-2.14
	(1.29)		(1.34)	(1.53)	(1.41)	(1.45)	(1.61)
Share high-skill		0.21	0.41	0.51*	0.42	1.51**	1.60***
		(0.34)	(0.32)	(0.29)	(0.28)	(0.61)	(0.59)
Cohort shares (p–value)	0.45		0.24	0.03	0.01	0.02	0.00
Skill share (p–value)		0.54	0.21	0.09	0.14	0.01	0.01
First-stage F–statistic					22.8	9.9	1.8
Hansen test (p–value)					—	0.33	0.31
Countries	32	32	32	32	32	32	32
Observations	341	341	341	318	341	341	341
				(b) Non-OECD countries			
Share < 20	-5.44***		-4.88***	-5.15**	-6.88***	-3.03*	-4.47**
	(1.85)		(1.84)	(2.54)	(2.07)	(1.78)	(1.89)
Share 20–24	-3.14*		-2.85	-3.24	-5.58**	-1.88	-4.42*
	(1.77)		(1.77)	(2.88)	(2.40)	(1.76)	(2.27)
Share 25–29	-4.35*		-4.11*	-4.57*	-5.80**	-3.29*	-4.54**
	(2.23)		(2.18)	(2.77)	(2.54)	(1.99)	(2.24)
Share 30–34	-4.48**		-4.30**	-4.70*	-6.55**	-3.67**	-6.21**
	(1.88)		(1.85)	(2.84)	(2.67)	(1.78)	(2.59)
Share 35–39	-7.06***		-6.89***	-7.04**	-6.87**	-6.30***	-5.29*
	(2.48)		(2.47)	(2.96)	(3.19)	(2.32)	(3.01)
Share 40–44	-1.71		-1.66	-1.86	-3.80	-1.48	-3.88
	(2.46)		(2.42)	(3.09)	(3.06)	(2.26)	(2.98)
Share 45–49	-4.12		-4.18	-5.02	-8.14**	-4.37*	-7.74**
	(2.55)		(2.53)	(3.98)	(3.54)	(2.43)	(3.25)
Share 55–59	-7.70***		-7.56***	-8.24*	-9.02***	-7.11***	-8.37***
	(2.83)		(2.78)	(4.58)	(3.16)	(2.61)	(3.02)
Share 60–64	-7.66***		-7.67***	-8.96***	-9.60***	-7.73***	-10.11***
	(2.34)		(2.27)	(2.97)	(2.92)	(2.04)	(2.67)
Share 65+	-4.47		-4.41	-4.92	-6.23*	-4.21	-5.95*
	(2.81)		(2.76)	(3.58)	(3.21)	(2.63)	(3.12)
Share high-skill		0.96*	0.92	0.60	0.50	4.03***	3.98***
		(0.50)	(0.68)	(0.67)	(0.73)	(1.45)	(1.49)
Cohort shares (p–value)	0.01		0.01	0.00	0.00	0.00	0.03
Skill share (p–value)		0.06	0.18	0.37	0.49	0.01	0.01
First-stage F–statistic					6.3	28.2	4.6
Hansen test (p–value)					—	0.83	0.90
Countries	88	88	88	88	88	88	88
Observations	757	757	757	735	757	757	757

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). Panel (a) reports results for OECD countries, Panel (b) for non-OECD countries. The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms (coefficients unreported). Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged shares of high skills of cohorts at the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.19: Heterogeneity: Sample Split Before and After 1990 (Levels)

	Demography & Skills		Bias Correction		Skills Instrumented	
	–1990	1990+	–1990	1990+	–1990	1990+
	(1)	(2)	(3)	(4)	(5)	(6)
Share < 20	2.49	-5.21***	2.83	-4.40**	2.50*	-1.38
	(1.51)	(1.68)	(2.15)	(1.96)	(1.50)	(2.19)
Share 20–24	1.64	-3.60**	1.60	-2.92	1.63	-0.27
	(1.30)	(1.58)	(2.08)	(2.32)	(1.27)	(1.89)
Share 25–29	0.76	-5.05***	0.88	-4.65**	0.76	-2.82
	(1.82)	(1.79)	(2.64)	(2.35)	(1.78)	(1.82)
Share 30–34	1.30	-4.72***	1.59	-4.40	1.29	-1.72
	(1.68)	(1.64)	(2.15)	(2.70)	(1.69)	(2.16)
Share 35–39	0.66	-5.31***	1.44	-5.00*	0.65	-3.89**
	(1.60)	(1.74)	(2.69)	(3.02)	(1.61)	(1.76)
Share 40–44	2.83**	-5.45***	3.20	-5.47**	2.83**	-1.95
	(1.37)	(1.76)	(2.49)	(2.65)	(1.34)	(2.11)
Share 45–49	1.36	-5.56***	0.52	-4.85	1.38	-4.61**
	(1.52)	(1.87)	(2.95)	(3.34)	(1.49)	(1.99)
Share 55–59	1.49	-5.60***	1.68	-5.80**	1.49	-4.18**
	(1.85)	(2.10)	(3.30)	(2.88)	(1.81)	(2.02)
Share 60–64	-1.56	-7.34***	-1.56	-8.29***	-1.55	-8.21***
	(1.62)	(1.89)	(2.57)	(2.22)	(1.57)	(1.84)
Share 65+	1.66	-7.56***	3.00	-7.53**	1.65	-8.40***
	(1.67)	(2.12)	(2.97)	(3.25)	(1.60)	(2.54)
Share high-skill	1.10	0.56	0.28	0.42	1.17	6.36**
	(1.01)	(0.81)	(0.90)	(0.92)	(2.03)	(2.53)
Output p.w. (t–1)	0.54***	0.23***	0.80***	0.42***	0.54***	0.20***
	(0.10)	(0.06)	(0.07)	(0.08)	(0.10)	(0.06)
Capital p.w.	0.28***	0.46***	0.19***	0.41***	0.28***	0.46***
	(0.07)	(0.06)	(0.04)	(0.04)	(0.07)	(0.06)
Cohort shares (p–value)	0.15	0.02	0.27	0.00	0.27	0.00
Skill share (p–value)	0.28	0.49	0.75	0.65	0.56	0.01
First-stage F–statistic					19.0	14.3
Hansen test (p–value)					0.29	0.52
Countries	103	120	103	120	85	120
Observations	516	582	471	479	498	582
R^2	0.82	0.74			0.82	0.70

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The sample is split in periods before 1990 (1955–1985) and after 1990 (1990–2010). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Columns (3) and (4) correct for the dynamic-panel bias using the Bruno (2005) estimator. Instruments are the lagged shares of high skills of cohorts at the edge of the working-age population in Columns (5) and (6); see Panel (c) of Figure A.3 for an illustration. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) and the first-stage Kleibergen-Paap rk Wald F–statistic are reported. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.20: Heterogeneity: Sample Split Before and After 1990 (Levels, Barro-Lee Data)

	Demography & Skills		Bias Correction		Skills Instrumented	
	–1990	1990+	–1990	1990+	–1990	1990+
	(1)	(2)	(3)	(4)	(5)	(6)
Share < 20	0.05	-3.78*	0.74	-2.74	0.05	-2.56
	(1.38)	(2.15)	(2.31)	(1.70)	(1.35)	(2.20)
Share 20–24	1.76	-2.24	1.01	-1.23	1.77	-0.74
	(1.31)	(1.78)	(2.36)	(1.92)	(1.30)	(1.89)
Share 25–29	-1.84	-4.49**	-1.52	-3.68*	-1.83	-4.05*
	(1.61)	(2.03)	(2.74)	(2.08)	(1.57)	(2.08)
Share 30–34	-0.49	-2.52	-0.14	-1.75	-0.47	-1.40
	(1.54)	(2.21)	(2.28)	(2.31)	(1.56)	(2.28)
Share 35–39	-0.13	-1.95	0.13	-1.43	-0.12	-1.14
	(1.27)	(2.17)	(2.50)	(2.46)	(1.24)	(2.19)
Share 40–44	1.01	-5.42**	1.57	-5.32**	1.00	-4.26**
	(1.33)	(2.19)	(2.45)	(2.60)	(1.30)	(2.17)
Share 45–49	-0.14	-2.58	-0.45	-0.99	-0.14	-2.33
	(1.34)	(2.72)	(3.30)	(2.86)	(1.32)	(2.71)
Share 55–59	1.61	-4.39	1.34	-4.22	1.61	-3.56
	(1.54)	(2.73)	(3.51)	(2.67)	(1.51)	(2.78)
Share 60–64	-1.12	-7.16***	-1.76	-7.98***	-1.10	-6.92***
	(1.60)	(1.81)	(2.79)	(1.83)	(1.56)	(1.81)
Share 65+	0.54	-3.98	1.76	-3.68	0.56	-3.38
	(1.89)	(2.54)	(3.14)	(3.03)	(1.83)	(2.68)
Share high-skill	0.18	0.69	-0.01	0.59	0.13	2.11**
	(0.38)	(0.55)	(0.44)	(0.46)	(0.51)	(1.05)
Output p.w. (t–1)	0.54***	0.23***	0.84***	0.48***	0.54***	0.21***
	(0.08)	(0.06)	(0.06)	(0.06)	(0.08)	(0.05)
Capital p.w.	0.31***	0.35***	0.18***	0.29***	0.31***	0.36***
	(0.06)	(0.05)	(0.04)	(0.03)	(0.06)	(0.05)
Cohort shares (p–value)	0.36	0.01	0.32	0.00	0.36	0.01
Skill share (p–value)	0.64	0.21	0.99	0.20	0.80	0.04
First-stage F–statistic					8.3	15.6
Hansen test (p–value)					0.31	0.18
Countries	122	139	122	139	97	139
Observations	582	677	534	555	557	677
R^2	0.81	0.67			0.81	0.66

Notes: This table reports results for demographic and human capital data by Barro and Lee (2013). The sample is split in periods before 1990 (1955–1985) and after 1990 (1990–2010). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms. Columns (3) and (4) correct for the dynamic-panel bias using the Bruno (2005) estimator. Instruments are the lagged shares of high skills of cohorts at the edge of the working-age population in Columns (5) and (6); see Panel (c) of Figure A.3 for an illustration. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) and the first-stage Kleibergen-Paap rk Wald F–statistic are reported. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.21: Robustness: Old Versus Young Populations (Levels)

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Old populations (above median of the young-age dependency ratio)						
Share < 20	-4.26***		-3.92**	-4.12***	-3.35**	-3.57**	-2.86*
	(1.54)		(1.55)	(1.58)	(1.38)	(1.60)	(1.47)
Share 20–24	-1.93		-1.76	-1.97	-3.62**	-1.58	-3.51**
	(1.27)		(1.28)	(1.91)	(1.47)	(1.27)	(1.51)
Share 25–29	-4.79**		-4.64**	-5.08***	-2.24	-4.48***	-2.03
	(1.82)		(1.79)	(1.82)	(1.53)	(1.74)	(1.54)
Share 30–34	-3.32**		-3.22**	-3.23*	-5.28***	-3.12**	-5.18***
	(1.32)		(1.31)	(1.72)	(1.72)	(1.28)	(1.72)
Share 35–39	-4.51**		-4.39**	-4.54**	-2.27	-4.27**	-2.11
	(1.81)		(1.80)	(2.16)	(1.78)	(1.75)	(1.75)
Share 40–44	-3.27**		-3.23**	-3.59**	-3.85***	-3.19**	-3.86***
	(1.32)		(1.31)	(1.53)	(1.28)	(1.27)	(1.29)
Share 45–49	-5.28***		-5.25***	-5.65**	-4.43***	-5.22***	-4.43***
	(1.80)		(1.80)	(2.33)	(1.65)	(1.74)	(1.64)
Share 55–59	-5.47**		-5.38**	-6.19**	-5.38***	-5.29***	-5.32***
	(2.10)		(2.09)	(2.42)	(1.92)	(2.02)	(1.91)
Share 60–64	-6.22***		-6.29***	-7.11***	-6.32***	-6.37***	-6.54***
	(1.59)		(1.57)	(1.63)	(1.67)	(1.51)	(1.68)
Share 65+	-6.13***		-6.13***	-6.29***	-5.90***	-6.13***	-5.95***
	(1.54)		(1.54)	(1.83)	(1.54)	(1.52)	(1.56)
Share high-skill		0.15	0.54	0.47	0.41	1.10	1.13
		(0.38)	(0.42)	(0.43)	(0.42)	(0.95)	(0.92)
Cohort shares (p–value)	0.02		0.01	0.00	0.00	0.00	0.00
Skill share (p–value)		0.69	0.21	0.28	0.33	0.25	0.22
First-stage F–statistic					6.5	16.2	2.6
Hansen test (p–value)					—	0.24	0.18
Observations	549	549	549	525	543	543	543
R^2	0.93	0.92	0.93		0.93	0.93	0.93
	(b) Young populations (below median of the young-age dependency ratio)						
Share < 20	3.42		3.67	2.88	-0.49	4.45	0.65
	(3.11)		(3.18)	(3.80)	(5.27)	(3.01)	(5.35)
Share 20–24	3.42		3.46	2.80	-1.90	3.58	-1.47
	(3.09)		(3.09)	(4.01)	(5.33)	(3.02)	(5.41)
Share 25–29	5.23		5.24	4.24	1.70	5.26*	1.88
	(3.26)		(3.25)	(3.99)	(6.07)	(3.13)	(6.14)
Share 30–34	0.50		0.53	-0.20	-2.22	0.63	-1.88
	(3.94)		(3.96)	(4.62)	(5.29)	(3.93)	(5.41)
Share 35–39	4.45		4.45	4.25	1.23	4.46	1.41
	(2.84)		(2.84)	(4.24)	(6.31)	(2.80)	(6.41)
Share 40–44	7.66*		7.60*	6.65	4.25	7.39*	3.99
	(3.90)		(3.89)	(4.54)	(6.70)	(3.77)	(6.71)
Share 45–49	3.66		3.55	2.95	-6.74	3.22	-6.71
	(3.94)		(3.95)	(5.89)	(7.94)	(3.91)	(8.04)
Share 55–59	-0.84		-0.67	-1.37	-7.54	-0.17	-7.34
	(5.22)		(5.19)	(7.04)	(10.07)	(4.95)	(10.20)
Share 60–64	1.76		2.04	1.21	-2.17	2.87	-1.75
	(6.15)		(6.13)	(7.46)	(11.31)	(5.81)	(11.26)
Share 65+	9.94		9.91	9.45	14.80	9.84*	17.70
	(6.26)		(6.22)	(6.43)	(13.30)	(5.97)	(13.16)
Share high-skill		0.88	0.77	0.48	0.63	3.09**	3.05*
		(0.92)	(1.07)	(1.11)	(1.11)	(1.44)	(1.61)
Cohort shares (p–value)	0.07		0.07	0.24	0.66	0.11	0.66
Skill share (p–value)		0.34	0.47	0.66	0.57	0.03	0.06
First-stage F–statistic					0.9	32.2	0.9
Hansen test (p–value)					—	0.50	0.53
Observations	549	549	549	528	545	545	545
R^2	0.70	0.69	0.70		0.69	0.70	0.69

Notes: This table reports results for demographic and human capital data by IIASA-VID (Lutz et al., 2007). The sample has been split with respect to the young-age dependency ratio. Panel (a) reports results for observations for which the young-age dependency ratio is above the median, Panel (b) for observations below the median. The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w. are measured in logarithms (coefficients unreported). Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged share of high skills of the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A.22: Robustness: Labor Force Shares

	Demography	Skills	Demography & Skills	Bias Correction	Demography Instrumented	Skills Instrumented	Both Instrumented
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Labor Force Share < 20	-4.96***		-4.91***	-4.53***	-5.64***	-3.14*	-1.85
	(1.34)		(1.37)	(1.44)	(1.74)	(1.63)	(2.39)
Labor Force Share 20–24	-4.16***		-4.07***	-3.76***	-6.00***	-0.85	-3.96**
	(1.27)		(1.32)	(1.43)	(1.41)	(1.77)	(1.66)
Labor Force Share 25–29	-4.03***		-3.96***	-4.25**	-4.67***	-1.45	-2.10
	(1.34)		(1.38)	(1.70)	(1.60)	(1.65)	(2.10)
Labor Force Share 30–34	-4.28***		-4.21***	-4.05**	-7.13***	-1.79	-5.36**
	(1.51)		(1.52)	(1.69)	(1.83)	(1.81)	(2.31)
Labor Force Share 35–39	-5.54***		-5.48***	-5.49***	-5.68***	-3.50**	-4.43**
	(1.49)		(1.48)	(1.89)	(1.90)	(1.65)	(2.15)
Labor Force Share 40–44	-4.25***		-4.22***	-4.30**	-5.63***	-3.01**	-5.01***
	(1.38)		(1.39)	(1.67)	(1.64)	(1.42)	(1.93)
Labor Force Share 45–49	-4.60**		-4.57**	-4.62*	-6.59***	-3.70*	-5.49**
	(1.77)		(1.77)	(2.40)	(1.81)	(1.94)	(2.14)
Labor Force Share 55–59	-5.78***		-5.77***	-5.79**	-6.27***	-5.35***	-6.78***
	(1.88)		(1.89)	(2.52)	(1.92)	(2.01)	(2.11)
Labor Force Share 60–64	-6.93***		-6.92***	-7.68***	-9.28***	-6.71**	-10.16***
	(2.43)		(2.44)	(2.24)	(2.71)	(3.04)	(3.41)
Labor Force Share 65+	-2.00		-1.98	-1.28	-6.80	-1.57	-11.76*
	(2.12)		(2.13)	(1.98)	(4.74)	(3.16)	(6.69)
Share high-skill		0.85**	0.19	-0.29	0.08	6.99***	6.29**
		(0.33)	(0.66)	(0.59)	(0.69)	(2.54)	(2.93)
Output p.w. (t–1)	0.25***	0.50***	0.25***	0.41***	0.23***	0.18***	0.19***
	(0.05)	(0.05)	(0.05)	(0.04)	(0.06)	(0.06)	(0.07)
Capital p.w.	0.46***	0.33***	0.46***	0.42***	0.47***	0.46***	0.48***
	(0.06)	(0.04)	(0.06)	(0.03)	(0.06)	(0.06)	(0.06)
Cohort shares (p–value)	0.03		0.04	0.00	0.00	0.12	0.02
Skill share (p–value)		0.01	0.77	0.62	0.91	0.01	0.03
First-stage F–statistic					2.8	10.9	1.1
Hansen test (p–value)					—	0.60	0.73
Countries	120	120	120	120	120	120	120
Observations	645	1,098	645	645	645	645	645
R^2	0.76	0.86	0.76		0.75	0.68	0.65

Notes: This table reports results for demographic data by International Labour Organization (2011) and human capital data by IIASA-VID (Lutz et al., 2007). The dependent variable is log output per worker. All regressions include country-specific fixed and time effects. Lagged output p.w. and capital p.w., measured in logarithms, are included as controls in all specifications. Column (4) corrects for the dynamic-panel bias using the Bruno (2005) estimator. The p–value for a Wald test whether coefficients of workforce shares (proxied by the working-age population) or high-skill shares are jointly different from zero are reported. Instruments are shifted age cohorts in Column (5), the lagged shares of high skills of cohorts at the edge of the working-age population in Column (6), and a combination of both in Column (7). See Figure A.3 for an illustration. First-stage F–statistic reports the first-stage Kleibergen-Paap rk Wald F–statistic. Hansen test p–values refer to the robust overidentifying restriction test. Standard errors are clustered at the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix B

Appendix to Chapter 3

B.1 Additional Tables

Table B.1: Robustness: Return of Health and Education to Productivity

	No Controls	Adding Controls	Fixed Effects	Adding Gini	Lagged Controls	IV $\Delta(h_{i,t})$	IV $\Delta(s_{i,t})$	IV $\Delta(h_{i,t})$, $\Delta(s_{i,t})$
Regressors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln(y_{i,t-1})$	-0.051***	-0.15***	-0.33***	-0.18***	-0.20***	-0.14***	-0.14***	-0.14***
	(0.0095)	(0.017)	(0.052)	(0.024)	(0.022)	(0.020)	(0.016)	(0.020)
$\Delta \ln(k_{i,t})$	0.51***	0.35***	0.27**	0.23***	0.48***	0.35***	0.35***	0.35***
	(0.052)	(0.063)	(0.11)	(0.066)	(0.075)	(0.063)	(0.068)	(0.068)
$\Delta(h_{i,t})$	1.22***	0.63**	0.66**	0.76**	0.72**	0.97	0.60**	0.56
	(0.30)	(0.30)	(0.27)	(0.35)	(0.28)	(1.13)	(0.30)	(1.18)
$\Delta(s_{i,t})$	0.057**	0.038	0.035	0.020	0.035	0.043*	-0.019	-0.019
	(0.024)	(0.023)	(0.029)	(0.022)	(0.023)	(0.023)	(0.087)	(0.088)
$\Delta(a_{i,t})$	0.0063	0.0061	-0.010	0.0082	0.0077	0.0051	0.0083	0.0084
	(0.0089)	(0.0088)	(0.0098)	(0.010)	(0.0086)	(0.0087)	(0.0082)	(0.0085)
$\Delta(a^2_{i,t})$	-0.00052	-0.00051	0.00028	-0.00054	-0.00057*	-0.00048	-0.00055*	-0.00056*
	(0.00036)	(0.00036)	(0.00037)	(0.00039)	(0.00032)	(0.00034)	(0.00033)	(0.00033)
$\Delta(\sigma^2_{i,t})$	—	—	—	-0.63	—	—	—	—
	—	—	—	(0.45)	—	—	—	—
R^2	0.29	0.38	0.38	—	0.41	0.37	0.36	0.36
First-stage F	—	—	—	—	—	19.6	25.3	10.0
Countries	116	116	116	109	116	116	116	116
Observations	613	613	613	461	613	613	613	613
Controls	—	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the growth rate of log output per worker $\Delta \ln(y_t)$. All specifications include lagged schooling $s_{i,t-1}$ and a full set of time effects. Columns (2) to (8) add further controls $x_{i,t}$ for the quality of economic institutions, the value added by the agricultural sector, the percentage of land area in the tropics, and a full set of regional dummies. First-stage F refers to the Kleibergen-Paap first-stage F-statistic. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

B.2 Coding of Compulsory Schooling Laws

Table B.2: Coding Description

This table describes coding choices for countries in which compulsory schooling laws differ by schooling type or target group, and countries which experienced longer spells of turbulence and civil war. This list contains all countries for which information on compulsory schooling was available and, thus, even those that do not enter the estimation sample.

Albania: United Nations Educational, Scientific, and Cultural Organization (UNESCO) yearbooks report four plus an additional three years of compulsory schooling for 1963 and 1964. From 1965 to 1967, four plus an additional other four years of compulsory schooling are reported. We code these as seven and eights years of schooling because "[f]our years' schooling is compulsory for all children; a second period of three (four) years is compulsory for children in towns and villages where a seven-grade (eight-grade) school is available" (UNESCO, 1963–1968).

Andorra: Andorra's educational system is split into French and Spanish schools. However, because both schooling system differ in terms of compulsory schooling, we follow UNESCO's convention and code values as missing until 1977. Afterwards, both schools require a minimum of ten years of schooling so that we code a value of ten.

Angola: "The school system in the Portuguese Overseas Provinces forms part of the general pattern of Portuguese education. It is consequently the same as in metropolitan territory, but not all the levels and types of education provided in Portugal are to be found overseas" (UNESCO, 1964–1967). Therefore, we assume that for the years 1964 to 1967, compulsory schooling amounts to four years as it is the case for Portugal.

Argentina: In 1972, compulsory schooling takes a value of eights years, while before and afterwards compulsory schooling is consistently reported with seven years. Because the structure of the educational system did not change in 1972, we code a value of seven years.

Australia: For 1963 and 1964, UNESCO yearbooks report eight to ten years of compulsory schooling, varying by state. We take the figure of New South Wales, the most densely populated state, and thus code nine years. From 1968 onward, the yearbooks report values between nine and eleven years, varying by state or whether kindergarten counts toward primary education. We code ten years of compulsory schooling. This figure is consistent with more recent data published by the World Bank (2017). Moreover, the number reflects average compulsory schooling.

Bahrain: For 1971 and 1972, eight years of compulsory schooling are reported. However, change "will be applied in 1973/1974" (UNESCO, 1971). Following the value of the preceding years, we code a value of zero. For the period 1987 to 1993, zero compulsory schooling is reported. This figure contrasts with the high values before and afterwards. We thus code a missing rather than a zero value. Based on the age range, twelve years of compulsory are reported for the period 1995 to 1997. However, the compulsory program only contains six years of primary schooling with a general academic curriculum combined with religious instruction, which continues to nine years. Correspondingly, we code nine instead of twelve years for the period 1995 to 1997.

Barbados: There is no compulsory schooling from 1963 to 1967; however, the value for 1966 is missing. We impute this value to be zero. For the years 1995 to 1997, UNESCO yearbooks report twelve years of compulsory schooling instead of eleven in preceding and subsequent periods.

Belgium: In 1985 and 1986, UNESCO yearbooks report eight and nine years of compulsory schooling. Based on the preceding years and the age range, nine years of compulsory schooling are implausible, however. Therefore, we code eight years in 1985 and 1986.

Benin: After independence in 1960, there was a longer spell of political turbulence. In particular, several changes in power occurred at the beginning of the 1970s. According to the UNESCO yearbooks, compulsory schooling amounts to six years until 1970, zero years from 1971 to 1974, and seven years from 1975 onward. Due to the unstable nature of government, the exact role of compulsory schooling and whether it was enforced is unclear. Therefore, we decide to code the years 1971 to 1974 as missing rather than a clean zero.

Brazil: For 1963 and 1964, UNESCO yearbooks report compulsory schooling values of four and five years. From 1965 onward, the level remains consistently at four years. Because Brazil follows the Portuguese educational system, we code a value of four for 1963 and 1964.

Brunei: For the years 1995 to 1997, UNESCO yearbooks report compulsory schooling levels of twelve years. These stand in contrast to nine years of compulsory schooling before and afterwards. Because neither the educational system nor the age range of compulsory schooling changed during this period, we code nine instead of twelve years.

Cameroon: Historically, the educational system consisted of French schools in the Eastern and British schools in the Western part of Cameroon. In 1976, the British system was adopted in the entire country. We use the British system's compulsory schooling regulations throughout all periods. UNESCO yearbooks list eight years of compulsory schooling in 1969 and 1970. Given the subsequent period without any compulsory schooling, it is unlikely that this regulation has been enforced. We thus code a zero value for both 1969 and 1970.

Canada: Compulsory schooling "[...] figures vary slightly from one Province to another" (UNESCO, 1963). Values range from seven to ten years between 1963 and 1968, and eight to ten years between 1969 and 1994. We take a slightly conservative view and code a value of eight years for 1963–1968 and nine years for 1969–1994.

Cape Verde: "The school system in the Portuguese Overseas Provinces forms part of the general pattern of Portuguese education. It is consequently the same as in metropolitan territory, but not all the levels and types of education provided in Portugal are to be found overseas" (UNESCO, 1964–1967). Therefore, we assume that for the years 1964 to 1967, compulsory schooling amounts to four years as it is the case for Portugal.

Czech Republic and Slovakia: We use compulsory schooling regulations of former Czechoslovakia for both countries prior to 1994.

Egypt: For 1989 to 1991, UNESCO yearbooks report nine years of compulsory education. However, these figures are implausible given five years of primary and three years of lower secondary education. Therefore, we code eight rather than nine years. For 1995 and 1996, the yearbooks report five years of compulsory schooling. This figure does not reflect lower secondary education, which is also compulsory since the educational reforms in the early 1990s. Hence, we set the corresponding value to eight years instead of five.

Fiji: Between 1975 and 1997, UNESCO yearbooks report zero years of compulsory schooling in contrast to eight years from 1963 to 1974. We follow the World Bank convention, which codes missing values of compulsory schooling between 1998 and 2015 (World Bank, 2017). Thus, we code missing values for the years 1975 to 1997.

Finland: For 1967, 1971, and 1972, UNESCO yearbooks report eight instead of formerly nine years of compulsory schooling. Based on the age range and structure of the educational system, these shifts seem implausible. Thus, we code nine year of compulsory schooling.

Germany: Figures are based on West Germany prior to 1990. We code twelve rather than nine years of compulsory schooling in 1968–1970 and 1973–1988. This coding includes nine years of compulsory schooling plus an additional three years of "[...] part time vocational education" (UNESCO, 1973).

Guinea: In 1971/1972, compulsory schooling increases from eight to twelve years before it dropped again back to eight years in 1973. Throughout this period, the overall structure of the educational system remained unaltered. The only detectable change was the range of compulsory schooling from 7–15 to 7–22 which is implausible in comparison to other countries and also Guinea's legal age. Therefore, compulsory schooling is coded to remain at eight years instead of twelve.

Table B.2: . . . Continued

Guinea-Bissau: For 1981 and 1982, UNESCO yearbooks report seven years of compulsory schooling in contrast to six years in preceding and subsequent periods. Because the educational system remained unaltered during these years, this change seems implausible. Hence, we code six rather than seven years.

Guyana: Throughout the period 1963 to 1997, compulsory schooling takes a value of eight years with the exception of 1981 and 1982 (nine years), 1983 (six years) and 1995 to 1997 (ten years). However, the shifts are inconsistent with the relative stability of the educational system between 1980 and 1984, as well as the age range of compulsory schooling from six to 14 for the period 1995 to 1997. We thus code eight years over the entire period.

India: In 1971 and 1972, the UNESCO yearbooks report various levels of compulsory schooling. In the years thereafter, only a uniform level of five years is reported. This change is justified by the fact that "[t]his information pertains to the majority of states" (UNESCO, 1975). Therefore, we code also a value of five years for 1971 and 1972.

Indonesia: In 1973 and 1974, UNESCO yearbooks report zero values for compulsory schooling. These figures stand in contrast to six years of compulsory schooling before and thereafter. Moreover, the educational system remained unaltered during this period. Hence, we code a value of six instead of zero years.

Iran: For the years 1966, 1967, 1973, and 1974, UNESCO yearbooks report five years of compulsory schooling in contrast to six years in preceding and intermediate periods. However, these figures seem implausible, because the educational remained unaltered during this period. Therefore, we code six rather than five years.

Iraq: UNESCO yearbooks consistently report six years of compulsory schooling. In 1983, however, five years are reported, although the educational structure did not change. We code six instead of five years. Moreover, compulsory schooling is missing in 1973 and 1974. Because the educational system remained unaltered, we set the value to six years—the same as in the preceding and following years.

Israel: Between 1981 and 1987, UNESCO yearbooks report nine years of compulsory schooling in contrast to eleven years in preceding and subsequent periods. Moreover, this figure seems implausible given the age range from five to 15. Hence, we code eleven instead of nine years for this period.

Jordan: According to the UNESCO yearbooks, compulsory schooling increased from six to nine years in 1964 based on a widening of the age range. However, this increase is not observed in 1965 where the age range is again six years. Therefore, we code six instead of nine years.

Kiribati and Tuvalu: Until 1976, the islands were a British protectorate under the name Gilbert and Ellice Islands. We thus use compulsory schooling of the former protectorate for both Kiribati and Tuvalu. For the years 1975 to 1980, during which the islands became independent, we code missing instead of the reported zero values. For the years 1985 and 1986, UNESCO yearbooks report five years of schooling in contrast to nine years in the preceding and subsequent periods. Because the educational system remained unaltered during this time, we code nine instead of five years.

Kuwait: For 1982 and 1983, UNESCO yearbooks report four in contrast to eight years in preceding and subsequent periods. Because the educational system remained unaltered during this period, we code eight instead of four years.

Laos: For the years 1990 to 1994, UNESCO yearbooks report eight in contrast to five years of compulsory schooling in preceding and subsequent periods. Because the educational system with five years of compulsory primary schooling remained unaltered during this period, we code five instead of eight years.

Lebanon: Throughout the period 1963 to 1997, compulsory schooling is consistently zero years, except for 1971 where UNESCO yearbooks report a value of twelve. Given the overall trend, this value seems implausible so that we code zero years.

Lesotho: The UNESCO yearbooks report compulsory schooling of eight years for the former British Crown colony Basutoland in 1964 and 1965. However, there was no compulsory schooling for the independent state of Lesotho between 1966 and 1984. Moreover, the yearbooks also report a value of zero for the colony in 1963. We thus set the value for compulsory schooling to zero for 1964 and 1965.

Malawi: For 1963 to 1965, UNESCO yearbooks report eight years of compulsory schooling based on the English schools in the former British colony. From 1966 onward, zero years of schooling are reported. Because Malawi became independent in 1964, eight years of compulsory schooling seem implausible. Hence, we code zero rather eight years.

Malaysia: From 1968 to 1984, UNESCO yearbooks report six years of compulsory schooling for some and zero or missing values for other regions. Because there is no compulsory schooling in the most populous regions, we code zero years from 1968 to 1984.

Malta: In 1986 and 1987, UNESCO yearbooks report twelve years of compulsory schooling. Based on the stable educational system, the age range, and subsequent values, these figures seem implausible. We code ten instead of twelve years.

Mauritius: For the years 1981 to 1983, UNESCO yearbooks report eight years of compulsory schooling in contrast to seven years in preceding and subsequent years. Because the educational system remained unaltered during this period, we code seven rather than eight years. Between 1987 and 1994, figures for compulsory schooling drop to zero. However, these values seem implausible, because the educational system did not change in this period either. We code missing instead of zero values.

Monaco: For 1973 and 1974, UNESCO yearbooks report eleven years of compulsory schooling in contrast to ten years before and afterwards. Because the educational system remained unaltered during this period, we code ten rather than eleven years.

Mozambique: "The school system in the Portuguese Overseas Provinces forms part of the general pattern of Portuguese education. It is consequently the same as in metropolitan territory, but not all the levels and types of education provided in Portugal are to be found overseas" (UNESCO, 1964–1967). Therefore, we assume that for the years 1964 to 1967, compulsory schooling amounts to four years as it is the case for Portugal.

Nauru: For the years 1963 to 1970, UNESCO yearbooks report nine years of compulsory schooling for European and ten years for Nauruan schools. We code a value of ten years.

Nepal: Historically, the Nepalese educational system consisted of English and Sanskrit schools. Until 1967, there was no compulsory schooling for either of these schools. Beginning in 1968, the English school system prescribed five years of schooling while attendance at Sanskrit schools was not compulsory. Following the UNESCO's convention to document compulsory schooling based on the English system from 1973 onward (UNESCO, 1973), we code five years of compulsory schooling.

New Zealand: For the years 1994 to 1997, UNESCO yearbooks report eleven years of compulsory schooling. However, the educational system consists of six years of primary and four years of lower secondary schooling. For this reason, we code ten rather than eleven years. This coding choice is consistent with preceding and subsequent periods and the stability of the educational system overall.

Niger: For the period 1973 to 1979, the UNESCO yearbooks report compulsory schooling of 12/13 years in contrast to eight years in the preceding and subsequent periods. This substantial change is not reflected in a corresponding transformation of the educational system and only represents shifts in the age range for compulsory schooling. Therefore, this extreme increase seems implausible so that we code compulsory schooling to remain at eight years throughout 1973 to 1979.

Norway: From 1968 to 1970, values of seven and nine years are being reported, because "[a] law passed in 1968 extended compulsory education from seven to nine years. This has been applied in most municipalities" (UNESCO, 1968).

Table B.2: ...Continued

Philippines: In 1963 and 1964, a missing value of compulsory schooling is reported. However, we decide to code a zero value, because "[i]n implementation of Republic Act No. 1124, Department Order No. 1, s.1957, Article 2 states that elementary education shall ultimately be made available for all children between 7 and 13 years" (UNESCO, 1963). Hence, compulsory schooling was not yet implemented in 1963 and 1964.

Poland: Between 1963 and 1970, UNESCO yearbooks report various values of compulsory schooling. We take a conservative view and code 1963 and 1964 with a value of seven years and 1965 to 1970 with a value of eight years.

Republic of Congo: For the period 1973 and 1974, compulsory schooling dropped from an initial value of ten to six years. From 1975 onward, compulsory schooling reverted back to a value of ten years. Throughout this entire time, compulsory schooling age ranges from six to 16 years for boys and six to 17 years for girls. Therefore, we also code a value of ten years for 1973 and 1974.

Romania: For 1963 and 1964, UNESCO yearbooks report seven or eight years of compulsory schooling. In subsequent years, educational regulations prescribe eight years of compulsory schooling. Based on this stability in the educational system, we set values to eight years for 1963 and 1964.

Saint Lucia: For 1985 and 1986, UNESCO yearbooks report eleven years of compulsory education in contrast to ten years in preceding and subsequent periods. Because the structure of the educational system with seven years of primary and three years of lower secondary schooling did not change during these years, this shifts seems implausible. Hence, we code ten rather than eleven years.

Sao Tome and Principe: "The school system in the Portuguese Overseas Provinces forms part of the general pattern of Portuguese education. It is consequently the same as in metropolitan territory, but not all the levels and types of education provided in Portugal are to be found overseas" (UNESCO, 1964–1967). Therefore, we assume that for the years 1964 to 1967, compulsory schooling amounts to four years as it is the case for Portugal.

Senegal: UNESCO yearbooks report seven years of compulsory education for 1971 and 1972 and six years for 1973 and 1974. However, compulsory primary education corresponded only to six and five years. Therefore, we code six and five years rather than seven and six.

Singapore: Compulsory schooling has only been introduced in 2003. Hence, we code one missing value as zero before 2003.

South Africa: Between 1963 and 1984, UNESCO yearbooks report seven and nine years of compulsory schooling, varying by state and race. We code seven years of schooling as the corresponding figure for the black population, which constitutes approximately 80 percent of the total population.

Sri Lanka: From 1995 to 1997, UNESCO yearbooks report eleven years of compulsory schooling in contrast to ten years beforehand. Based on the age limits that remained unaltered over this period, we code ten instead of eleven years.

St. Vincent and The Grenadines: UNESCO yearbooks report ten years of compulsory schooling for 1968–1974 as well as 1978–1985, and zero years for the periods 1963–1967, 1975–1977, and 1986–1995. Between 1996 and 2004, no values are reported. The overall structure of the educational system did not change substantially throughout all these periods so that large shifts in compulsory schooling appear implausible. We thus code values for 1963–1967, 1975–1977 and 1986–1995 to be missing rather than zero.

Suriname: UNESCO yearbooks report eleven years of compulsory schooling for the period 1995 to 1997. This figure stands in stark contrast to only six years before and afterwards. Because the educational system with six years of compulsory primary schooling remained unaltered during these years, we code six instead of eleven years.

Swaziland: In the early years until 1965, the educational system consisted of European, African, and Eurafrican schools. Because education was compulsory only at European schools, which were abolished from 1966 onward, and not for the other school types, we code a value of zero.

Switzerland: According to the UNESCO yearbooks, compulsory schooling varies between seven and nine years across Swiss cantons from 1963 to 1997. In some cantons, students are additionally required to take up at least two years of "complementary part-time schooling" (UNESCO, 1963). Hence, the reported figures are likely too low. Thus, we follow the convention of UNESCO reports from 1975 to 1981 and code nine years of compulsory schooling throughout the entire period.

Thailand: In 1963 and 1964, UNESCO yearbooks report between four and seven years of compulsory schooling. Based on the age range and subsequent values, we code both observations as seven.

Tonga: For the years 1995 to 1997, UNESCO yearbooks report eight years of compulsory schooling. However, this figure seems implausible in comparison to six years in preceding and subsequent periods. Moreover, the educational system remained unaltered during these years. Hence, we code six rather than eight years. For 2012 to 2015, the World Bank (2017) reports eight and then 15 years of compulsory schooling. These figures are implausible, because only primary education, which requires six years of schooling, is compulsory in Tonga. Therefore, we code also six years of compulsory schooling for the period 2012 to 2015.

Trinidad and Tobago: In 1973 and 1974, compulsory schooling is reported to possess a value of ten years. Before 1973 and after 1974, this figure corresponds to seven years. Because only primary schooling is compulsory with a standard duration of seven years given entry ages for primary and secondary schooling, we code a value of seven for 1973 and 1974.

Turkey: Between 1965 and 1967, eight years of compulsory schooling is reported. However, only five years of primary schooling were compulsory. In line with preceding and subsequent periods, we thus code a compulsory schooling of five years.

Tunisia: From 1968 to 1981, UNESCO yearbooks report six years of compulsory schooling. For 1982 and 1983, no values are reported. From 1984 onward, compulsory schooling is documented with a value of zero until 1992. The yearbooks show eleven years of compulsory schooling for 1993/1994, and nine years from 1995 onward. The educational system consists of six years of primary schooling, three years of lower secondary schooling, and a further four years of upper secondary schooling. This structure is maintained throughout the entire period 1981 to 1995. Because zero values are implausible, we code them as missing. For the years 1993 and 1994, we set compulsory schooling to nine instead of eleven years.

United States: For the years 1963 to 1997, UNESCO yearbooks present values ranging from ten to twelve years for the U.S. Minimum compulsory schooling corresponds to ten years, formally from age six to 16. Some states require students to remain in school until coming of age, implying two further years. However, there are also exemption regulations for religious groups and homeschooling. We take a conservative view and set the compulsory schooling thus to the minimum value of ten years, which is fulfilled by all states.

Vanuatu: Historically, the educational system consists of English and French schools. Compulsory schooling years refer to regulations with respect to English schools.

Yemen: Figures are based on compulsory schooling of the former Arab Republic of Yemen and the Republic of Yemen.

Zambia: For the years 1963 to 1966, UNESCO yearbooks report compulsory schooling of eight years with zero years from 1967 onward. Because "[e]ducation is compulsory in certain areas only" (UNESCO, 1963–1966), we code the years 1963 to 1966 as zero.

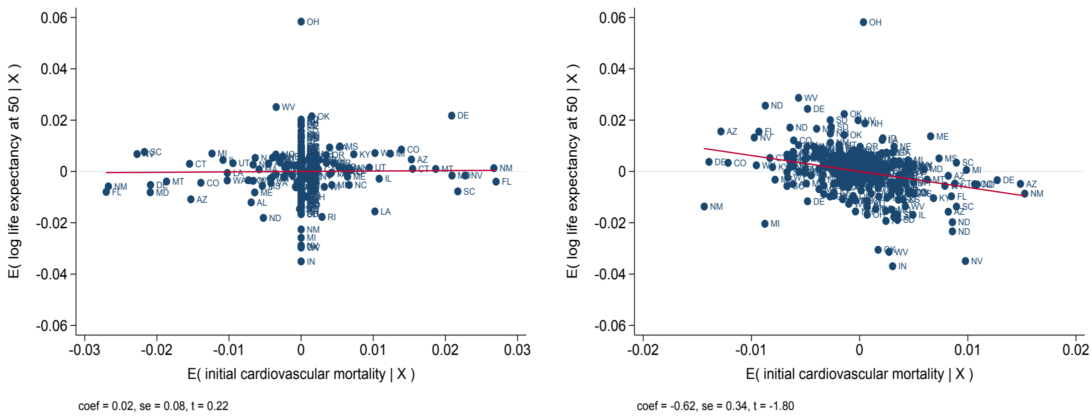
Azerbaijan, Armenia, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan: Prior to 1992, we code compulsory schooling according to the values of the former Soviet Union. Between 1963 and 1966, UNESCO yearbooks report eight and nine years of compulsory schooling. Because primary schooling comprises only eight grades, we code eight rather than nine years.

Bosnia and Herzegovina, Croatia, Macedonia, Montenegro, Serbia, Slovenia: Prior to 1993, we code compulsory schooling according to values of former Yugoslavia. Figures of Serbia and Montenegro are taken from the Federal Republic of Yugoslavia for the years 1993 to 1997.

Appendix C

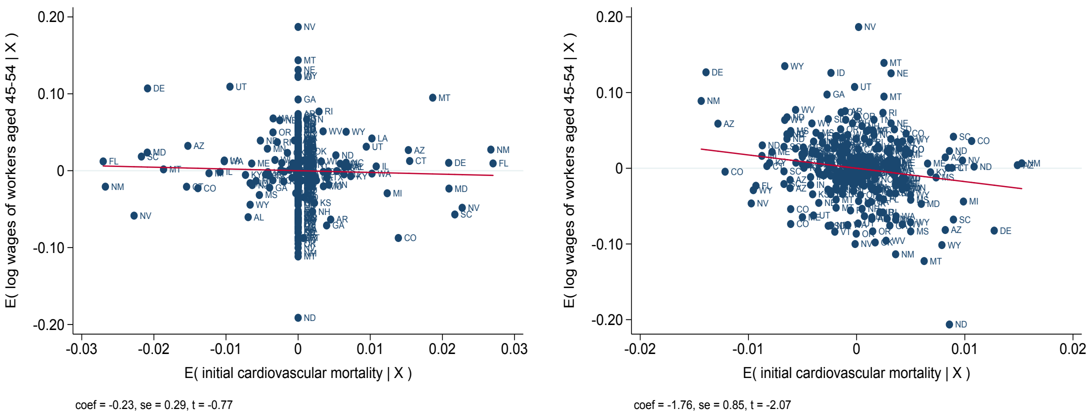
Appendix to Chapter 4

C.1 Additional Figures



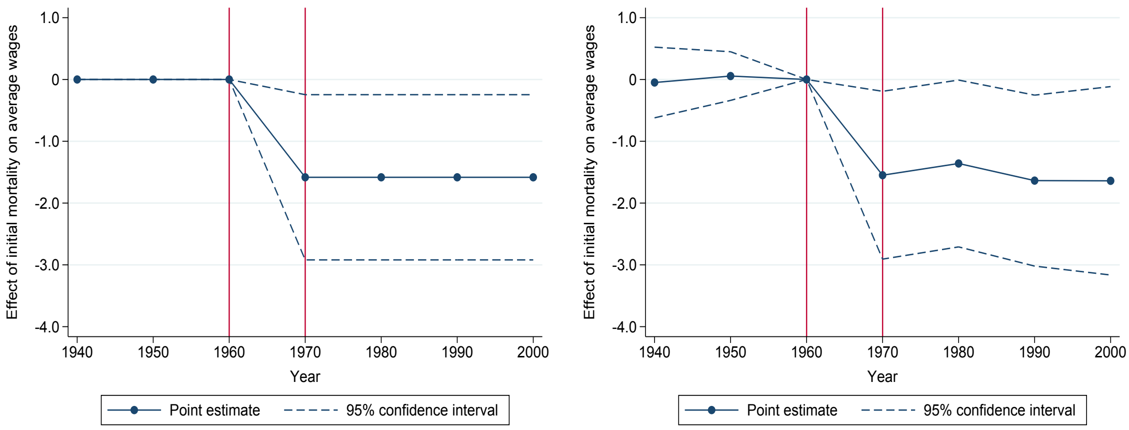
(a) Pre Treatment (Table 4.1b, Column 5) (b) Post Treatment (Table 4.1b, Column 5)

Figure C.1: Partial Correlation Plots: First Stage (Flexible Model)



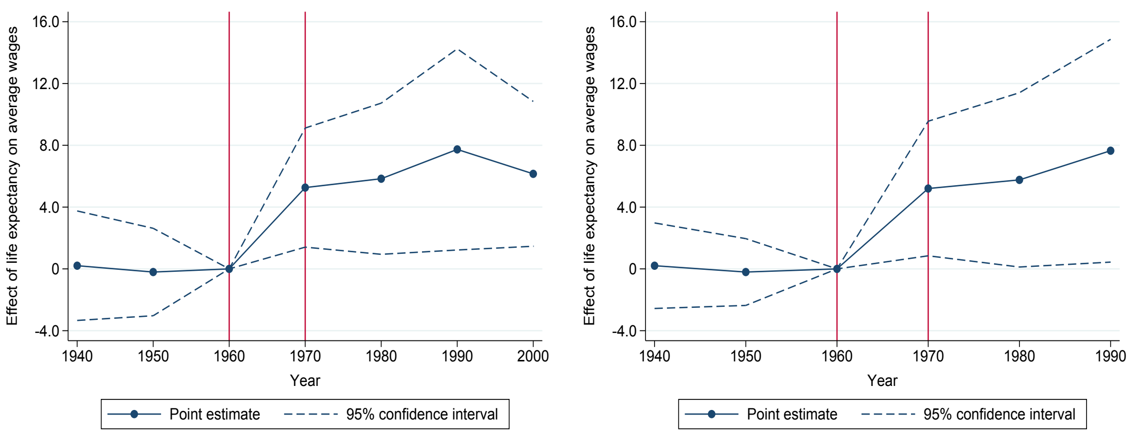
(a) Pre Treatment (Table 4.2b, Column 5) (b) Post Treatment (Table 4.2b, Column 5)

Figure C.2: Partial Correlation Plots: Reduced Form (Flexible Model)



(a) DD Model (Table C.3a, Column 5) (b) Flexible Model (Table C.3b, Column 5)

Figure C.3: Robustness: Reduced Form – Total Workforce



(a) 1940–2000 (Table 4.10, Column 3) (b) 1940–1990 (Table 4.10, Column 4)

Figure C.4: Illustration: Effect of Life Expectancy over Time (Total Workforce)

C.2 Additional Tables

Table C.1: Descriptive Statistics by Age Group

	Total	15–24	25–34	35–44	45–54	55–64	65+
Log wages	10.20	9.41	10.24	10.44	10.46	10.36	9.90
	(0.37)	(0.26)	(0.35)	(0.37)	(0.40)	(0.40)	(0.31)
Labor force participation	58.43	56.78	71.71	73.26	71.23	56.42	17.86
	(6.25)	(9.90)	(9.81)	(9.54)	(9.03)	(5.51)	(5.79)
Usual work hours per week	17.68	19.19	29.59	30.32	29.47	23.23	6.84
	(3.63)	(4.85)	(5.76)	(5.29)	(5.03)	(3.21)	(1.93)
Usual work weeks per year	18.78	19.99	31.27	31.99	31.50	25.46	8.03
	(5.97)	(6.29)	(10.18)	(10.44)	(9.76)	(6.93)	(2.43)
Average work hours per year	723.93	641.58	1250.97	1304.04	1283.79	993.46	254.08
	(254.28)	(240.45)	(445.50)	(452.71)	(424.22)	(282.46)	(87.69)
Log life expectancy	3.65	4.00	3.81	3.58	3.29	2.95	2.74
	(0.07)	(0.05)	(0.06)	(0.07)	(0.09)	(0.11)	(0.12)
Mortality from CVD in 1960	0.23	—	—	—	—	—	—
× Post 1960	(0.20)	—	—	—	—	—	—
Controls in 1960 × Post 1960:							
Initial log life expectancy at 50	2.07	2.27	2.16	2.03	1.86	1.66	1.53
	(1.80)	(1.97)	(1.88)	(1.76)	(1.61)	(1.44)	(1.33)
Initial mortality other than CVD	0.21	—	—	—	—	—	—
	(0.18)	—	—	—	—	—	—
Initial share college graduates	0.03	—	—	—	—	—	—
	(0.03)	—	—	—	—	—	—
Initial share college enrollment	0.07	—	—	—	—	—	—
	(0.06)	—	—	—	—	—	—
Initial population density ($\times \frac{1}{100}$)	0.75	—	—	—	—	—	—
	(1.60)	—	—	—	—	—	—
Initial log wages	5.80	—	—	—	—	—	—
	(5.03)	—	—	—	—	—	—
Sample weights:							
Initial white population ($\times \frac{1}{100000}$)	32.95	4.36	4.17	4.47	3.83	2.94	3.18
	(34.26)	(4.31)	(4.42)	(4.84)	(4.18)	(3.27)	(3.41)

Notes: Descriptive statistics for balanced panel of the 48 contiguous states from 1940–2000 with a total number of 336 observations. Numbers are means for the respective variable in the total population or a specific age group. Standard deviations are in parentheses. CVD is an abbreviation for cardiovascular diseases.

Table C.2: OLS: Adult Life Expectancy and Average Wages of Workers Aged 45–54

	Dependent variable: log wages of whites 45–54				
	(1)	(2)	(3)	(4)	(5)
	Ordinary Least Squares				
Log life expectancy at 50	-0.43 (0.26)	-0.05 (0.37)	-0.04 (0.38)	0.23 (0.41)	-0.01 (0.37)
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Estimates are weighted by the initial white population of 45- to 54-year-olds. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.3: Robustness: Reduced Form for Total Workforce

	Dependent variable: log wages of the total workforce				
	(1)	(2)	(3)	(4)	(5)
	(a) Differences-in-differences model				
Mortality × Post 1960	-0.62***	-0.65	-1.34**	-1.18**	-1.58**
	(0.19)	(0.43)	(0.61)	(0.56)	(0.66)
	(b) Flexible model				
Mortality × 1940	0.36	0.36	0.35	-0.33	-0.05
	(0.23)	(0.23)	(0.24)	(0.22)	(0.28)
Mortality × 1950	0.13	0.13	0.12	0.11	0.06
	(0.17)	(0.17)	(0.17)	(0.21)	(0.20)
Mortality × 1970	-0.05	-0.08	-0.75	-1.11*	-1.55**
	(0.17)	(0.41)	(0.58)	(0.56)	(0.68)
Mortality × 1980	-0.79***	-0.81	-1.49**	-0.77	-1.36**
	(0.27)	(0.50)	(0.64)	(0.54)	(0.67)
Mortality × 1990	-0.42	-0.44	-1.15	-1.49**	-1.64**
	(0.35)	(0.54)	(0.72)	(0.64)	(0.69)
Mortality × 2000	-0.57	-0.60	-1.34*	-1.67**	-1.64**
	(0.39)	(0.57)	(0.76)	(0.72)	(0.76)
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Estimates are weighted by the initial white population. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.4: Robustness: Linear Specification of Life Expectancy

	Dependent variable: log wages of whites 45–54				
	(1)	(2)	(3)	(4)	(5)
	(a) Differences-in-differences model (2SLS)				
Life expectancy at 50	-0.08***	0.39	0.22**	0.10***	0.15**
	(0.02)	(0.29)	(0.10)	(0.03)	(0.06)
First-stage F–statistic	46.8	2.1	6.8	35.5	9.4
	(b) Flexible model (2SLS)				
Life expectancy at 50	-0.05***	0.03	0.04*	0.09***	0.12*
	(0.02)	(0.02)	(0.02)	(0.03)	(0.06)
First-stage F–statistic	10.9	7.7	9.8	7.9	2.0
Hansen test p–value	0.04	0.002	0.004	0.2	0.8
	(c) Flexible model (LIML)				
Life expectancy at 50	-0.06***	0.12	0.10*	0.12***	0.15*
	(0.02)	(0.14)	(0.06)	(0.04)	(0.09)
First-stage F–statistic	10.9	7.7	9.8	7.9	2.0
Hansen test p–value	0.04	0.05	0.05	0.3	0.9
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Estimates are weighted by the initial white population of 45- to 54-year-olds. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.5: Adult Life Expectancy and Average Wages by Age Cohorts

	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Flexible model (2SLS)						
Log life expectancy	1.00	3.61***	2.71***	2.35***	0.84	0.66	3.13***
(of specific age group)	(2.05)	(1.29)	(0.80)	(0.89)	(0.74)	(1.10)	(0.94)
First-stage F–statistic	3.4	5.3	8.5	3.5	5.9	2.4	5.8
Hansen test p–value	0.4	0.8	1.0	0.8	0.7	0.02	0.8
	(b) Flexible model (LIML)						
Log life expectancy	1.13	3.84***	2.75***	2.57**	0.90	1.48	3.30***
(of specific age group)	(2.52)	(1.40)	(0.81)	(1.00)	(0.78)	(2.80)	(1.00)
First-stage F–statistic	3.4	5.3	8.5	3.5	5.9	2.4	5.8
Hansen test p–value	0.4	0.8	1.0	0.8	0.7	0.02	0.8
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with a full set of time dummies with the year 1960 as reference category. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.6: Adult Life Expectancy and Average Wages: No Migration

	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Flexible model (2SLS)						
Log life expectancy	1.71	4.36***	4.83***	3.91***	1.85**	6.41**	4.88***
(of specific age group)	(1.91)	(1.19)	(0.97)	(1.20)	(0.93)	(2.84)	(1.08)
First-stage F–statistic	4.7	6.1	6.6	3.2	5.4	3.8	5.9
Hansen test p–value	0.01	0.008	0.009	0.02	0.006	0.1	0.002
	(b) Flexible model (LIML)						
Log life expectancy	3.67	6.24***	6.89***	5.51***	2.59**	7.44**	6.60***
(of specific age group)	(3.61)	(2.06)	(1.72)	(2.07)	(1.27)	7.44**	(1.65)
First-stage F–statistic	4.7	6.1	6.6	3.2	5.4	3.8	5.9
Hansen test p–value	0.02	0.03	0.03	0.07	0.009	0.1	0.008
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with a full set of time dummies with the year 1960 as reference category. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.7: Adult Life Expectancy and Average Wages: No Old-Age Migration

	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Flexible model (2SLS)						
Log life expectancy	0.25	3.93**	2.59***	2.22**	0.65	1.72	3.00***
(of specific age group)	(3.28)	(1.61)	(0.93)	(0.89)	(0.65)	(1.43)	(0.97)
First-stage F–statistic	1.4	3.2	6.3	3.9	7.8	2.7	4.7
Hansen test p–value	0.5	0.8	0.3	0.5	0.3	0.05	0.3
	(b) Flexible model (LIML)						
Log life expectancy	-0.97	4.42**	2.91***	2.63**	0.75	3.90	3.45***
(of specific age group)	(9.13)	(1.90)	(1.07)	(1.10)	(0.71)	(4.16)	(1.18)
First-stage F–statistic	1.4	3.2	6.3	3.9	7.8	2.7	4.7
Hansen test p–value	0.5	0.8	0.3	0.5	0.3	0.08	0.4
States	45	45	45	45	45	45	45
Observations	315	315	315	315	315	315	315
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with a full set of time dummies with the year 1960 as reference category. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.8: Adult Life Expectancy and Average Wages: Metropolitan Areas

	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Flexible model (2SLS)						
Log life expectancy	2.48	5.01***	3.73***	4.35***	2.40**	0.60	4.97***
(of specific age group)	(2.16)	(1.33)	(0.91)	(1.64)	(0.94)	(1.35)	(1.15)
First-stage F–statistic	6.2	12.8	15.3	4.4	7.5	3.3	11.8
Hansen test p–value	0.02	0.2	0.4	0.1	0.5	0.01	0.2
	(b) Flexible model (LIML)						
Log life expectancy	2.99	5.40***	3.82***	5.27**	2.47**	0.58	5.29***
(of specific age group)	(3.13)	(1.48)	(0.95)	(2.17)	(0.99)	(1.58)	(1.27)
First-stage F–statistic	6.2	12.8	15.3	4.4	7.5	3.3	11.8
Hansen test p–value	0.02	0.2	0.4	0.2	0.5	0.01	0.2
States	33	33	33	33	33	33	33
Metropolitan Areas	89	89	89	89	89	89	89
Observations	623	623	623	623	623	623	623
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include metropolitan-area-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with a full set of time dummies with the year 1960 as reference category. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.9: Heterogeneity: White-Collar and Blue-Collar Workers

	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) White-collar workers: flexible model (2SLS)						
Log life expectancy	-3.29	6.40***	3.63**	1.00	-0.53	4.83**	3.60***
(of specific age group)	(4.66)	(2.13)	(1.45)	(1.12)	(0.74)	(2.04)	(1.29)
First-stage F–statistic	2.0	3.5	6.4	2.6	5.2	5.3	4.7
Hansen test p–value	0.6	0.4	0.5	0.4	0.09	0.04	0.3
	(b) White-collar workers: flexible model (LIML)						
Log life expectancy	-5.82	8.31***	3.97**	1.26	-0.60	6.74**	4.21***
(of specific age group)	(8.09)	(3.14)	(1.60)	(1.48)	(0.86)	(3.11)	(1.62)
First-stage F–statistic	2.0	3.5	6.4	2.6	5.2	5.3	4.7
Hansen test p–value	0.6	0.6	0.5	0.4	0.09	0.07	0.3
	(c) Blue-collar workers: flexible model (2SLS)						
Log life expectancy	5.37	4.65**	3.41**	4.12**	3.57**	-1.28	4.32**
(of specific age group)	(3.50)	(2.17)	(1.38)	(2.02)	(1.43)	(1.94)	(1.74)
First-stage F–statistic	2.3	3.2	5.9	2.4	5.0	5.3	4.4
Hansen test p–value	0.7	0.8	0.9	0.9	0.8	0.4	1.0
	(d) Blue-collar workers: flexible model (LIML)						
Log life expectancy	6.09	5.30**	3.62**	4.55**	3.81**	-1.35	4.44**
(of specific age group)	(4.00)	(2.50)	(1.47)	(2.31)	(1.54)	(2.21)	(1.79)
First-stage F–statistic	2.3	3.2	5.9	2.4	5.0	5.3	4.4
Hansen test p–value	0.7	0.8	0.9	0.9	0.8	0.4	1.0
States	48	48	48	48	48	48	48
Observations	288	288	288	288	288	288	288
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Average wages contain observations from white-collar workers in Panels (a) and (b) and from blue-collar workers in Panels (c) and (d). Control variables are measured in 1960 and interacted with a full set of time dummies with the year 1960 as reference category. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.10: Heterogeneity: College and Non-College Workers

	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) College workers: flexible model (2SLS)						
Log life expectancy	-1.54	2.63*	2.12**	1.16	-0.01	2.86	2.27**
(of specific age group)	(3.39)	(1.59)	(1.06)	(0.96)	(1.30)	(2.97)	(0.90)
First-stage F–statistic	3.1	5.5	8.7	3.8	6.1	2.8	5.9
Hansen test p–value	0.7	0.8	0.6	0.4	0.09	0.06	0.7
	(b) College workers: flexible model (LIML)						
Log life expectancy	-1.74	2.76	2.19**	1.28	0.01	4.52	2.34**
(of specific age group)	(3.71)	(1.70)	(1.10)	(1.10)	(1.62)	(5.69)	(0.94)
First-stage F–statistic	3.1	5.5	8.7	3.8	6.1	2.8	5.9
Hansen test p–value	0.7	0.8	0.6	0.4	0.09	0.08	0.7
	(c) Non-college workers: flexible model (2SLS)						
Log life expectancy	0.14	1.48	0.94	1.67*	0.88	0.28	1.20
(of specific age group)	(2.21)	(1.11)	(0.76)	(0.86)	(0.82)	(1.06)	(0.85)
First-stage F–statistic	3.4	5.3	8.5	3.4	5.9	2.4	5.7
Hansen test p–value	0.4	0.6	0.7	0.6	1.0	0.09	0.6
	(d) Non-college workers: flexible model (LIML)						
Log life expectancy	0.22	1.56	1.01	1.81*	0.90	0.54	1.28
(of specific age group)	(2.65)	(1.19)	(0.80)	(0.93)	(0.83)	(1.55)	(0.89)
First-stage F–statistic	3.4	5.3	8.5	3.4	5.9	2.4	5.7
Hansen test p–value	0.4	0.6	0.7	0.6	1.0	0.09	0.6
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Average wages contain observations from workers with at least some college education in Panels (a) and (b) and from workers without any college education in Panels (c) and (d). Control variables are measured in 1960 and interacted with a full set of time dummies with the year 1960 as reference category. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.11: Adult Life Expectancy and Labor Supply by Age Cohorts: Flexible Model (2SLS)

	Flexible model (2SLS)						
	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Labor force participation (0 to 100 percent)						
Log life expectancy	-92.22	-227.50***	-108.21***	-25.16	-106.01***	-63.96*	-170.43***
(of specific age group)	(71.25)	(53.62)	(29.89)	(27.30)	(40.48)	(36.56)	(41.13)
Hansen test p–value	0.5	0.4	0.3	0.7	0.10	0.2	0.7
	(b) Usual hours per week						
Log life expectancy	-83.21**	-66.01***	-32.40**	-0.84	-34.18**	-15.37	-26.91*
(of specific age group)	(36.75)	(20.93)	(13.91)	(11.40)	(17.03)	(12.83)	(14.53)
Hansen test p–value	0.05	0.5	0.9	0.5	0.01	0.2	0.5
	(c) Usual weeks per year						
Log life expectancy	-81.63**	-85.75***	-43.09***	-4.91	-48.33**	-21.19	-30.87**
(of specific age group)	(41.08)	(25.03)	(15.14)	(12.86)	(19.01)	(16.54)	(15.14)
Hansen test p–value	0.7	0.6	0.5	0.8	0.03	0.03	0.8
	(d) Labor supply of those working (weeks $\times$ hours)						
Log life expectancy	-2384.18	-2820.10***	-1519.95**	193.90	-2094.78**	-643.37	-956.40
(of specific age group)	(1466.70)	(1043.86)	(688.12)	(581.45)	(858.40)	(631.58)	(671.45)
Hansen test p–value	0.5	0.6	1.0	0.7	0.02	0.06	0.6
First-stage F–statistic	3.4	5.3	8.5	3.5	5.9	2.4	5.8
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the labor force participation in Panel (a), usual hours worked per week in Panel (b), usual weeks worked per year in Panel (c), and hours worked per year of those working in Panel (d). All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with a full set of time dummies with the year 1960 as reference category. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.12: Adult Life Expectancy and Labor Supply by Age Cohorts: Flexible Model (LIML)

	Flexible model (LIML)						
	15–24	25–34	35–44	45–54	55–64	65+	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(a) Labor force participation (0 to 100 percent)						
Log life expectancy	-104.64	-261.05***	-118.71***	-36.96	-213.46***	-97.21	-177.53***
(of specific age group)	(85.79)	(65.51)	(33.20)	(36.26)	(80.06)	(71.79)	(43.29)
Hansen test p–value	0.5	0.5	0.3	0.7	0.2	0.4	0.7
	(b) Usual hours per week						
Log life expectancy	-132.30*	-76.23***	-33.37**	-5.18	-95.30**	-24.01	-30.39*
(of specific age group)	(73.57)	(24.70)	(14.30)	(20.70)	(47.01)	(22.59)	(16.26)
Hansen test p–value	0.1	0.5	0.9	0.5	0.08	0.3	0.5
	(c) Usual weeks per year						
Log life expectancy	-94.45*	-93.64***	-46.03***	-6.19	-88.84**	-57.58	-33.69**
(of specific age group)	(49.23)	(27.78)	(16.11)	(14.58)	(35.89)	(79.40)	(16.32)
Hansen test p–value	0.7	0.6	0.5	0.8	0.09	0.3	0.8
	(d) Labor supply of those working (weeks $\times$ hours)						
Log life expectancy	-2685.52	-3106.16***	-1543.29**	212.52	-4100.54**	-1474.06	-1056.13
(of specific age group)	(1696.27)	(1146.91)	(697.33)	(691.18)	(1750.78)	(1974.76)	(727.12)
Hansen test p–value	0.5	0.6	1.0	0.7	0.07	0.2	0.6
First-stage F–statistic	3.4	5.3	8.5	3.5	5.9	2.4	5.8
States	48	48	48	48	48	48	48
Observations	336	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the labor force participation in Panel (a), usual hours worked per week in Panel (b), usual weeks worked per year in Panel (c), and hours worked per year of those working in Panel (d). All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of the respective age group. Control variables are measured in 1960 and interacted with a full set of time dummies with the year 1960 as reference category. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.13: Adult Life Expectancy and College Enrollment

	Dependent variable: college enrollment 15–24					
	Diff-in-Diff. Model (2SLS)		Flexible Model (2SLS)		Flexible Model (LIML)	
	(1)	(2)	(3)	(4)	(5)	(6)
Log life expectancy at 30	1.19**		0.85*		0.91	
	(0.56)		(0.49)		(0.58)	
Log life expectancy at 50		0.77**		0.65**		0.84*
		(0.33)		(0.31)		(0.43)
First-stage F–statistic	22.6	15.0	5.4	3.2	5.4	3.2
Hansen test p–value	—	—	0.1	0.3	0.1	0.4
States	48	48	48	48	48	48
Observations	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of 15- to 24-year-olds. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college enrollment, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.14: Robustness: Moving-Window Model

	Dependent variable: log wages of whites 45–54			
	1940–1970	1940–1980	1940–1990	1940–2000
	(1)	(2)	(3)	(4)
Log life expectancy at 50	3.05**	2.09**	2.45***	2.54***
	(1.37)	(0.96)	(0.87)	(0.95)
First-stage F–statistic	23.6	37.5	47.3	14.2
States	48	48	48	48
Observations	192	240	288	336
Full controls	✓	✓	✓	✓

Notes: Regression results for moving window model which adds one post-treatment period at a time. All regressions include state-fixed, time, and region-year-fixed effects. Estimates are weighted by the initial white population of 45- to 54-year-olds. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.15: Robustness: Long-Differences Model

	Dependent variable: log wages of whites 45–54			
	1960–1970	1960–1980	1960–1990	1960–2000
	(1)	(2)	(3)	(4)
Log life expectancy at 50	5.96***	2.61*	4.72***	4.13
	(2.16)	(1.57)	(1.62)	(4.38)
First-stage F–statistic	13.7	24.3	14.6	0.8
States	48	48	48	48
Observations	96	96	96	96
Full controls	✓	✓	✓	✓

Notes: Regression results for long differences model. All regressions include state-fixed and region-year-fixed effects. Estimates are weighted by the initial white population of 45- to 54-year-olds. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table C.16: Effect of Individual Health on Wages

	Dependent variable: respondents' log wages					
	(1)	(2)	(3)	(4)	(5)	(6)
High blood pressure	-0.05**	-0.05**	-0.05**	-0.05**	-0.01	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
× born before 1910		-1.39***				
		(0.22)				
× born before 1920			-0.02			
			(0.29)			
× born before 1930				-0.01		
				(0.10)		
× born before 1940					-0.08*	
					(0.04)	
× born before 1950						-0.03
						(0.04)
Individuals	22213	22213	22213	22213	22213	22213
Born before cutoff year	—	72	1061	5942	36103	63142
Observations with high blood pressure	34171	34171	34171	34171	34171	34171
Total observations	84016	84016	84016	84016	84016	84016

Notes: All regressions include individual-fixed, state-fixed, wave, and census-region-wave effects as well as a quartic age trend. High blood pressure is a binary indicator that takes value one, if respondents report to have ever had high blood pressure been diagnosed, and zero else. High blood pressure is interacted with a dummy indicator that takes value one, if the individual has been born before a certain threshold level, for example., 1910, and zero else. Standard errors are clustered at the individual level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix D

Appendix to Chapter 5

Table D.1: Descriptive Statistics

	Mean	Std. Dev.	Min.	Max.	Obs.
Wage Gini	0.45	0.03	0.35	0.52	336
Total income Gini	0.48	0.02	0.40	0.54	288
Total family income Gini	0.37	0.03	0.31	0.47	288
Log life expectancy at 30	3.81	0.06	3.65	3.93	336
Mortality from CVD in 1960 × Post 1960	0.23	0.20	0	0.46	336
Median age	29.95	4.74	18	42	336
Share 45–64 in working-age population	0.32	0.03	0.22	0.39	336
Controls in 1960 × Post 1960:					
Initial log life expectancy at 30	2.16	1.88	0	3.81	336
Initial mortality not CVD	0.21	0.18	0	0.47	336
Initial share college graduates	0.03	0.03	0	0.07	336
Initial population density ($\times \frac{1}{100}$)	0.75	1.60	0	8.23	336
Initial log wages	5.80	5.03	0	10.41	336
Initial wage Gini	0.25	0.22	0	0.51	336
Initial income Gini	0.28	0.24	0	0.52	288
Initial family income Gini	0.20	0.17	0	0.41	288
Initial median age	16.35	14.30	0	33	366
Initial share 45–64 in working-age population	0.19	0.17	0	0.38	366
Sample weights:					
Initial white population ($\times \frac{1}{100000}$)	32.95	34.26	2.63	152.87	336

Table D.2: Adult Life Expectancy and Wage Inequality – OLS Estimates

	Dependent variable: wage Gini				
	(1)	(2)	(3)	(4)	(5)
Log life expectancy at 30	0.52***	1.09***	1.09***	0.64***	0.21*
	(0.05)	(0.09)	(0.09)	(0.14)	(0.12)
Controls in 1960 × Post 1960:					
Initial life expectancy		✓	✓	✓	✓
Initial mortality (not CVD)			✓	✓	✓
Initial share college					✓
Initial population density					✓
Initial income					✓
Region-year FE				✓	✓
FE & TE	✓	✓	✓	✓	✓
States	48	48	48	48	48
Observations	336	336	336	336	336

Notes: All regressions include state-fixed and time effects. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table D.3: Robustness: Different Inequality Measures

	Inequality is measured by Gini coefficient of					
	wage income		total income		total family income	
	(1)	(2)	(3)	(4)	(5)	(6)
	(a) Differences-in-differences model (2SLS)					
Log life expectancy at 30	0.90***	0.87***	1.49***	1.57**	1.39***	1.33***
	(0.32)	(0.31)	(0.56)	(0.62)	(0.54)	(0.51)
Initial inequality × Post 1960		-0.19		0.21		0.33**
		(0.16)		(0.29)		(0.14)
First-stage F–statistic	23.3	25.2	12.5	11.1	12.5	14.1
	(b) Flexible model (2SLS)					
Log life expectancy at 30	1.10***	0.95***	1.49***	1.51***	1.19***	1.16***
	(0.32)	(0.29)	(0.48)	(0.51)	(0.41)	(0.40)
Initial inequality × Post 1960		-0.36***		0.07		0.31**
		(0.13)		(0.27)		(0.13)
First-stage F–statistic	5.4	6.0	3.7	3.4	3.7	3.9
Hansen test p–value	0.9	0.9	0.7	0.7	0.2	0.3
	(c) Flexible model (LIML)					
Log life expectancy at 30	1.14***	0.97***	1.66***	1.71***	1.57**	1.50**
	(0.34)	(0.30)	(0.57)	(0.61)	(0.64)	(0.59)
Initial inequality × Post 1960		-0.37***		0.09		0.37**
		(0.13)		(0.29)		(0.16)
First-stage F–statistic	5.4	6.0	3.7	3.4	3.7	3.9
Hansen test p–value	0.9	0.9	0.7	0.7	0.4	0.4
States	48	48	48	48	48	48
Observations	336	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table D.4: Robustness: Different Measures of Population Aging

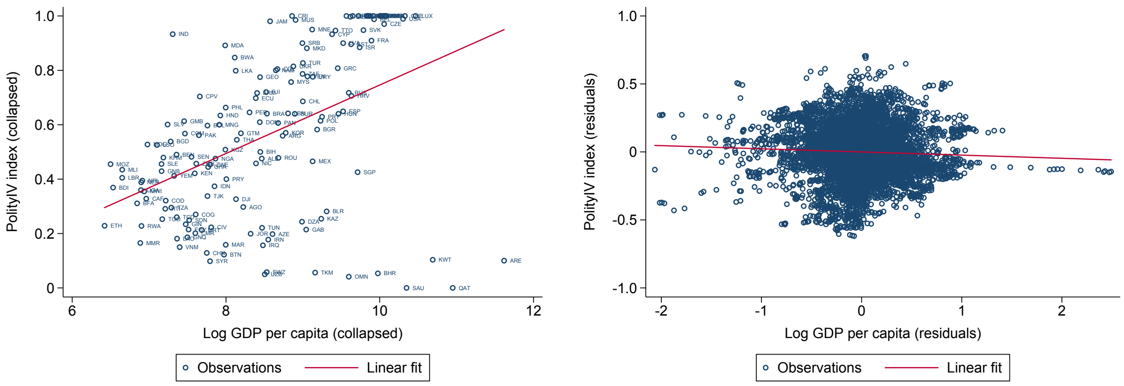
	Dependent variable: wage Gini				
	Differences-in-differences model			Flexible model	
		2SLS		2SLS	LIML
	(1)	(2)	(3)	(4)	(5)
	(a) Median age				
Median age	0.008***	0.007**	0.007**	0.009***	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Initial median age × Post 1960		-0.001	-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)
Initial inequality × Post 1960			-0.071	-0.269*	-0.270*
			(0.154)	(0.148)	(0.151)
First-stage F–statistic	18.6	19.3	20.0	3.7	3.7
Hansen test p–value	—	—	—	1.0	1.0
	(b) Share of 45 to 64 in the workforce				
Share 45–64	0.62***	0.64**	0.66**	0.64***	0.74***
	(0.23)	(0.25)	(0.26)	(0.24)	(0.28)
Initial share 45–64 × Post 1960		0.06	0.06	0.12	0.13
		(0.10)	(0.10)	(0.09)	(0.10)
Initial inequality × Post 1960			0.17	-0.02	0.00
			(0.17)	(0.16)	(0.17)
First-stage F–statistic	19.0	16.7	16.3	3.6	3.6
Hansen test p–value	—	—	—	0.7	0.8
States	48	48	48	48	48
Observations	336	336	336	336	336
Full controls	✓	✓	✓	✓	✓

Notes: All regressions include state-fixed, time, and region-year-fixed effects. Control variables are measured in 1960 and interacted with the post-1960 treatment dummy. The full set of controls comprises log initial life expectancy, initial mortality from non-cardiovascular diseases, the initial share of college graduates, initial population density, and log initial income. Standard errors are clustered on the state level and reported in parentheses. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix E

Appendix to Chapter 6

E.1 Additional Figures



(a) Unconditional Correlation (b) Residuals (w/o Fixed and Time Effects)

Figure E.1: Reproduction of Correlations Following Acemoglu et al. (2008)

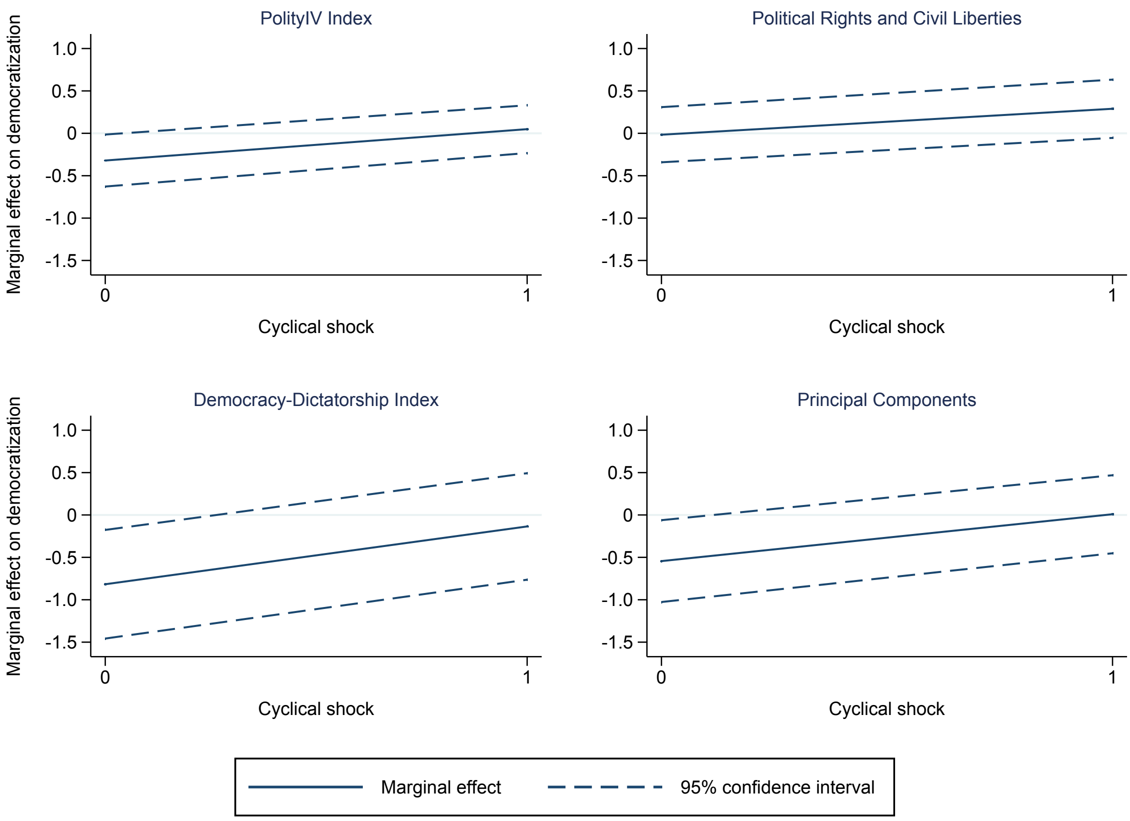


Figure E.2: Effect of Inequality on Changes in Democratic Quality

Notes: The marginal effects of negative income shocks are based on the estimates in Table 6.2.

E.2 Additional Tables

Table E.1: Descriptive Statistics

	Transitions, shocks or trends in sample	Mean	Std. Dev.	Min	Max	Obs.
		Democracy indicators				
PolityIV index (PIV)		0.72	0.32	0	1	3678
Political Rights & Civil Liberties (FH)		0.64	0.31	0	1	3563
Democracy-Dictatorship index (DD)		0.62	0.49	0	1	2971
Principal Components (PCA)		0.66	0.34	0.03	1	2856
		Democratization				
Democratization (PIV)	76	0.02	0.14	0	1	3678
Democratization (FH)	81	0.02	0.15	0	1	3480
Democratization (DD)	114	0.04	0.19	0	1	2970
Democratization (PCA)	79	0.03	0.17	0	1	2773
		Income, shocks and trends				
Income p.c.		8.92	1.17	5.82	11.98	3678
Negative cyclical shock	[332,450]	0.12	0.33	0	1	3678
Positive cyclical shock	[394,534]	0.15	0.35	0	1	3678
Negative income trend	[247,273]	0.07	0.26	0	1	3678
High inflation shock	[741,878]	0.24	0.43	0	1	3612
		Inequality and human capital				
Market Gini		0.46	0.06	0.28	0.68	3591
Share 15–24		0.18	0.03	0.09	0.26	3678
Average years of schooling		7.34	2.92	0.95	13.18	3678

Table E.2: Robustness: Multiple Imputation

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
Democratic Quality$_{t-3}$	-0.29***	-0.35***	-0.32***	-0.39***
	(0.04)	(0.05)	(0.03)	(0.05)
Shock$_{t-3}$	-0.15***	-0.11*	-0.30***	-0.23***
	(0.05)	(0.07)	(0.08)	(0.07)
Inequality$_{t-3}$	-0.22	-0.02	-0.50*	-0.35
	(0.16)	(0.17)	(0.29)	(0.25)
(Shock·Inequality)$_{t-3}$	0.33***	0.27*	0.61***	0.49***
	(0.12)	(0.15)	(0.17)	(0.16)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Shocks	448	450	349	332
Countries	128	133	129	124
Observations	3678	3575	3036	2773

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a negative cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.3: Robustness: Different Coding for Economic Shocks

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
	(a) Negative cyclical shock within the current year			
Democratic Quality$_{t-3}$	-0.29***	-0.35***	-0.32***	-0.39***
	(0.04)	(0.05)	(0.03)	(0.05)
Shock$_{t-3}$	-0.15***	-0.13	-0.29***	-0.25***
	(0.05)	(0.08)	(0.08)	(0.07)
Inequality$_{t-3}$	0.28*	0.00	-0.77**	-0.51**
	(0.15)	(0.16)	(0.32)	(0.24)
(Shock·Inequality)$_{t-3}$	0.33**	0.31*	0.64***	0.57***
	(0.13)	(0.18)	(0.18)	(0.17)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Shocks	178	175	137	128
Countries	128	133	129	124
Observations	3695	3587	3053	2784
R^2	0.14	0.12	0.20	0.16
	(b) Negative cyclical shock at least once within last five years			
Democratic Quality$_{t-3}$	-0.30***	-0.36***	-0.33***	-0.40***
	(0.04)	(0.05)	(0.03)	(0.06)
Inequality$_{t-3}$	-0.35**	-0.02	-0.88**	-0.57**
	(0.16)	(0.17)	(0.34)	(0.26)
Shock$_{t-3}$	-0.14**	-0.07	-0.31***	-0.18**
	(0.07)	(0.06)	(0.10)	(0.08)
(Shock·Inequality)$_{t-3}$	0.31**	0.17	0.65***	0.38**
	(0.15)	(0.15)	(0.22)	(0.18)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Shocks	684	685	535	507
Countries	128	133	129	124
Observations	3646	3544	3005	2742
R^2	0.15	0.12	0.21	0.16

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. In Panel (a), the shock indicator takes a value of one, if, within one year, there is a cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. In Panel (b), the shock indicator takes a value of one, if, within a time interval of five years, there is a cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.4: Robustness: Smoothing Parameter λ and Negative Cyclical Shocks

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
	(a) $\lambda = 1$			
Democratic Quality$_{t-3}$	-0.29***	-0.35***	-0.32***	-0.39***
	(0.04)	(0.05)	(0.03)	(0.05)
Shock$_{t-3}$	-0.24**	-0.14	-0.36*	-0.32*
	(0.11)	(0.13)	(0.19)	(0.16)
Inequality$_{t-3}$	-0.31**	0.00	-0.79**	-0.53**
	(0.15)	(0.16)	(0.32)	(0.24)
(Shock·Inequality)$_{t-3}$	0.53**	0.34	0.75*	0.68*
	(0.25)	(0.30)	(0.42)	(0.37)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Shocks	220	218	150	141
Countries	128	133	129	124
Observations	3678	3575	3036	2773
R^2	0.14	0.12	0.20	0.16
	(b) $\lambda = 100$			
Democratic Quality$_{t-3}$	-0.29***	-0.35***	-0.32***	-0.39***
	(0.04)	(0.05)	(0.03)	(0.05)
Shock$_{t-3}$	-0.04	-0.02	-0.22***	-0.10
	(0.06)	(0.06)	(0.08)	(0.09)
Inequality$_{t-3}$	-0.28*	0.04	-0.83**	-0.53**
	(0.15)	(0.17)	(0.32)	(0.24)
(Shock·Inequality)$_{t-3}$	0.07	0.01	0.43**	0.18
	(0.14)	(0.14)	(0.16)	(0.20)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Shocks	916	912	819	756
Countries	128	133	129	124
Observations	3678	3575	3036	2773
R^2	0.14	0.12	0.21	0.16

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. The smoothing parameter for the HP filter is set to $\lambda = 1$ in Panel (a) and $\lambda = 100$ in Panel (b). Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.5: Robustness: Different Coding for Democratization Period

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
	(a) Democratization is completed within one year			
Democratic Quality$_{t-1}$	-0.08***	-0.07***	-0.14***	-0.12***
	(0.02)	(0.02)	(0.01)	(0.02)
Shock$_{t-1}$	-0.05**	-0.09***	-0.13***	-0.09**
	(0.02)	(0.03)	(0.04)	(0.04)
Inequality$_{t-1}$	-0.02	-0.03	-0.38***	-0.24**
	(0.07)	(0.04)	(0.14)	(0.10)
(Shock·Inequality)$_{t-1}$	0.10*	0.21***	0.29***	0.22**
	(0.06)	(0.08)	(0.09)	(0.09)
Controls	✓	✓	✓	✓
Transitions	21	17	49	28
Negative Shocks	492	490	400	382
Countries	128	133	131	125
Observations	3881	3782	3290	3015
R^2	0.05	0.04	0.10	0.05
	(b) Democratization is completed within five years			
Democratic Quality$_{t-5}$	-0.47***	-0.56***	-0.47***	-0.63***
	(0.05)	(0.07)	(0.03)	(0.07)
Shock$_{t-5}$	-0.19*	-0.15	-0.32**	-0.21*
	(0.10)	(0.10)	(0.13)	(0.11)
Inequality$_{t-5}$	-0.60**	-0.32	-1.20**	-0.85**
	(0.25)	(0.29)	(0.49)	(0.40)
(Shock·Inequality)$_{t-5}$	0.42*	0.35	0.67**	0.45*
	(0.24)	(0.24)	(0.31)	(0.27)
Controls	✓	✓	✓	✓
Transitions	117	126	164	124
Shocks	403	399	298	278
Countries	126	132	124	119
Observations	3451	3342	2785	2532
R^2	0.23	0.19	0.30	0.25

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. In Panel (a), the dependent variable takes a value of one, if a positive change in democratic quality occurs which larger or equal to the respective threshold over a period of one year, or zero else. In Panel (b), the dependent variable takes a value of one, if a positive change in democratic quality occurs which is larger or equal to the respective threshold over a period of five years, and zero else. The shock indicator takes a value of one, if, within a time interval of five years, there is at least once a cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.6: Robustness: Negative Cyclical Income Shocks, Inequality, and Democratization

	Democratization indicator based on			
	PolityIV Index $\Delta \geq 0.5$	Political Rights & Civil Liberties $\Delta \geq 0.5$	Democracy-Dictatorship $\Delta = 1$	Principal Components $\Delta \geq 0.5$
	(1)	(2)	(3)	(4)
Democratic Quality$_{t-3}$	-0.29***	-0.09***	-0.32***	-0.32***
	(0.04)	(0.02)	(0.03)	(0.05)
Shock$_{t-3}$	-0.17***	-0.09**	-0.33***	-0.18***
	(0.05)	(0.04)	(0.08)	(0.06)
Inequality$_{t-3}$	-0.32**	0.02	-0.82**	-0.47**
	(0.16)	(0.06)	(0.33)	(0.22)
(Shock·Inequality)$_{t-3}$	0.37***	0.21**	0.68***	0.37**
	(0.12)	(0.09)	(0.16)	(0.14)
Controls	✓	✓	✓	✓
Transitions	76	15	114	56
Shocks	448	450	349	332
Countries	128	133	129	124
Observations	3678	3575	3036	2773
R^2	0.14	0.06	0.21	0.14

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a negative cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.7: Robustness: Evidence for Alternative Democracy Codings (Data from Acemoglu et al., 2016)

	Democratization indicator based on						
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components	Papaioannou & Siourounis (2008)	Acemoglu et al. (2016)	Principal Components (Full)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Democratic Quality$_{t-3}$	-0.36***	-0.28***	-0.33***	-0.31***	-0.33***	-0.36***	-0.40***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.05)
Shock$_{t-3}$	-0.34***	-0.15	-0.31***	-0.22***	-0.32***	-0.36***	-0.42***
	(0.10)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)	(0.11)
Inequality$_{t-3}$	-0.95***	-0.14	-0.84**	-0.52**	-0.46	-0.88***	-1.05**
	(0.30)	(0.22)	(0.33)	(0.25)	(0.29)	(0.33)	(0.40)
(Shock·Inequality)$_{t-3}$	0.72***	0.36*	0.65***	0.47***	0.72***	0.81***	0.94***
	(0.21)	(0.21)	(0.17)	(0.17)	(0.19)	(0.20)	(0.24)
Controls	✓	✓	✓	✓	✓	✓	✓
Transitions	138	118	115	88	94	144	117
Shocks	363	360	331	315	331	369	284
Countries	120	127	126	120	117	127	110
Observations	3091	3016	2974	2684	2969	3219	2472
R^2	0.24	0.18	0.21	0.17	0.24	0.25	0.19

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a negative cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.8: Robustness: Overlap Between Cyclical Shocks and Democratization

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
	(a) One-year overlap			
Democratic Quality$_{t-3}$	-0.29***	-0.36***	-0.33***	-0.40***
	(0.04)	(0.05)	(0.03)	(0.05)
Shock$_{t-2}$	-0.25***	-0.20**	-0.37***	-0.34***
	(0.07)	(0.08)	(0.10)	(0.10)
Inequality$_{t-2}$	-0.34**	0.00	-0.85**	-0.55**
	(0.15)	(0.17)	(0.33)	(0.26)
(Shock·Inequality)$_{t-2}$	0.57***	0.49**	0.83***	0.78***
	(0.16)	(0.19)	(0.22)	(0.23)
Controls	✓	✓	✓	✓
Transitions	80	86	121	84
Shocks	469	465	374	352
Countries	128	133	130	124
Observations	3779	3645	3159	2855
R^2	0.15	0.13	0.21	0.18
	(b) Two-year overlap			
Democratic Quality$_{t-3}$	-0.30***	-0.37***	-0.33***	-0.42***
	(0.04)	(0.05)	(0.03)	(0.05)
Shock$_{t-1}$	-0.20***	-0.16**	-0.34***	-0.31***
	(0.07)	(0.07)	(0.09)	(0.08)
Inequality$_{t-1}$	-0.33**	0.06	-0.90***	-0.56**
	(0.16)	(0.17)	(0.34)	(0.28)
(Shock·Inequality)$_{t-1}$	0.46***	0.40**	0.77***	0.70***
	(0.16)	(0.18)	(0.20)	(0.19)
Controls	✓	✓	✓	✓
Transitions	86	92	129	91
Shocks	481	471	389	363
Countries	128	133	130	124
Observations	3860	3692	3271	2924
R^2	0.16	0.14	0.22	0.19

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. In Panel (a), the democratization period and the economic shock overlap for one year. Explicitly, the specification allows cyclical shocks that occur during the first year of democratization to affect the overall likelihood of a democratic transition. Panel (b) allows an overlap of two years. Correspondingly, cyclical shocks during the first and second year of democratization may also affect the overall likelihood of a democratic transition. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.9: Robustness: Logistic Regression Model

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
	(a) Fixed effects			
Democratic Quality$_{t-3}$	-19.51***	-15.28***	-32.94	-11.70***
	(3.09)	(1.83)	(1205.34)	(1.65)
Shock$_{t-3}$	-4.08	-3.85	-21.97***	-14.75***
	(5.76)	(3.31)	(6.21)	(5.42)
Inequality$_{t-3}$	15.16	4.60	-23.46	-21.71
	(16.57)	(14.07)	(17.46)	(19.17)
(Shock·Inequality)$_{t-3}$	7.13	7.89	44.32***	29.88***
	(11.69)	(6.89)	(12.66)	(10.74)
Controls	✓	✓	✓	✓
Transitions	76	81	113	79
Shocks	176	148	143	142
Countries	31	34	36	31
Observations	1134	1070	1065	846
	(b) Fixed effects without lagged democratic quality			
Shock$_{t-3}$	-8.74**	-5.42*	-11.16***	-11.45***
	(3.74)	(2.98)	(3.64)	(4.11)
Inequality$_{t-3}$	-12.15*	-19.44**	-25.38***	-27.09***
	(7.28)	(7.76)	(7.14)	(9.15)
(Shock·Inequality)$_{t-3}$	17.52**	11.57*	21.49***	22.38***
	(7.47)	(6.08)	(7.26)	(8.05)
Controls	✓	✓	✓	✓
Transitions	76	81	113	79
Shocks	176	148	143	142
Countries	31	34	36	31
Observations	1134	1070	1065	846

Notes: All regressions include country and time fixed effects. Results are reported for the full fixed-effects logit specification in Panel (a) and fixed-effects logit without the lagged dependent variable in Panel (b). Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a cyclical shock of at least minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. In Panel (a), the democratization period and the economic shock overlap for one year. Explicitly, the specification allows cyclical shocks that occur during the first year of democratization to affect the overall likelihood of a democratic transition. Panel (b) allows an overlap of two years. Correspondingly, cyclical shocks during the first and second year of democratization may also affect the overall likelihood of a democratic transition. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.10: Demographic Change as Determinant of Democratization

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
Democratic Quality$_{t-3}$	-0.30***	-0.35***	-0.33***	-0.40***
	(0.04)	(0.05)	(0.03)	(0.06)
Inequality$_{t-3}$	-1.18**	-0.92	-0.79	-1.08*
	(0.49)	(0.57)	(0.81)	(0.64)
(Sh.15–24)$_{t-3}$	-1.89	-2.50*	0.49	-1.26
	(1.16)	(1.36)	(2.08)	(1.54)
(Sh.15–24·Inequality)$_{t-3}$	5.85**	5.46*	0.70	3.74
	(2.57)	(3.04)	(4.40)	(3.41)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Countries	128	133	129	124
Observations	3695	3587	3053	2784
R^2	0.15	0.12	0.20	0.16

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. Sh.15–24 measures the share of 15- to 24-year-olds in the total population. Inequality refers to the market Gini coefficient. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.11: Triple-Interaction Between Shocks, Inequality, and Share of Youths

	Democratization indicator based on			
	PolityIV Index	Political Rights & Civil Liberties	Democracy-Dictatorship	Principal Components
	(1)	(2)	(3)	(4)
Democratic Quality$_{t-3}$	-0.31***	-0.35***	-0.33***	-0.40***
	(0.04)	(0.05)	(0.03)	(0.06)
Shock$_{t-3}$	-0.61	-0.16	-0.76	-0.74
	(0.41)	(0.44)	(0.75)	(0.76)
Inequality$_{t-3}$	-1.41***	-1.00	-1.05	-1.37*
	(0.50)	(0.61)	(0.78)	(0.70)
(Sh.15–24)$_{t-3}$	-2.33**	-2.61*	-0.03	-1.87
	(1.16)	(1.43)	(2.01)	(1.71)
(Shock·Sh.15–24)$_{t-3}$	2.31	0.14	2.29	2.50
	(2.13)	(2.26)	(3.86)	(3.95)
(Shock·Inequality)$_{t-3}$	1.22	0.22	1.67	1.61
	(0.89)	(0.97)	(1.66)	(1.67)
(Sh.15–24·Inequality)$_{t-3}$	6.79***	5.66*	1.81	5.03
	(2.56)	(3.21)	(4.25)	(3.75)
(Shock·Sh.15–24·Inequality)$_{t-3}$	-4.45	0.46	-5.16	-5.44
	(4.58)	(4.99)	(8.51)	(8.59)
Controls	✓	✓	✓	✓
Transitions	76	81	114	79
Shocks	448	450	349	332
Countries	128	133	129	124
Observations	3678	3575	3036	2773
R^2	0.15	0.12	0.21	0.16

Notes: All regressions include country and time fixed effects. Log GDP p.c. and average years of schooling are added as controls for income and education. The dependent variable takes a value of one, if a positive change in democratic quality occurred that over a period of three years equaled or exceeded the respective threshold described in Section 6.2.3, and zero else. The shock indicator takes a value of one, if, within a time interval of three years, there is at least once a negative cyclical shock of minus five percent, as expressed by the cyclical component of the HP filter. Following the Ravn and Uhlig (2002) rule, the smoothing parameter for the HP filter is $\lambda = 6.25$. Sh. 15–29 measures the share of 15- to 24-year-olds in the total population. Inequality refers to the market Gini coefficient. Standard errors are clustered on the country level. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix F

Appendix to Chapter 7

F.1 Estimation Models

In order to estimate this model, five different dynamic panel models are proposed, fixed effects, random effects, bias corrected fixed effects (CFE), differences GMM (DGMM), and system GMM (SGMM), each having advantages and disadvantages. The material in this section draws strongly on Cameron and Trivedi (2005) and Bruno (2005). Throughout this paper, vectors and matrices are indicated by bold characters.

Random Effects and Fixed Effects. The random effects model can be written as

$$y_{i,t} = \alpha\, y_{i,t-1} + \mathbf{x'}_{i,t-1}\boldsymbol{\theta} + \eta_i + \epsilon_{i,t}.$$

For a static version of the model (that is, without the lagged dependent variable), consistency requires regressors be uncorrelated with individual-specific effects,

$$E[\mathbf{x}_{i,t-1}(\eta_i + \epsilon_{i,t})] = 0.$$

Under these assumptions, feasible GLS which is more efficient than OLS, yields consistent estimates. If a lagged dependent variable is added to the regression equation, $E[y_{i,t-1}(\eta_i + \epsilon_{i,t})] \neq 0$, because by definition $E[y_{i,t-1},\, \eta_i] \neq 0$. Identification requires the regressors to be uncorrelated with the unobserved individual fixed effect. This represents a restrictive assumption whose failure leads to biased estimates. Hence, for $\alpha > 0$ the autoregressive parameter will be overstated. In contrast, explanatory variables with $\beta_j > 0$ will exhibit downward bias. The direction of bias in explanatory variables is ambiguous, because the inclusion of an interaction term allows for non-monotonous overall effect.

The fixed effects or within estimator is obtained by subtracting the time-averaged model from the original model

$$y_{i,t} - \bar{y}_i = \alpha\,(y_{i,t-1} - \bar{y}_i) + (\mathbf{x_{i,t-1}} - \mathbf{\bar{x}_i})'\boldsymbol{\theta} + (\epsilon_{i,t} - \bar{\epsilon}_i),$$

where time-average variables are indicated by upper bars. For notational convenience, explanatory variables, time effects and controls are now summarized by the vector $\mathbf{x}_{i,t-1}$. Consistency of the fixed effects estimator requires either the number of countries $N \to \infty$ or the count of observation periods $T \to \infty$ and differences of error terms from their time average to be uncorrelated with the respective differences of the right-hand side variables, or formally

$$E[(\epsilon_{i,t} - \bar{\epsilon}_i)|(\mathbf{x_{i,t-1}} - \mathbf{\bar{x}_i})] = 0 \quad and \quad E[(\epsilon_{i,t} - \bar{\epsilon}_i)|(y_{i,t-1} - \bar{y}_i)] = 0. \tag{F.1}$$

In a static model (without lagged dependent variable as explanatory variables), a sufficient condition for equation (F.1) to hold is strict exogeneity of the regressors. In the fixed effects model, the problem of correlation between the error and the unobserved individual effect is dealt with mechanically when the unobserved fixed effect is eliminated by the within-transformation. Correspondingly, a potential omitted variable must be correlated with the demeaned regressors to bias the results. In the presence of time effects, omitted variables therefore have to be both time- and country-varying, which restricts the set of potential confounds considerably. Estimates obtained with these estimators can be informative regarding the existence of a non-monotonic effect in a differences or level model, respectively. However, including a lagged dependent variable automatically introduces bias, because $\bar{y}_i$ and $\bar{\epsilon}_i$ are by definition correlated through the averaged model. Consequently, the second part of condition (F.1) is necessarily violated. For fixed effects to be nonetheless consistent $\bar{\epsilon}_i$ has to become very small relative to $\epsilon_{i,t}$, which is the case for $T \to \infty$. This is clearly not the case for the short panel at hand. Correspondingly, the autoregressive parameter α will be downward biased, while the bias for the effect of democracy and equality is ambiguous due to the inclusion of an interaction term.[1]

Another problem is the requirement of strong exogeneity of the error terms; that is, all regressors, including lagged and future realizations, must be uncorrelated with current-period error terms. This assumption is substantially stronger as the weak exogeneity condition needed in panel GMM models discussed below, which only requires lagged instruments to be uncorrelated with the current-period error terms.

Bias-Corrected Fixed Effects. One way to deal with the inherent endogeneity problem that is introduced by the lagged dependent variable is to perform a bias correction. For instance, Bun and Kiviet (2003) propose a bias-corrected fixed effects estimator for balanced

[1]For further information see Nickell (1981).

panels that has been generalized to the unbalanced case by Bruno (2005). The resulting estimator approximates and eliminates the bias that results from the fixed panel length. Consider the standard dynamic model from above

$$y_{i,t} = \alpha\, y_{i,t-1} + \mathbf{x_{i,t}}'\boldsymbol{\theta} + \epsilon_{i,t},$$

Stacking observations across countries and over time yields

$$\mathbf{y} = \mathbf{D}\boldsymbol{\eta} + \mathbf{W}\boldsymbol{\delta} + \boldsymbol{\epsilon},$$

where $W = (\mathbf{y_{-1}} \vdots \mathbf{X})$ is a $(NT \times k)$ matrix of stacked observations; $\mathbf{D} = \mathbf{I_N} \otimes \boldsymbol{\iota_T}$ is a $(NT \times N)$ matrix of individual dummies (where $\boldsymbol{\iota_T}$ is a vector of unity elements); $\boldsymbol{\delta} = (\alpha \vdots \theta')'$ denotes the $(k \times 1)$ vector of coefficients; and $\mathbf{y}$ as well as $\boldsymbol{\eta}$ are stacked vectors of the dependent variable and the unobserved country fixed effects.[2]

Define a selection rule $s_{i,t}$ that only selects those observations that are usable for the dynamic panel, that is, those with observations for the current as well as the lagged period. Stacking the selection indicator into a matrix of dimension $(NT \times NT)$, the dynamic model can be written as

$$\mathbf{Sy} = \mathbf{SD}\boldsymbol{\eta} + \mathbf{SW}\boldsymbol{\delta} + \mathbf{S}\boldsymbol{\epsilon}.$$

The fixed effect estimator for this unbalanced panel is then given by

$$\boldsymbol{\delta} = (\mathbf{W}'\mathbf{M}_s\mathbf{W})^{-1}\mathbf{W}'\mathbf{M}_s\mathbf{y}$$

with

$$\mathbf{M}_s = \mathbf{S}\big[\mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{SD})^{-1}\mathbf{D}'\big]\mathbf{S}$$

being the symmetric and idempotent $(NT \times NT)$ transformation matrix that removes the unobserved country-specific fixed effects.

The bias of the fixed effects estimator can then be approximated by

$$\mathbf{c}_1(\bar{T}^{-1}) = \sigma_\epsilon^2 tr(\mathbf{\Pi})\mathbf{q_1},$$

$$\begin{aligned}\mathbf{c}_2(N^{-1}\bar{T}^{-1}) = -\sigma_\epsilon^2 \big[&\mathbf{Q}\bar{\mathbf{W}}'\mathbf{\Pi}\mathbf{M}_s\mathbf{W} + tr(\mathbf{Q}\bar{\mathbf{W}}'\mathbf{\Pi}\mathbf{M} - s\bar{\mathbf{W}})\mathbf{I}_{k+1} \\ &+ 2\sigma_\epsilon^2 q_{11} tr(\mathbf{\Pi}'\mathbf{\Pi}\mathbf{\Pi})\mathbf{I}_{k+1}\big]\mathbf{q_1},\end{aligned}$$

[2]Note that the corresponding elements of the vector $\boldsymbol{\theta}$ are 0 for $\mathbf{x_{i,t}}$, because contemporaneous effects are excluded from the model. This notation is chosen to follow Bruno (2005) as closely as possible.

and

$$\mathbf{c}_3(N^{-1}\bar{T}^{-2}) = \sigma_\epsilon^4 tr(\mathbf{\Pi})\Big[2q_{11}\mathbf{Q}\bar{\mathbf{W}}'\mathbf{\Pi}\mathbf{\Pi}'\bar{\mathbf{W}}\mathbf{q_1} + \big[(\mathbf{q_1}'\bar{\mathbf{W}}'\mathbf{\Pi}\mathbf{\Pi}'\bar{\mathbf{W}}\mathbf{q_1}) + q_{11}tr(\mathbf{Q}\bar{\mathbf{W}}'\mathbf{\Pi}\mathbf{\Pi}'\bar{\mathbf{W}}) + 2tr(\mathbf{\Pi}'\mathbf{\Pi}\mathbf{\Pi}'\mathbf{\Pi})q_{11}^2\big]\mathbf{q_1}\Big],$$

where $\mathbf{Q} = \left[E(\mathbf{W}'\mathbf{M}_s\mathbf{W})\right]^{-1} = \left[\bar{\mathbf{W}}'\mathbf{M}_s\bar{\mathbf{W}} + \sigma_\epsilon^2 tr(\mathbf{\Pi}'\mathbf{\Pi})\mathbf{e_1}\mathbf{e_1}'\right]^{-1}$; $\bar{\mathbf{W}} = E(\mathbf{W})$; $\mathbf{e}_1 = (1, 0, \ldots, 0)'$ is the $(k \times 1)$ unit vector; $\mathbf{q}_1 = \mathbf{Q}\mathbf{e}_1$; $q_{11} = \mathbf{e}_1'$; $\mathbf{L_T}$ is a $(T \times T)$ matrix for which the first lower sub-diagonal is unity and all other entries take value zero; $\mathbf{L} = \mathbf{I}_N \otimes \mathbf{L}_T$; $\mathbf{A}_T = (\mathbf{I}_T - \alpha\mathbf{L}_T)^{-1}$; $\mathbf{A} = \mathbf{I}_N \otimes \mathbf{A}_T$; $\mathbf{\Pi} = \mathbf{M}_s\mathbf{L}\mathbf{A}$; and r is an indicator whether the observation is non-missing.

This leads to bias approximations with increasing accuracy of:

$$\mathbf{B}_1 = \mathbf{c}_1(\bar{T}^{-1});\ \mathbf{B}_2 = \mathbf{B}_1 + \mathbf{c}_2(N^{-1}\bar{T}^{-1});\ \mathbf{B}_3 = \mathbf{B}_2 + \mathbf{c}_3(N^{-1}\bar{T}^{-2}).$$

Using a consistent estimate for σ_ϵ^2 and α, one can estimate the approximate bias $\hat{\mathbf{B}}_i$ and obtain the bias-corrected fixed effects estimator

$$\boldsymbol{\delta}_i = \boldsymbol{\delta}_{FE} - \hat{\mathbf{B}}_i,\ i = 1,\ 2,\ 3.$$

Estimates for α can be obtained using either the Anderson and Hsiao (1981), the Arellano and Bond (1991) or the Blundell and Bond (1998) estimator. For the model at hand, the Arellano-Bond estimator is used with the computed variance

$$\hat{\sigma}_{AB}^2 = \frac{(\mathbf{y} - \mathbf{W}\boldsymbol{\delta}_{AB})'\mathbf{M}_s(\mathbf{y} - \mathbf{W}\boldsymbol{\delta}_{AB})}{(N - k - T)}.$$

In this study, the variance-covariance matrix is estimated using bootstrap procedures with 100 repetitions. Therefore, the corrected estimator yields more credible results than the standard OLS based dynamic panel models.[3]

Differences GMM. Alternatively, the inconsistency problem of standard panel regressors can be resolved by applying an IV variant of a first-differences OLS estimator. The differences GMM estimator identifies the coefficients of interest from changes in the explanatory variables (Arellano and Bond, 1991). For identification, the estimation exploits the lagged values of the explanatory variables as instruments for the endogenous regressors. Taking the first difference of the model presented in equation (1) in the main text yields

$$y_{i,t} - y_{i,t-1} = \alpha\left(y_{i,t-1} - y_{i,t-2}\right) + (\mathbf{x}_{i,t-1} - \mathbf{x}_{i,t-2})'\boldsymbol{\theta} + (\epsilon_{i,t} - \epsilon_{i,t-1}), \qquad t = 3, ..., T.$$

[3]For more information see Bun and Carree (2005) and Bun and Kiviet (2006).

Instrumenting $(y_{i,t-1} - y_{i,t-2})$ with the level lags $(y_{i,t-2}, y_{i,t-3}, \dots)$ removes the mechanical bias between the difference of the lagged dependent variable and the error term.[4] The second level lag is a good instrument, because it is correlated with the lagged difference $(y_{i,t-1} - y_{i,t-2})$ through the level $y_{i,t-2}$ and unrelated to the differenced error term. The count of available instruments increases with the time period t such that the model is overidentified.

Using all potential instruments gives the differences GMM (Arellano-Bond) estimator that yields consistent estimates under the moment conditions

$$E[y_{i,s}(\epsilon_{i,t} - \epsilon_{i,t-1})] = 0 \quad with \quad s \leq t-2.$$

A typical and well-known problem that may arise from using lagged values as instruments is that these may only be weakly correlated after the country fixed effects are removed. In other words, instruments might be weak because either α moves toward unity or as the relative variance of the fixed effects η_i increases.[5] This problem is more severe in more persistent time series. In the case of weak instruments, the estimates for the autoregressive parameter are biased toward the fixed effects estimator. Since α is overestimated in a standard dynamic fixed effects model, the resulting coefficient can be seen as a lower bound for the true parameter.

System GMM. More efficient results can be achieved, if one uses the additional moment conditions

$$E[(\eta_i + \epsilon_{i,t})(y_{i,s} - y_{i,s-1})] = 0 \quad with \quad s \leq t-1.$$

For these additional conditions to be satisfied, a simple stationarity condition is required: deviations from the initial steady state must be uncorrelated with the level itself. Under these assumptions, not only the level lags but also the lagged first difference can be used as instruments, which may result in substantial efficiency gains. Under a stationarity assumption regarding the dependent variable, this so-called system GMM estimator by Blundell and Bond (1998) is preferable to the differences GMM estimator, because it is more efficient and provides more stable results for highly persistent variables, that is, for estimates of α close to unity. Furthermore, the estimator can better accommodate for moderately high values of the autoregressive parameter and relatively low counts of time periods T.

In addition to the identifying assumptions of differences GMM (DGMM), system GMM (SGMM) also demands a stationarity assumption that requires changes in the dependent variable be uncorrelated with the country fixed effect. This condition likely holds only

[4]Alternatively, it is also possible to use the lagged difference as an instrument. In practice, however, this often leads to weak instruments so that level lags are the preferred choice. For more information see Anderson and Hsiao (1981) and Arellano and Bond (1991).

[5]For further information see Blundell and Bond (1998).

approximately in the present application, as there are cases where democratic transitions lead to a change in the long-run steady state.

This is not too problematic, if countries are sufficiently close to their long-run steady state at the end of the panel. As a heuristic approach to test whether stationarity may be violated, SGMM is run on a subsample of the data, where, one after another, the first two, the last two, and the first as well as the last sample periods are removed to see whether the resulting regression coefficients are substantially affected by restricting the analysis to particular subsamples. The results reveal that the coefficient estimates do respond to these sample changes with the largest effects emerging when the first two periods are removed and only little effect when the last two periods are removed. This indicates that the stationarity assumption required by the SGMM is not necessarily satisfied. Therefore, DGMM is expected to provide more reliable estimates than SGMM. Nevertheless, results for both estimators are reported for completeness. Moreover, this condition is probably not crucially violated, if the estimates of DGMM and SGMM are reasonably close to each other, because the differences model does not rely on the stationarity assumption.

In order to further reduce the danger of weak instruments in the SGMM specification, also time-invariant variables are included in the level regressions in order to reduce the relative variance of the country fixed effects to the idiosyncratic error term. One first indication whether instruments are weak is a comparison of results of differences and system GMM. The parameter estimate of SGMM is a weighted sum of the difference and level equation with more weight being put on the differenced model, if identification is strong. Therefore, quantitatively similar estimates of both models are an indication for instrument relevance.

The employed number of instruments is relevant for both DGMM and SGMM models. First, consistency of the estimated parameters is not affected, but efficiency increases with the number of moments available for estimation. Increasing the number of instruments can also lead to biases in the Hansen J-test whether instruments are statistically different from each other in the case of overidentification, however. Moreover, using lags that have little explanatory power for contemporaneous variables can weaken the instrument set. Hence, there is a trade-off between the efficiency of results and the strength of instruments as well as the validity of the test statistics for the validation of identification assumptions. According to Roodman (2009), as a rule, the number of instruments should not exceed the number of cross-sectional units (here: countries) in the sample.[6]

In the empirical analysis, we conduct Granger (1969, 1980) tests of different lag specifications in combination with the Akaike and Bayesian Information Criteria to determine which lags are most informative and prima facie causal. Finally, Arellano and Bover (1995) note that the choice of the transformation matrix that removes the

[6]Note that this rule of thumb is not conservative such that instrument counts close to the respective threshold are no guarantee that test statistics are not biased.

unobserved individual fixed effects does not matter, if the complete instrument set is used. However, as it is necessary to limit the number of lags, the choice of the transformation matrix may be of importance. For the baseline specifications, the forward orthogonalized deviations method is used, because it eliminates less data than first differences.[7]

F.2 Additional Robustness Checks

This section reports the results of extensive robustness checks in order to test whether the results are sensitive to the use of alternative indicators of institutional quality, to different measures of democracy or equality, or to different sets of control variables. Moreover, results are presented for standard errors that are corrected for multiple imputation of the reversed Gini coefficients, as well as for different lag specifications and transformations in the context of the GMM estimators. Finally, the baseline interaction of democracy and equality is tested against additional interactions with income per capita or years of schooling. The respective Tables are reported below.

Alternative Measures of Institutional Quality. The baseline results in Tables 1 and 2 indicate an interaction between democracy, measured by Constraints on the Executive, and equality, regardless of whether institutional quality is measured by the Economic Freedom index or the composite index of institutional quality. To investigate the robustness of this finding, we also replicated the analysis using alternative indices. In the main text, Table 4 presents robustness results obtained with the bias-corrected fixed effects estimator. Tables F.4 and F.5 below show the respective results for the DGMM estimator and for alternative estimators, respectively. The first alternative measure of institutional quality is the Civil Liberties index. However, it should be noted that this measure includes components that are more closely linked to political liberties than economic freedom. Nevertheless, the main finding of a significant interaction between executive constraints and equality remains robust.[8] Alternative measures for the quality of economic institutions are the regulations of credit, labor and business, or the soundness of money as a measure of reliability of the economic environment. Both of these measures are taken from subindices underlying the Economic Freedom index. Estimates for these institutional measures as dependent variable also deliver evidence for a significant interaction between democracy and equality.[9] From the ICRG, alternative measures of the quality of economic institutions are the quality of property rights, or the protection against corruption. However, these variables are only available for a substantially shorter time span (1985–2010), which limits the variation over time. Also for these measures, the estimates systematically reveal

[7]See Hayakawa (2009) for a comparison of different transformation techniques.
[8]See Table F.5, Panel (a).
[9]See Panels (b) and (c) of Table F.5.

a positive interaction term, which is not statistically significant in all cases; however, presumably due to the limited variation that can be used for estimation.[10]

Alternative Measures of Political Institutions. In order to gauge the importance of the particular democracy indicator used in the estimation, we also conducted robustness checks with alternative measures of political institutions. Table 5 in the main text presents robustness results obtained with the bias-corrected fixed effects estimator. Table F.7 below shows the respective results for the DGMM estimator, and Table F.6 presents additional results. The robustness analysis uses various measures of political institutions. In parallel to the methodology for institutional quality, we constructed a composite measure based on a principal component analysis of political institutions from the Constraints on the Executive index and of the political competition component of the democracy index by Vanhanen and Lundell (2014). Alternative measures of democracy are the PolityIV composite index (for democracy and autocracy), the Political Rights index, a composite measure of political institutions based on the principal component of the index by Vanhanen and Lundell (2014), the PolityIV index, and the Political Rights index, or the composite index as well as the political competition component of the index by Vanhanen and Lundell (2014). In related studies, some scholars use a discrete democratic transition variable, which takes a value of 1, if a country democratizes, and 0 otherwise, instead of a continuous measure (see, for example, Persson and Tabellini, 2006, and Papaioannou and Siourounis, 2008). A drawback of such a measure is that identification is based only on substantial changes in political institutions that occur during the observation period, while many such changes and transitions to democracy have already taken place before the time institutional indicators are available. Moreover, constructing an appropriate reference group is non-trivial. With reference to the motivating figures in the introduction, the first robustness check uses the binary Democracy-Dictatorship indicator of Cheibub, Gandhi, and Vreeland (2010). The results are qualitatively and quantitatively similar to the baseline results and reveal a positive interaction term between democracy and equality.[11] Regardless of the measure of democracy used, the estimation results deliver a positive interaction effect between democracy and inequality on the quality of institutions.[12] Overall, these results also indicate that the finding of a non-monotonic effect of democracy in relation to equality is not sensitive to the particular democracy indicator that is used in the estimation.

[10] The results are reported in Panels (d) and (e) of Table F.5. Mixed findings are obtained for rule of law (the Law and Order index) as dependent variable. Unreported estimates deliver mostly insignificant coefficients for the interaction term, however, which also differ quantitatively across level and differences models. This might have several reasons, including the fact that the construction of the rule of law index is rather intransparent, as it is based on several subindices that do not necessarily reflect aspects of economic institutions and that are not available independently. This has raised doubts about what is actually measured with these indices (see, for example, Voigt, 2013, for a detailed discussion of this point).

[11] See Table F.6.

[12] The results for Economic Freedom as measure of institutional quality are contained in Table F.8, whereas the results for the composite index of institutional quality are contained in Table F.9.

Also the results for the static panel model replicate for diverse measures of political institutions.[13]

Robustness with Respect to the Inequality Measure. The SWIID data set by Solt (2009, 2016b) uses multiple imputation techniques in order to reduce the number of missing values in the data set. To account for concerns about the quality of the inequality measure, we conducted several robustness checks. In the main text, the results of these checks obtained with the bias-corrected fixed effects estimator are contained in Table 6. Table F.11 shows the respective results for the DGMM estimator.

First, in order to alleviate the concern that the results might be driven by the interpolation procedure, which might add noise to the measure, we conducted a robustness check using a binary measure of inequality in addition to using a binary measure of democracy. The results are unaffected.[14] In light of the multiple imputation, treating imputations as regular data points disregards the uncertainty regarding the imputed value, which is itself a random variable. Adjusting for this fact usually results in larger standard errors and less significant findings. To account for this issue, we conduct robustness checks that accommodate for imputation procedures. To enable corrections for multiple imputation, from version 4.0 onward the SWIID data set provides 100 values for the Gini coefficient for every country-year cell, instead of just a single value. These multiple observations for each country-year cell can be used to estimate the coefficients up to 100 times, for every single Gini realization, respectively. The final estimates for coefficients and standard errors are obtained by taking the average over all estimated coefficients. This procedure adequately reflects the underlying uncertainty in the equality measure that has been introduced by the imputation procedure. For the problem at hand, all available 100 Gini realizations are used, even though a smaller number is typically already sufficient. Intuitively, the baseline regressions used the mean over all 100 realizations and estimated the respective model. The alternative procedure estimates each model 100 times and takes the mean over the estimated parameters thereafter.[15] While accounting for multiple imputation mainly affects standard errors, the parameter estimates can vary compared to the baseline, because the estimated coefficient for the average reversed Gini need not be identical to the mean over all 100 Gini data points. Nevertheless, the results regarding the interaction effect between democracy and inequality are virtually unaffected.[16] This suggests that the main finding of non-monotonicity is not driven by neglecting the variation in the equality

[13]See, for instance, Table F.10 for results for the standard PolityIV composite index (for democracy and autocracy).

[14]See Table F.12.

[15]In practice, this is done in STATA using the "mi estimate" prefix before running the desired regression type. For a more technical discussion see Rubin (1996).

[16]The results for replicating the baseline regressions using the multiply-imputed reversed Gini data are contained in Table F.13. The estimated autoregressive coefficients are very similar to the baseline estimates from Table 1. The estimated parameters for the reversed Gini coefficients are somewhat closer to zero than in the main specifications and become in some cases insignificant.

time series due to multiple imputation.[17] In particular, the respective threshold levels of equality and democracy for a positive marginal effect the estimates are quantitatively almost unaffected compared to the baseline results. Taken together, the results from this analysis indicate that the estimated interaction effect does not hinge on standard error corrections for the Gini time series.

Instead of measuring economic inequality using net incomes, one might argue that it is more appropriate to use gross income inequality as proxy for the de facto political power and, hence, the relevant determinants of economic institutions. Both measures, the net and gross Gini indices, are highly correlated in the SWIID data set, however.[18] Correspondingly, the estimation results are very similar when using a measure of equality based on the reversed gross Gini coefficient.[19]

Economic inequality is also reflected by the distribution of skills that are available for production. Therefore, an alternative measure of equality is provided by the reversed human capital Gini coefficient described in Section 3.3 of the paper. As in the main results, the effect of equality and democracy is negative once the model allows for heterogeneous effects through an interaction term, while the interaction between democracy and equality is positive throughout all specifications—even though not always significant. The estimated parameters for the direct effects of democracy, equality, and the interaction are quantitatively somewhat smaller than in the baseline.[20]

Alternatively, we estimated the empirical model separately on subsamples that were split by the level of equality (into the lowest quintile, the three intermediate quintiles, and the highest quintile) without an interaction term. Consistent with the earlier findings, the effect of democracy on institutional quality is negative (although not significant) in the subsample in which equality is lowest, whereas the effect gets more positive in the subsamples reflecting intermediate inequality, and is most positive in the subsample where equality is highest.[21] The positive effect is largest and significant in the subsample that exhibits the highest level of equality, providing additional evidence for the heterogeneous effect of democracy.

A final set of robustness checks uses the (reversed) top-10-percent income shares constructed by Piketty (2014) as measure for inequality. Again, the interaction between

[17]To the knowledge of the authors only few existing papers account for multiple imputation or interpolation of inequality data. To the extent that the effect for many variables might be overstated without accounting for the imputation noise, the analysis in this study also provides a contribution in this respect by applying more extensive corrections for standard errors than what is standard in the literature. At the same time, the analysis abstracts from the problem of sample selection that also arises from imputation procedures, if the pattern of missing observations is not random and not adequately modeled in the imputation procedure, see Cameron and Trivedi (2005).

[18]See Table 8 in the Appendix of the paper.

[19]See Table F.14.

[20]The results are contained in Table F.15. Interestingly, the findings are consistent with recent evidence by Castelló-Climent and Doménech (2014), who find divergent trends in income inequality and education inequality.

[21]See the results in Tables F.16 and F.17.

democracy and equality is positive throughout but not always significant, probably partly due to the small number of countries for which this information is available.[22]

Taken together, the results deliver a coherent pattern of heterogeneity in the effect of democracy on institutional quality that is related to equality.

Robustness of GMM Results. A potential concern with the results from the GMM estimators is that the findings might be influenced by the choice of the transformation that removes the country fixed effect, or by the choice of the instrument set. Robustness checks in this direction have been conducted with forward orthogonalized deviations (FOD) and first differences (FD) as choices of the transformation matrix to remove the unobserved fixed effect, and different specifications of lags that are used for identification. The estimates appear not to be particularly sensitive to the use of alternative sets of lags and vary in a very moderate range for the autoregressive parameter, the democracy index as well as the interaction term. Consistent with the baseline results, there is a positive interaction effect throughout all difference and system GMM specifications, which is quantitatively similar to the baseline findings of Table 1 and significant in almost all specifications.[23]

Another potential concern with the GMM estimates is weak identification of the coefficients of interest due to the use of inflated instrument sets. Extensive robustness checks with more parsimonious sets of instruments confirm the finding of a robust positive interaction effect.[24]

Robustness to Alternative Estimation Methods. The estimation of linear models might be overly restrictive in light of the fact that institutional quality is measured as index on a discrete grid. Alternatively, we estimated the model using interval regression techniques that account for the clustered measurement of a continuous latent variable in discrete bins. The results are largely unaffected by these modifications.[25]

[22]The results are presented in Table F.18.

[23]An overview is contained in Table F.19. Detailed results are reported in Tables F.20 and F.21. Panel (a) reports different specifications for DGMM, Panel (b) for SGMM. Columns (1) and (2) replicate the results from the baseline specification; Columns (3) and (4) add one additional lag dimension compared to the baseline; Columns (5) and (6) use one lag dimension less, respectively; and the instrument set is collapsed in Columns (7) and (8). The AR(2), Hansen J-test statistics and the difference-in-Hansen test indicate that endogeneity of instruments is not a threat for the consistency of the estimates except potentially for the collapsed instrument set. This last finding may indicate that very far lags have little explanatory power and thus render the instrument set weak or that the estimated variance-covariance matrix becomes somewhat unstable due to the inclusion of many similar and strongly correlated (but potentially uninformative) right-hand side variables.

[24]Table F.22 in the Appendix contains results for different minimalist sets of instruments for the lagged dependent variable as well as for the explanatory variables, which contain substantially fewer instruments than the rule of thumb of the number cross-sectional units in the sample suggested by Roodman (2009).

[25]See Tables F.23 and F.24 for details.

Different Sets of Controls. Another dimension for robustness checks concerns the specification of the estimation framework in terms of the included control variables. For robustness, we estimated the empirical model model without any further controls except country-fixed and time effects, or with additional controls. These specifications provide robustness checks in two dimensions. The parsimonious specifications are likely to be most affected by potential bias from relevant omitted variables. At the same time, these specifications suffer least from problems of endogeneity induced by control variables ("bad controls"). The reverse holds for the extensive specifications. Regardless, the specification of the estimation framework in terms of controls does not appear to affect the results and leaves the finding of a significant interaction effect unaltered.[26]

Given the emphasis on the distributive conflict in motivating the potential role of economic equality (and its interaction with democracy), redistribution might be a potential omitted factor that drives the estimation results. However, controlling for redistribution (measured in terms of the difference between the gross and net Gini index, or the share of central government revenues and expenditures including social security as percentage of GDP) does not affect the findings.[27]

Similarly, accounting for income growth as additional control above and beyond income levels leaves the results unaffected.[28]

Overidentification: Adding Interactions between Democracy and Other Variables. A final robustness check is to compare the results obtained with the baseline specification to findings for specifications that include alternative interactions of democracy with income as well as with human capital in addition to the interaction with equality. The purpose of this exercise is twofold. First, these estimates provide a sort of overidentification test to examine whether the interaction effects found so far potentially take up heterogeneous effects of democracy in some other dimension than inequality. Second, the specifications conduct a horse race between different channels that affect institutional quality. This horse race is motivated by the arguments forwarded by Lipset (1959), who suggested that a multitude of factors might be relevant for institutions to work successfully. The estimates also account for other factors that have been identified in the previous literature. For example, Acemoglu, Johnson, Robinson, and Yared (2008, 2009) investigate

[26]Tables F.25 and F.26 report the respective findings. Panel (a) of the Tables shows estimated coefficients of specifications without any further controls except country-fixed and time effects. Panel (b) includes the same controls as the specifications in Table 1 (log income per capita, average years of schooling, as well as oil producer and former socialism in the level equations) and adds log population size; a colonial history dummy, which is unity, if the country was a former colony, and zero else, as well as ethnic polarization in the level regressions; an inflation dummy that takes a value of 1 for price changes larger than 4 percent and 0 otherwise; and a dummy for deflation, which is equal to unity whenever the inflation rate is negative, and zero otherwise. The reason for using binary measures is the huge variation as some countries in the sample experienced periods of extremely high inflation (for example, hyper-inflation in the case of Argentina), while most of the countries had a moderate development of prices.

[27]Detailed results are available upon request.

[28]Respective estimation results are reported in Table F.27.

the importance of income for political institutions, while Murtin and Wacziarg (2014) and Fortunato and Panizza (2015) identify human capital as a central determinant of institutional quality. Up to this point, in the literature, these channels have been investigated in isolation but have not been compared to each other in terms of their relevance in the same estimation framework. The results of this exercise indicate that including additional sources of heterogeneity does not affect the point estimates, in particular those of the interaction term between equality and democracy (Eq×Demo).[29] Moreover, standard errors remain almost unaffected when the additional interactions are added. In addition, the results indicate no evidence for a heterogeneous effect of income and democracy, as the respective interaction (GDP×Demo) is insignificant throughout all regressions and does not affect the interaction of democracy and equality quantitatively. Similarly, the heterogeneous effect for average years of schooling and democracy (HC×Demo) is not significantly different from zero. The finding that democracy has a heterogeneous effect on institutional quality conditional on the degree of economic equality is robust to accounting for alternative sources of heterogeneity.

Additional Results: Redistribution, Stability of Democracy, and Effect Heterogeneity. According to several of the theories motivating the empirical analysis mentioned in the Introduction, the findings might be interpreted as the consequence of an influential rich elite inducing lower institutional quality in democracy in order to protect itself from excessive distributive pressure. If this is correct, one would expect the (seemingly counterintuitive) result that tax revenues and redistribution is lower in democracies in which inequality is high. In other words, in an estimation framework with redistribution as dependent variable, one would expect a negative interaction effect between democracy and equality. In fact, there is some tentative evidence pointing in this direction.[30]

A second set of additional results addresses the question of stability of democracy. Ultimately, the arguments mentioned in the Introduction also imply that democracies might become unstable, if inequality becomes too large, referring to a hypothesis that goes

[29]The results are reported in Table F.28. Columns (1)–(4) contain the results for bias-corrected fixed effects Columns (5)–(8) those from DGMM estimators, and Columns (9)–(12) contain those from SGMM estimators. The first column of each block replicates the baseline results. Subsequently, an interaction with income or human capital is added, respectively, and in the last column of each block, all three interactions are estimated jointly. For the sake of brevity, results for random or fixed effects are not reported. The findings are similar and available upon request.

[30]Table F.29 presents the estimation results regarding this conjecture, using the the share of expenditures on social security funds or the share on social spending as percentage of GDP as dependent variables. This information is only available for a small sample of developed countries (data are from the OECD or IMF data bases, respectively), which limits the usefulness in the estimation. Nevertheless, throughout all specifications, the interaction term between democracy and equality has a negative sign; however, with the exception of bias-corrected fixed effects, the coefficient estimates are not statistically significant. Additional unreported results using a relative political extraction indicator for industrial countries by Hendrix (2010) and Arbetman-Rabinowitz et al. (2013) as measure for taxation and redistribution reveal a positive interaction effect, but the coefficient is estimated with insufficient precision to deliver statistically significant results.

back to Lipset and that has been expressed in modified form more recently by Piketty (2014).[31] To investigate the role of inequality for the stability of democracy, we re-estimate the empirical framework with an interaction between democracy and equality but with the level of political institutions as dependent variable. Because political instability and major changes in political institutions are infrequent, the estimation is conducted on baseline sample 1970–2010 in five-year intervals, as well as for a longer period from 1870–2010 in ten-year intervals. The results are indicative of a positive interaction effect of democracy and equality on political stability, particularly over the longer horizon.[32] We view this as evidence that is suggestive, or at least not inconsistent, with the theories underlying the main hypothesis of this paper, that democratic institutions are self-reinforcing, but that this effect is stronger the greater the level of equality in society.

In an attempt to investigate the possibility of heterogeneity in the interaction effect, we conducted the estimation separately for countries that democratized before and after 1974, following the classification by Huntington (1993). While the interaction appears positive throughout, the coefficient estimate is indeed somewhat larger and more significant for the sample that democratized after 1974 (the "third wave" of democratization), indicating the interaction between a democratic political regime and an equal distribution of income had a particularly large impact on institutional quality in countries where the institution building process was on the way or not yet finished.[33] An alternative dimension for heterogeneous effects is the distinction into OECD and non-OECD countries. Also here, the interaction effect is positive throughout. The interaction appears to be somewhat larger and more significant among the OECD countries for the Economic Freedom index, whereas the interaction appears somewhat larger for non-OECD countries for the composite measure as dependent variable.[34]

[31]In Lipset‘s words: "[f]rom Aristotle down to the present, men have argued that only in a wealthy society in which relatively few citizens lived in real poverty could a situation exist in which the mass of the population could intelligently participate in politics and could develop the self-restraint necessary to avoid succumbing to the appeals of irresponsible demagogues. A society divided between a large impoverished mass and a small favored elite would result either in oligarchy [...] or in tyranny" (Lipset, 1959, p.75).

[32]See Tables F.30 and F.31.

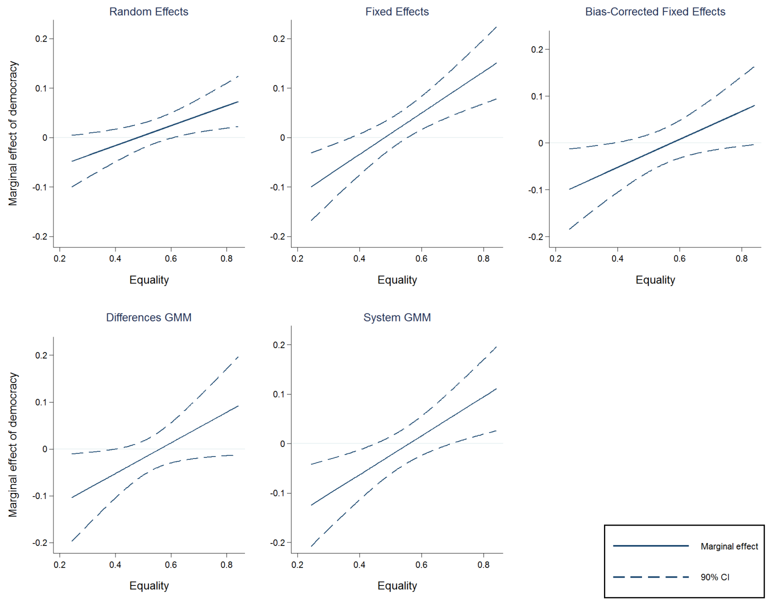
[33]Detailed results are reported in Tables F.32 and F.33.

[34]For detailed results see Tables F.34 and F.35.

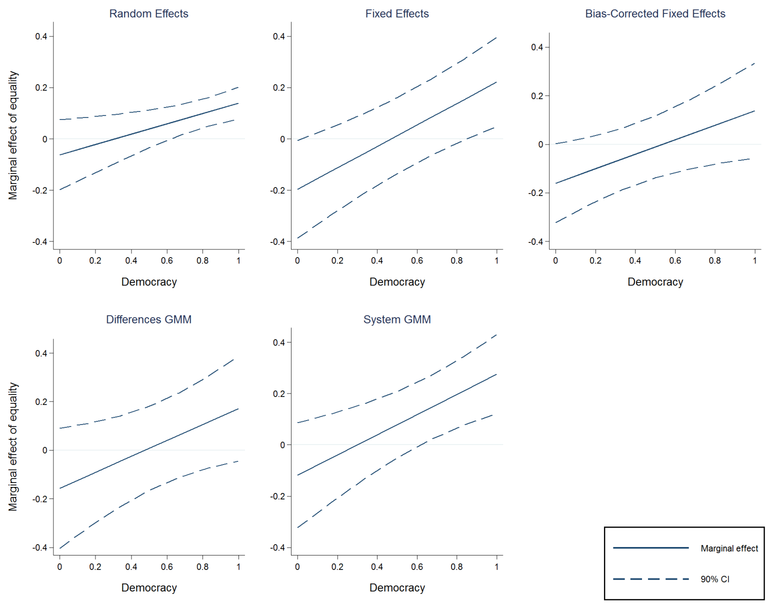
F.3 Additional Material: Tables and Figures

This section contains the following figures:

This section contains the following tables:

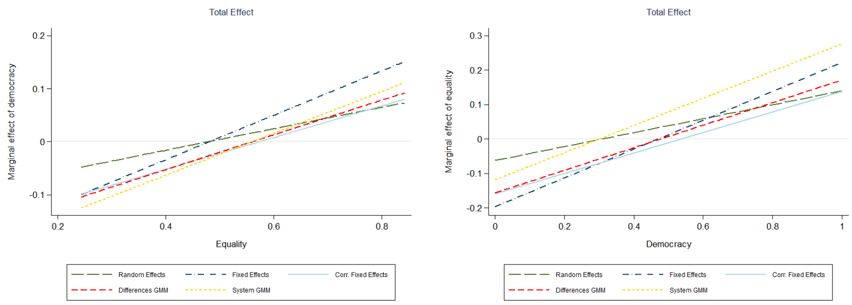


(a) Democracy and Institutional Quality



(b) Equality and Institutional Quality

Figure F.1: Marginal Effects Table 1(b)



(a) Democracy and Institutional Quality (b) Equality and Institutional Quality

Figure F.2: Marginal Effects Across Estimators Table 1(b)

Table F.1: Appendix: Data Sources

Economic Freedom in the World (EF): 1970–2010. Raw data range from 0 to 10, where higher scores represent better economic institutions. Values are divided by the factor 10 to normalize the range from 0 to 1. The index is composed of 42 distinct variables in five general categories: size of government and taxation; private property and the rule of law; soundness of money; trade regulation and tariffs; regulation of business, labor, and capital markets. In the analysis we use the composite indicator and the subcomponents for regulation of business, labor and capital markets, and private property and rule of law. Source: Gwartney, Lawson, and Hall (2013), web link: http://www.freetheworld.com/.

Civil Liberties (CL): 1972–2010. Raw data are coded as integer values from 1 (most free) to 7 (least free). The ratings are based on the evaluation of 15 questions with respect to four subcategories: freedom of expression and belief; associational and organizational rights; rule of law; and personal autonomy and individual rights. In particular, the last subcomponent encompasses aspects regarding choices of employment and higher education, rights to own property and establish private business, and government influence by government officials. The scale is reversed and normalized to range from 0 to 1 so that higher values indicate better economic institutions. The country-year observations of 1972 are used for the year 1970. Source: http://www.freedomhouse.org/report-types/freedom-world.

Political Rights (PR): 1972–2010. Raw data are coded as integer values from 1 (wide range of political rights) to 7 (virtually no political rights). The ratings are based on the evaluation of 10 questions with respect to three subcategories: electoral process, political pluralism and participation, and functioning of government. The scale is reversed and normalized to range from 0 to 1 so that higher values indicate more democratic institutions. The country-year observations of 1972 are used for the year 1970. Source: http://www.freedomhouse.org/report-types/freedom-world.

International Country Risk Guide (ICRG) Political Risk Rating: 1984–2010. The rating takes integer values from 0 to 100 and is aimed to provide a means of assessing the political stability of the countries covered by ICRG on a comparable basis. In the analysis, we use the subcomponents investment profile, which measures the degree of property rights, profits repatriation and payment delays (takes values 0–12), protection against corruption (0–6), and the prevalence of law and order (0–6). The subcomponents are normalized to range from 0 to 1, where higher values indicate better economic institutions. Source: http://www.prsgroup.com/about-us/our-two-methodologies/prs.

PolityIV Combined Democracy-Autocracy indicator: 1800–2010. The raw combined index is the sum of the Polity autocracy and the democracy indicators, which take integer values from –10 to 0 and 0 to 10, respectively, and thus ranges from –10 to 10. The indicator is normalized to range from 0 to 1 where higher values represent more democratic institutions. Source: Marshall, Jaggers, and Gurr (2013), web link: http://www.systemicpeace.org/inscrdata.html.

PolityIV Constraints on Executive indicator: 1800–2010. Raw data take integer values from 1 to 7, which refer to institutionalized constraints on decision-making powers of chief executives. The scale is reversed and normalized to range from 0 to 1 so that higher values indicate stronger constraints on the executive and, thus, more democratic institutions. Source: Marshall, Jaggers, and Gurr (2013), web link: http://www.systemicpeace.org/inscrdata.html.

Vanhanen Democracy Index, v.2.0: 1810–2010. Raw data are based on two subcomponents, competition and participation, each ranging from 0 to 100. The competition variable is defined as 100 minus the number of votes the largest party has won. Participation is defined as the the share of total population that has participated in elections. As competition and participation are considered complementary, the total index is calculated of the product of both measures. Indices are normalized to range from 0 to 1, where higher values indicate more democratic institutions. Source: Vanhanen and Lundell (2014), web link http://www.prio.org/Data/Governance/Vanhanens-index-of-democracy/.

Democracy-Dictatorship Indicator: 1946–2008. Raw data are a dichotomous classification of political regimes according to whether the leading government positions are filled by elections and to what extent these elections are contested. Democratic regimes are assigned value 1 and autocratic regimes, that is, governments that do not fulfill requirements for being listed as a democracy, assigned value 0. Source: Cheibub, Gandhi, and Vreeland (2010), web link: http://sites.google.com/site/joseantoniocheibub/datasets/democracy-and-dictatorship-revisited.

SWIID (In-)Equality Measures, v.5.0: 1960–2010. Raw data are net and gross income Gini coefficients ranging from 0 to 100. For low values of the Gini coefficient, income is spread relatively even across individuals in the population, while for higher values income is concentrated in the hands of very few people. The scale is reverted and normalized to range from 0 to 1 so that higher score represent a higher degree of income equality. The measure for absolute redistribution is defined as the difference between gross and net Gini coefficients. The measure for relative redistribution corresponds to the absolute redistribution divided by the Gini coefficient. Source: Solt (2009, 2016b), web link: http://myweb.uiowa.edu/fsolt/swiid/swiid.html.

Log GDP per Capita: 1950–2010. Raw data are output-side real GDP at chained Purchasing Power Parities (in mil. 2005 US$) and population in millions constructed by Penn World Tables v.8.1. Source: Feenstra, Inklaar, and Timmer (2015), web link: http://www.rug.nl/research/ggdc/data/pwt/pwt-8.1.

Years of Schooling and Log Population: 1950–2010. Raw data are average years of schooling and population constructed by the Educational Attainment data set v.1.3. Source: Barro and Lee (2013), web link: http://www.barrolee.com/.

Ethnic Polarization: Constant across time. Raw data range from 0 to 1 with higher values indicating ethnically more polarized societies. Source: Reynal-Querol and Montalvo (2005).

Central Government Revenues, Expenditures, and Social Security Funds: 1970–2010 for central government statistics and 1990–2010 for social security funds. Raw data are the share of revenues and expenditures of the central government, and expenditures of social security funds as percentage of GDP by the Government Finance Statistics Manuel (2014). Source: International Monetary Fund, web link: http://data.imf.org/.

Social Expenditures: 1980–2010. Raw data are social expenditures as percentage of GDP constructed by the OECD Social Expenditure Database (SOCX, 2014). Source: Organization for Economic Co-operation and Development (OECD), web link: http://www.oecd.org/social/expenditure.htm.

Colonial History: Constant across time. Raw data constructed by CEPII Geo data set contain information on the countries that have colonized a specific other country. A dummy variable is constructed that takes a value of 1, if a country has been colonized by other countries, and 0 otherwise. Source: Mayer and Zignago (2011), web link: http://www.cepii.fr/cepii/en/bdd_modele/bdd.asp.

Inflation and Deflation: 1960–2010. Raw data on inflation rates are from the World Development Indicators by the World Bank. Dummies are constructed to take a value of 1, if inflation rates are higher than 4 percent or lower than 0 percent. Source: World Bank (2014) , web link: http://data.worldbank.org/.

Socialist: Constant across time. A dummy is constructed that takes a value of 1, if the country was a member of the former Eastern Bloc and 0 otherwise.

Oil Exporter: Constant across time. Raw data on annual crude oil exports measured in thousands of barrels per day from 1980 to 2010 are taken from the International Energy Statistics by the U.S. Energy Information Administration, web link: http://www.eia.gov/. In order to compute the value of exports, crude oil exports are multiplied by the crude oil first purchase price in the U.S., provided by the U.S. Energy Information Administration, and adjusted by the annual implicit price deflator provided by the U.S. Bureau of Economic Analysis, web link: http://www.bea.gov/. A dummy is constructed that takes a value of 1, if a country exports oil on average worth of 5 percent or more of its GDP over the time from 1980 to 2010.

Relative Political Extraction: 1960–2010. The measure approximates the ability of governments to appropriate portions of the national output to advance public goods. Values are normalized to range from 0 to 1. Source: Arbetman-Rabinowitz, Fisunoglu, Kugler, Abdollahian, Johnson, Kang, and Yang (2013), web link: http://dataverse.harvard.edu/dataverse/rpc.

Table F.2: Descriptive Statistics

	Mean	Std. Dev.	Min.	Max.	Obs.
Economic Institutions					
Economic Freedom	0.65	0.12	0.31	0.89	543
Civil Liberties	0.70	0.27	0	1	543
Composite Index (CL & EF)	0.62	0.21	0.06	1	543
Regulation of Credit, Labor, and Business (EF)	0.64	0.12	0.16	0.90	538
Soundness of Money (EF)	0.75	0.21	0	0.99	543
Investment Profile (ICRG)	0.65	0.20	0.15	1	460
Protection Against Corruption (ICRG)	0.56	0.23	0	1	460
Political Institutions					
Constraints on Executive (XC)	0.73	0.34	0	1	543
PolityIV (PIV)	0.74	0.33	0	1	543
Political Rights (PR)	0.71	0.32	0	1	543
Vanhanen Index (VH)	0.19	0.14	0	0.47	543
Vanhanen Competition (VHC)	0.43	0.22	0	0.70	543
Composite Index (XC & VHC)	0.67	0.31	0	1	543
Composite Index (PIV & PR & VH)	0.60	0.29	0.02	1	543
(In-)Equality Measures					
Reversed Gini (1 – Gini net)	0.62	0.11	0.24	0.84	543
Reversed Gini (1 – Gini market)	0.56	0.09	0.21	0.80	543
(1 – Human Capital Gini)	0.70	0.19	0.10	0.97	543
Other Variables					
Log GDP	8.81	1.13	5.97	11.04	543
Log Population	9.25	1.51	5.71	13.85	543
Years of Schooling	7.24	2.80	0.90	12.91	543
Inflation	0.71	0.46	0	1	543
Deflation	0.04	0.20	0	1	543
Socialist	0.09	0.28	0	1	543
Oil Exporter	0.12	0.33	0	1	543

Table F.3: Simple Correlations

	Economic Freedom	Civil Liberties	PC EF & CL	Executive Constraints	Polity IV Index	PC XC & VHC	PC PR & PIV & VH	Gini Net	Gini Gross
Economic Freedom	1.00								
Civil Liberties	0.58	1.00							
PC EF & CL	0.89	0.89	1.00						
Executive Constraints	0.50	0.80	0.73	1.00					
PolityIV Index	0.51	0.82	0.75	0.96	1.00				
PC XC & VHC	0.52	0.80	0.75	0.96	0.95	1.00			
PC PR & PIV & VH	0.55	0.89	0.82	0.91	0.95	0.94	1.00		
Net Gini	0.34	0.52	0.48	0.39	0.38	0.41	0.51	1.00	
Gross Gini	0.18	0.29	0.26	0.23	0.22	0.23	0.30	0.84	1.00
Human Capital Gini	0.51	0.60	0.63	0.59	0.60	0.61	0.63	0.42	0.25

Notes: PC refers to the principal component extracted from the respective variables in parentheses by ways of a principal component analysis (see main text for details). Democracy in the respective columns is proxied by Constraints on Executive; an artificial indicator based on the principal component of Constraints and Executive and the Vanhanen Competition indicator; the combined PolityIV indicator; and an artificial indicator based on the principal components of the combined PolityIV indicator, Political Rights, and the composite Vanhanen indicator.

Table F.4: Robustness: Economic Institutions (DGMM)

	Civil Liberties	Regulation	Soundness of Money	Investment Profile	Protection vs. Corruption
	(1)	(2)	(3)	(4)	(5)
L.Inst. Quality	0.59***	0.59***	0.60***	0.06	0.34***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.09)
L.Equality	-0.18	-0.11	-0.39*	-0.54**	0.12
	(0.16)	(0.09)	(0.21)	(0.26)	(0.25)
L.Democracy	-0.30**	-0.15**	-0.23	-0.27	-0.12
	(0.13)	(0.07)	(0.18)	(0.23)	(0.17)
L.(Eq×Demo)	0.47**	0.25**	0.50*	0.54	0.26
	(0.21)	(0.12)	(0.30)	(0.42)	(0.33)
Controls	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓
AR(2) p–value	0.82	0.10	0.25	0.04	0.30
Hansen p–value	0.24	0.48	0.19	0.01	0.13
Instruments	82	82	82	49	49
Groups	112	96	96	90	90
Observations	498	443	461	317	317

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. Results are computed using the differences GMM estimator by Arellano and Bond (1991). Dependent variables in the respective columns are the principal component of Economic Freedom and Civil Liberties; Civil Liberties; Regulation of Credit, Labor, and Business; Soundness of Money; Investment Profile (Property Rights); and Protection against Corruption. Democracy is measured by the Constraints on Executive indicator. Equality is proxied by $(1 - \text{Net Gini})$. Control variables are log GDP p.c. and average years of schooling. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.5: Robustness: Different Measures for Economic Institutions

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(a) Dependent Variable: Civil Liberties 1970–2010 (CL)										
L.Equality	0.11*	-0.18	-0.06	-0.34**	-0.05	-0.27*	0.08	-0.18	0.17	-0.04
	(0.05)	(0.11)	(0.11)	(0.15)	(0.11)	(0.16)	(0.14)	(0.16)	(0.14)	(0.19)
L.Democracy	0.02	-0.20***	0.01	-0.31***	-0.05	-0.30**	-0.06	-0.30**	-0.06	-0.30**
	(0.02)	(0.07)	(0.03)	(0.11)	(0.04)	(0.14)	(0.04)	(0.13)	(0.05)	(0.13)
L.(Eq×Demo)		0.43***		0.60***		0.48**		0.47**		0.45*
		(0.13)		(0.20)		(0.24)		(0.21)		(0.23)
AR(2) p–value							0.83	0.82	0.82	0.78
Hansen p–value							0.22	0.24	0.05	0.16
Diff.–in–Hansen p–value									0.35	0.34
Instruments							62	82	68	89
Groups	112	112	112	112	112	112	112	112	112	112
Observations	610	610	610	610	610	610	498	498	610	610
(b) Dependent Variable: Regulation of Credit, Labor, and Business 1970–2010 (EF)										
L.Equality	0.03	-0.07	-0.01	-0.23*	0.01	-0.14	0.01	-0.11	0.03	-0.16
	(0.03)	(0.08)	(0.08)	(0.12)	(0.06)	(0.10)	(0.11)	(0.09)	(0.06)	(0.11)
L.Democracy	-0.01	-0.08	-0.01	-0.24***	-0.01	-0.17**	-0.01	-0.15**	-0.01	-0.18**
	(0.01)	(0.05)	(0.02)	(0.08)	(0.02)	(0.08)	(0.02)	(0.07)	(0.02)	(0.08)
L.(Eq×Demo)		0.13		0.42**		0.29**		0.25**		0.32**
		(0.10)		(0.16)		(0.14)		(0.12)		(0.13)
AR(2) p–value							0.11	0.10	0.10	0.13
Hansen p–value							0.14	0.48	0.48	0.59
Diff.–in–Hansen p–value									0.87	0.64
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	539	539	539	539	539	539	443	443	539	539
(c) Dependent Variable: Soundness Money 1970–2010 (EF)										
L.Equality	0.16***	-0.05	0.05	-0.31*	0.05	-0.33*	-0.13	-0.39*	0.05	-0.31*
	(0.05)	(0.18)	(0.16)	(0.19)	(0.14)	(0.19)	(0.19)	(0.21)	(0.13)	(0.18)
L.Democracy	0.03	-0.13	0.06	-0.35**	0.05	-0.38**	0.04	-0.23	0.07*	-0.34**
	(0.02)	(0.14)	(0.04)	(0.18)	(0.05)	(0.15)	(0.05)	(0.18)	(0.05)	(0.15)
L.(Eq×Demo)		0.29		0.74**		0.78***		0.50*		0.72***
		(0.23)		(0.29)		(0.26)		(0.30)		(0.27)
AR(2) p–value							0.24	0.25	0.16	0.18
Hansen p–value							0.06	0.19	0.06	0.22
Diff.–in–Hansen p–value									0.61	0.79
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	557	557	557	557	557	557	461	461	557	557
(d) Dependent Variable: Property Rights, Profit Repatriation, Payment Delays 1985–2010 (ICRG)										
L.Equality	0.02	-0.43**	-0.13	-0.47**	-0.19	-0.56**	-0.15	-0.54**	0.00	-0.51
	(0.08)	(0.21)	(0.24)	(0.23)	(0.18)	(0.26)	(0.27)	(0.26)	(0.24)	(0.31)
L.Democracy	0.03	-0.32**	0.07	-0.31*	0.05	-0.36*	0.03	-0.27	0.07	-0.48**
	(0.03)	(0.14)	(0.05)	(0.16)	(0.04)	(0.21)	(0.07)	(0.23)	(0.07)	(0.23)
L.(Eq×Demo)		0.60**		0.65**		0.72**		0.54		0.93**
		(0.24)		(0.27)		(0.35)		(0.42)		(0.41)
AR(2) p–value							0.03	0.04	0.51	0.33
Hansen p–value							0.00	0.01	0.00	0.00
Diff.–in–Hansen p–value									0.00	0.24
(e) Dependent Variable: Protection against Corruption 1985–2010 (ICRG)										
L.Equality	0.31***	-0.11	0.27*	0.22	0.29*	0.21	0.31	0.12	0.36**	0.19
	(0.06)	(0.14)	(0.15)	(0.24)	(0.16)	(0.24)	(0.20)	(0.25)	(0.15)	(0.27)
L.Democracy	0.02	-0.32***	0.04	-0.02	0.01	-0.08	0.04	-0.12	-0.04	-0.21
	(0.02)	(0.10)	(0.04)	(0.18)	(0.03)	(0.19)	(0.05)	(0.17)	(0.05)	(0.20)
L.(Eq×Demo)		0.59***		0.10		0.17		0.26		0.34
		(0.19)		(0.33)		(0.32)		(0.33)		(0.37)
AR(2) p–value							0.32	0.30	0.11	0.11
Hansen p–value							0.06	0.13	0.13	0.18
Diff.–in–Hansen p–value									0.16	0.15
Instruments							37	49	43	56
Groups	90	90	90	90	90	90	90	90	90	90
Observations	407	407	407	407	407	407	317	317	407	407
Lagged Y	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects, as well as the lagged dependent variable as control. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by the (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.6: Robustness: Binary Democracy Indicator

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Economic Freedom									
L.Inst. Quality	0.71***	0.71***	0.49***	0.49***	0.84***	0.83***	0.69***	0.66***	0.75***	0.72***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.04)
L.Equality	0.02	-0.03	0.02	-0.08	-0.01	-0.09	0.00	-0.11	0.01	-0.15
	(0.02)	(0.05)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)	(0.10)	(0.06)	(0.10)
L.Democracy	0.01	-0.04	0.03**	-0.11**	0.03**	-0.10**	0.03*	-0.09	0.02	-0.12*
	(0.01)	(0.04)	(0.01)	(0.04)	(0.01)	(0.04)	(0.02)	(0.06)	(0.01)	(0.07)
L.(Eq×Demo)		0.08		0.24***		0.22***		0.21**		0.25**
		(0.07)		(0.08)		(0.07)		(0.11)		(0.11)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.96	0.85	0.92	0.80
Hansen p–value							0.15	0.25	0.06	0.17
Diff.–in–Hansen p–value									0.60	0.85
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	546	546	546	546	546	546	450	450	546	546
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
L.Inst. Quality	0.77***	0.76***	0.55***	0.53***	0.83***	0.81***	0.73***	0.69***	0.77***	0.72***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.04)
L.Equality	0.06*	-0.02	0.01	-0.08	-0.01	-0.06	0.04	-0.07	0.11	-0.09
	(0.04)	(0.06)	(0.09)	(0.10)	(0.08)	(0.09)	(0.10)	(0.11)	(0.08)	(0.12)
L.Democracy	0.02*	-0.07	0.03**	-0.09	0.02	-0.06	0.00	-0.10	-0.00	-0.18**
	(0.01)	(0.04)	(0.02)	(0.06)	(0.01)	(0.05)	(0.02)	(0.07)	(0.01)	(0.07)
L.(Eq×Demo)		0.15**		0.22**		0.14		0.19		0.33***
		(0.07)		(0.10)		(0.09)		(0.12)		(0.12)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.35	0.29	0.26	0.20
Hansen p–value							0.04	0.18	0.11	0.34
Diff.–in–Hansen p–value									0.91	1.00
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	546	546	546	546	546	546	450	450	546	546

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Equality is proxied by (1 − Net Gini). Democracy is proxied by the Democracy-Dictatorship measure of Cheibub, Gandhi, and Vreeland (2010). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.7: Robustness: Political Institutions (DGMM)

	Principal Comp. XC & VHC	PolityIV Index	Political Rights	Principal Comp. PIV & PR & VH	Democracy-Dictatorship
	(1)	(2)	(3)	(4)	(5)
L.Inst. Quality	0.62***	0.62***	0.60***	0.62***	0.66***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
L.Equality	-0.17	-0.16	-0.21*	-0.15	-0.11
	(0.13)	(0.12)	(0.13)	(0.12)	(0.10)
L.Democracy	-0.16	-0.09	-0.18*	-0.14	-0.09
	(0.11)	(0.09)	(0.09)	(0.12)	(0.06)
L.(Eq×Demo)	0.40**	0.27*	0.39**	0.37*	0.21**
	(0.19)	(0.16)	(0.16)	(0.19)	(0.11)
Controls	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓
AR(2) p–value	0.59	0.70	0.94	0.81	0.85
Hansen p–value	0.34	0.30	0.33	0.24	0.25
Instruments	82	82	82	82	82
Groups	96	96	96	96	96
Observations	447	454	454	454	450

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. Results are computed using the differences GMM estimator by Arellano and Bond (1991). The dependent variable is Economic Freedom. The democracy proxies in the respective columns are the principal component of Constraints on Executive and the Vanhanen Competition indicator; the combined PolityIV indicator; Political Rights; the principal component of the combined PolityIV indicator, Political Rights, and the composite Vanhanen Democracy indicator; and the dichotomous Democracy-Dictatorship indicator. Equality is proxied by $(1 - \text{Net Gini})$. Control variables are log GDP p.c. and average years of schooling. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.8: Robustness: Different Measures for Political Institutions

	Dependent Variable: Economic Freedom									
	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(a) Democracy Variable: Composite Index (PCA of Constraints on Executive & Vanhanen Competition Component)										
L.Equality	0.03	-0.05	0.03	-0.13	-0.01	-0.15**	0.05	-0.17	0.02	-0.22**
	(0.02)	(0.07)	(0.06)	(0.09)	(0.06)	(0.07)	(0.07)	(0.13)	(0.06)	(0.11)
L.Democracy	0.02*	-0.05	0.06***	-0.15**	0.05***	-0.14**	0.06**	-0.16	0.05**	-0.20**
	(0.01)	(0.06)	(0.02)	(0.07)	(0.02)	(0.06)	(0.02)	(0.11)	(0.02)	(0.09)
L.(Eq×Demo)		0.13		0.38***		0.35***		0.40**		0.46***
		(0.10)		(0.13)		(0.11)		(0.19)		(0.15)
AR(2) p–value							0.80	0.59	0.80	0.60
Hansen p–value							0.14	0.34	0.06	0.34
Diff.–in–Hansen p–value									0.56	0.88
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543
(b) Democracy Variable: PolityIV Composite Index										
L.Equality	0.02	-0.03	0.03	-0.10	-0.00	-0.10	-0.02	-0.16	0.00	-0.14
	(0.02)	(0.07)	(0.06)	(0.09)	(0.06)	(0.08)	(0.07)	(0.12)	(0.06)	(0.11)
L.Democracy	0.03**	-0.02	0.07***	-0.07	0.06***	-0.05	0.06***	-0.09	0.05***	-0.10
	(0.01)	(0.05)	(0.02)	(0.07)	(0.02)	(0.06)	(0.02)	(0.09)	(0.02)	(0.07)
L.(Eq×Demo)		0.07		0.25**		0.20*		0.27*		0.25*
		(0.09)		(0.12)		(0.11)		(0.16)		(0.14)
AR(2) p–value							0.80	0.70	0.80	0.72
Hansen p–value							0.12	0.30	0.10	0.37
Diff.–in–Hansen p–value									0.84	0.97
(c) Democracy Variable: Political Rights										
L.Equality	0.02	-0.09	0.02	-0.18*	-0.01	-0.20**	-0.01	-0.21*	0.03	-0.19
	(0.02)	(0.09)	(0.06)	(0.10)	(0.06)	(0.08)	(0.07)	(0.13)	(0.06)	(0.12)
L.Democracy	0.03*	-0.08	0.06***	-0.18**	0.06***	-0.18***	0.05*	-0.18*	0.03	-0.20**
	(0.02)	(0.07)	(0.02)	(0.08)	(0.02)	(0.07)	(0.03)	(0.09)	(0.02)	(0.10)
L.(Eq×Demo)		0.17		0.43***		0.40***		0.39**		0.41**
		(0.12)		(0.14)		(0.11)		(0.16)		(0.18)
AR(2) p–value							0.92	0.94	0.91	0.98
Hansen p–value							0.16	0.33	0.06	0.19
Diff.–in–Hansen p–value									0.76	0.72
(d) Democracy Variable: Composite Index (PCA of Political Rights & PolityIV & Vanhanen Index)										
L.Equality	0.02	-0.04	0.03	-0.11	-0.00	-0.13	-0.02	-0.15	0.01	-0.17
	(0.02)	(0.07)	(0.06)	(0.09)	(0.06)	(0.08)	(0.08)	(0.12)	(0.06)	(0.11)
L.Democracy	0.03*	-0.03	0.09***	-0.13	0.08***	-0.12	0.09***	-0.14	0.06**	-0.19*
	(0.02)	(0.06)	(0.02)	(0.08)	(0.02)	(0.08)	(0.03)	(0.12)	(0.03)	(0.11)
L.(Eq×Demo)		0.10		0.38***		0.34**		0.37*		0.42**
		(0.11)		(0.14)		(0.14)		(0.19)		(0.17)
AR(2) p–value							0.89	0.81	0.88	0.82
Hansen p–value							0.08	0.24	0.06	0.27
Diff.–in–Hansen p–value									0.63	0.74
(e) Democracy Variable: Vanhanen Composite Index										
L.Equality	0.02	-0.01	0.02	-0.03	-0.00	-0.05	0.00	-0.04	0.02	-0.08
	(0.02)	(0.04)	(0.06)	(0.07)	(0.06)	(0.07)	(0.06)	(0.07)	(0.06)	(0.09)
L.Democracy	0.05	-0.05	0.17***	-0.22	0.16***	-0.24	0.12*	-0.27	0.10*	-0.47
	(0.03)	(0.12)	(0.05)	(0.19)	(0.05)	(0.22)	(0.07)	(0.23)	(0.06)	(0.30)
L.(Eq×Demo)		0.15		0.64**		0.64*		0.65*		0.97**
		(0.20)		(0.29)		(0.35)		(0.36)		(0.45)
AR(2) p–value							0.94	0.81	0.92	0.71
Hansen p–value							0.20	0.37	0.08	0.28
Diff.–in–Hansen p–value									0.52	0.64
(f) Democracy Variable: Vanhanen Competition Component										
L.Equality	0.03	-0.04	0.02	-0.09	-0.01	-0.11	0.03	-0.11	0.04	-0.16*
	(0.02)	(0.06)	(0.06)	(0.08)	(0.06)	(0.07)	(0.07)	(0.10)	(0.07)	(0.09)
L.Democracy	0.03**	-0.07	0.07***	-0.20**	0.06***	-0.20**	0.07**	-0.21*	0.06**	-0.28**
	(0.02)	(0.07)	(0.02)	(0.09)	(0.02)	(0.10)	(0.03)	(0.12)	(0.03)	(0.13)
L.(Eq×Demo)		0.18		0.48***		0.47***		0.52**		0.60***
		(0.13)		(0.16)		(0.17)		(0.21)		(0.22)
AR(2) p–value							0.95	0.72	0.94	0.68
Hansen p–value							0.13	0.38	0.11	0.39
Diff.–in–Hansen p–value									0.68	0.94
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	550	550	550	550	550	550	454	454	550	550
Lagged Y	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Table F.8 and the principal component of Economic Freedom and Civil Liberties in Table F.9. Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM (continued on next page ...)

Table F.9: Robustness: Different Measures for Political Institutions II

	Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Democracy Variable: Combined Index: Constraints on Executive & Vanhanen Competition Component (PCA)									
L.Equality	0.08**	-0.05	0.00	-0.17	-0.03	-0.14	0.04	-0.18	0.11*	-0.15
	(0.04)	(0.08)	(0.09)	(0.11)	(0.08)	(0.09)	(0.11)	(0.14)	(0.06)	(0.13)
L.Democracy	0.02	-0.09	0.05**	-0.18**	-0.00	-0.16*	-0.00	-0.21**	-0.01	-0.26***
	(0.02)	(0.06)	(0.02)	(0.07)	(0.02)	(0.09)	(0.03)	(0.09)	(0.03)	(0.09)
L.(Eq×Demo)		0.21**		0.41***		0.31**		0.38**		0.47***
		(0.10)		(0.13)		(0.16)		(0.15)		(0.17)
AR(2) p–value							0.17	0.11	0.13	0.08
Hansen p–value							0.03	0.15	0.10	0.32
Diff.–in–Hansen p–value									0.87	0.82
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543
	(b) Democracy Variable: PolityIV Composite Index									
L.Equality	0.08**	-0.02	0.01	-0.12	-0.01	-0.08	0.03	-0.14	0.10	-0.12
	(0.04)	(0.09)	(0.09)	(0.12)	(0.08)	(0.11)	(0.10)	(0.16)	(0.07)	(0.14)
L.Democracy	0.03**	-0.04	0.06***	-0.10	0.01	-0.08	0.00	-0.15	0.01	-0.22**
	(0.02)	(0.06)	(0.02)	(0.07)	(0.02)	(0.09)	(0.03)	(0.11)	(0.03)	(0.08)
L.(Eq×Demo)		0.14		0.28**		0.17		0.30		0.40***
		(0.10)		(0.12)		(0.15)		(0.20)		(0.15)
AR(2) p–value							0.33	0.27	0.25	0.20
Hansen p–value							0.05	0.13	0.04	0.22
Diff.–in–Hansen p–value									0.76	0.87
	(c) Democracy Variable: Political Rights									
L.Equality	0.08**	-0.08	0.01	-0.21*	-0.01	-0.17	-0.00	-0.22	0.11	-0.12
	(0.03)	(0.09)	(0.09)	(0.12)	(0.08)	(0.12)	(0.10)	(0.14)	(0.07)	(0.16)
L.Democracy	0.04*	-0.10	0.06**	-0.21***	-0.00	-0.20**	-0.00	-0.27***	-0.02	-0.29**
	(0.02)	(0.06)	(0.02)	(0.07)	(0.03)	(0.09)	(0.03)	(0.10)	(0.03)	(0.12)
L.(Eq×Demo)		0.23**		0.47***		0.35**		0.48***		0.51**
		(0.11)		(0.12)		(0.16)		(0.17)		(0.20)
AR(2) p–value							0.34	0.34	0.25	0.26
Hansen p–value							0.02	0.16	0.06	0.17
Diff.–in–Hansen p–value									0.91	0.86
	(d) Democracy Variable: Combined Index: Political Rights & PolityIV & Vanhanen Index (PCA)									
L.Equality	0.07*	-0.03	0.01	-0.14	-0.01	-0.11	-0.01	-0.17	0.11	-0.13
	(0.04)	(0.08)	(0.09)	(0.11)	(0.08)	(0.11)	(0.11)	(0.16)	(0.07)	(0.13)
L.Democracy	0.04**	-0.06	0.08***	-0.16*	0.01	-0.16	-0.02	-0.25*	-0.02	-0.33***
	(0.02)	(0.07)	(0.03)	(0.08)	(0.03)	(0.11)	(0.04)	(0.14)	(0.04)	(0.10)
L.(Eq×Demo)		0.17		0.41***		0.31*		0.44*		0.56***
		(0.12)		(0.14)		(0.19)		(0.23)		(0.17)
AR(2) p–value							0.33	0.27	0.25	0.19
Hansen p–value							0.02	0.15	0.04	0.20
Diff.–in–Hansen p–value									0.83	0.80
	(e) Democracy Variable: Vanhanen Composite Index									
L.Equality	0.06*	0.02	0.01	-0.05	-0.01	-0.06	0.02	-0.06	0.08	-0.03
	(0.04)	(0.06)	(0.09)	(0.10)	(0.08)	(0.09)	(0.10)	(0.10)	(0.08)	(0.08)
L.Democracy	0.05	-0.15	0.14**	-0.28	0.02	-0.37	-0.12	-0.64*	-0.07	-0.69***
	(0.04)	(0.16)	(0.06)	(0.23)	(0.07)	(0.30)	(0.09)	(0.33)	(0.06)	(0.23)
L.(Eq×Demo)		0.31		0.68**		0.66		0.93*		1.07***
		(0.25)		(0.34)		(0.47)		(0.49)		(0.37)
AR(2) p–value							0.29	0.21	0.25	0.16
Hansen p–value							0.04	0.20	0.20	0.42
Diff.–in–Hansen p–value									0.87	0.84
	(f) Democracy Variable: Vanhanen Competition Component									
L.Equality	0.08**	-0.02	0.01	-0.10	-0.01	-0.08	0.03	-0.10	0.05	-0.12
	(0.04)	(0.07)	(0.09)	(0.10)	(0.08)	(0.09)	(0.08)	(0.12)	(0.07)	(0.10)
L.Democracy	0.03	-0.12	0.06*	-0.19*	0.01	-0.18	-0.02	-0.28**	-0.01	-0.31**
	(0.02)	(0.08)	(0.03)	(0.10)	(0.03)	(0.14)	(0.04)	(0.13)	(0.03)	(0.12)
L.(Eq×Demo)		0.26*		0.43**		0.34		0.50**		0.57***
		(0.14)		(0.17)		(0.24)		(0.21)		(0.21)
AR(2) p–value							0.28	0.20	0.26	0.17
Hansen p–value							0.06	0.19	0.17	0.46
Diff.–in–Hansen p–value									0.86	0.84
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	550	550	550	550	550	550	454	454	550	550
Lagged Y	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

(... continuation from previous page) Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.10: Robustness: Static Model – Effect of Democracy and Equality on Institutional Quality

	Random Effects		Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(a) Dependent Variable: Economic Freedom							
L.Equality	0.08	-0.08	0.06	-0.10	0.01	-0.18*	0.01	-0.18
	(0.05)	(0.11)	(0.08)	(0.11)	(0.07)	(0.10)	(0.07)	(0.14)
L.Democracy	0.06***	-0.09	0.07***	-0.12	0.06***	-0.12*	0.06***	-0.19**
	(0.02)	(0.08)	(0.02)	(0.09)	(0.02)	(0.07)	(0.02)	(0.08)
L.(Eq×Demo)		0.27**		0.32**		0.32***		0.44***
		(0.13)		(0.15)		(0.12)		(0.15)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓					✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value					0.29	0.23	0.30	0.23
Hansen p–value					0.27	0.30	0.46	0.24
Diff.–in–Hansen p–value							0.76	0.57
Instruments					49	69	54	75
Groups	96	96	96	96	96	96	96	96
Observations	560	560	560	560	464	464	560	560
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)							
L.Equality	0.15*	-0.19	0.06	-0.22*	-0.04	-0.37**	0.11	-0.22
	(0.08)	(0.13)	(0.10)	(0.13)	(0.12)	(0.17)	(0.09)	(0.15)
L.Democracy	0.21***	-0.12	0.18***	-0.14	0.12***	-0.25**	0.13***	-0.29***
	(0.03)	(0.09)	(0.03)	(0.09)	(0.04)	(0.11)	(0.03)	(0.10)
L.(Eq×Demo)		0.58***		0.56***		0.69***		0.77***
		(0.14)		(0.15)		(0.19)		(0.19)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓					✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value					0.46	0.20	0.37	0.21
Hansen p–value					0.20	0.07	0.33	0.23
Diff.–in–Hansen p–value							0.57	0.38
Instruments					49	69	54	75
Groups	96	96	96	96	96	96	96	96
Observations	560	560	560	560	464	464	560	560

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the combined PolityIV indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c., average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first two lags of the explanatory variables. In the level equation, the lagged difference of the regressors is used so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.11: Robustness: Economic (In-)Equality (DGMM)

	Multiple Imputation (1)	Gross Gini (2)	Binary Indicator (3)	Human Capital Gini (4)	Lowest Quintile (5)	Second–Fourth Quintile (6)	Highest Quintile (7)
L.Inst. Quality	0.65***	0.67***	0.70***	0.70***	0.68***	0.69***	0.68***
	(0.05)	(0.05)	(0.07)	(0.05)	(0.11)	(0.08)	(0.10)
L.Equality	-0.14	-0.11	-0.03	0.03			
	(0.11)	(0.13)	(0.02)	(0.07)			
L.Democracy	-0.11	-0.10	0.03	-0.00	0.05	0.07**	0.14***
	(0.08)	(0.08)	(0.02)	(0.04)	(0.03)	(0.03)	(0.04)
L.(Eq×Demo)	0.29**	0.28*	0.07***	0.07			
	(0.14)	(0.15)	(0.02)	(0.06)			
Controls	✓	✓	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value	0.71	0.74	0.68	0.56	0.82	0.99	0.84
Hansen p–value	0.39	0.39	0.35	0.14	0.50	0.13	0.32
Instruments	82	82	81	82	30	42	22
Groups	96	96	96	96	35	77	29
Observations	447	447	447	543	109	280	99

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. Results are computed using the differences GMM estimator by Arellano and Bond (1991). The dependent variable is Economic Freedom. Democracy is measured by the Constraints on Executive indicator. Equality measures in the respective columns are (1 − Net Gini); (1 − Gross Gini); a binary equality indicator which takes a value of 1, if (1 − Net Gini) is above the median of the respective distribution for a given year, and 0 else; (1 − Human Capital Gini Gini). Columns (5)–(7) split the sample with respect to quintiles of the distribution of (1 − Net Gini): results are shown for the lowest quintile in Column (5), the second to fourth quintile in Column (6), and the highest quintile in Column (7). Control variables are log GDP p.c. and average years of schooling. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). For Columns (1)–(4) and (6), instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables. In Columns (5) and (7), instruments for models are limited up to the second lag of the lagged dependent variable and up to the first lag of the explanatory variables. In Column (5), the second lag of the explanatory variables is used additionally for efficiency gains. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.12: Robustness: Binary Democracy and Equality Indicators

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Economic Freedom									
L.Inst. Quality	0.71***	0.71***	0.49***	0.49***	0.84***	0.83***	0.74***	0.71***	0.71***	0.68***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.08)	(0.08)	(0.08)	(0.07)	(0.06)	(0.05)
L.Equality	0.01	-0.01	-0.00	-0.01	-0.00	-0.01	0.00	-0.01	0.01	-0.01
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
L.Democracy	0.01	-0.00	0.03**	0.01	0.03**	0.02	0.03*	0.03**	0.02*	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
L.(Eq×Demo)		0.02*		0.04***		0.03**		0.03**		0.03**
		(0.01)		(0.01)		(0.02)		(0.01)		(0.01)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.97	0.88	0.91	0.83
Hansen p–value							0.12	0.39	0.11	0.21
Diff.–in–Hansen p–value									0.45	0.27
Instruments							61	79	67	84
Groups	96	96	96	96	96	96	96	96	96	96
Observations	546	546	546	546	546	546	450	450	546	546
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
L.Inst. Quality	0.77***	0.76***	0.55***	0.54***	0.84***	0.82***	0.73***	0.70***	0.76***	0.68***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.05)	(0.04)
L.Equality	0.01*	-0.01	-0.01	-0.02	-0.02	-0.02	-0.01	-0.03	0.01	-0.02
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
L.Democracy	0.02**	0.01	0.03**	0.02	0.01	0.01	0.00	0.00	0.00	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
L.(Eq×Demo)		0.03**		0.03		0.01		0.03		0.05**
		(0.02)		(0.02)		(0.02)		(0.03)		(0.02)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.39	0.35	0.31	0.29
Hansen p–value							0.06	0.14	0.06	0.30
Diff.–in–Hansen p–value									0.31	0.97
Instruments							61	79	67	84
Groups	96	96	96	96	96	96	96	96	96	96
Observations	546	546	546	546	546	546	450	450	546	546

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). The equality indicator is constructed to take a value of 1, if the time average of $(1 - \text{Net Gini})$ for a specific country is above the median of the time averaged overall distribution, and 0 otherwise. Democracy is proxied by the Democracy-Dictatorship measure of Cheibub, Gandhi, and Vreeland (2010). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.13: Robustness: Multiple Imputation

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Economic Freedom									
L.Inst. Quality	0.71***	0.71***	0.48***	0.48***	0.83***	0.81***	0.70***	0.65***	0.74***	0.68***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.09)	(0.08)	(0.06)	(0.05)	(0.06)	(0.04)
L.Equality	0.03	-0.05	0.02	-0.12	-0.01	-0.14*	0.01	-0.14	-0.01	-0.18*
	(0.02)	(0.08)	(0.06)	(0.10)	(0.05)	(0.07)	(0.08)	(0.11)	(0.07)	(0.11)
L.Democracy	0.02	-0.05	0.04***	-0.13*	0.04**	-0.12*	0.04**	-0.11	0.03	-0.14*
	(0.01)	(0.05)	(0.02)	(0.07)	(0.015)	(0.06)	(0.02)	(0.08)	(0.02)	(0.08)
L.(Eq×Demo)		0.11		0.31**		0.29**		0.29**		0.32**
		(0.10)		(0.12)		(0.10)		(0.14)		(0.15)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.80	0.71	0.87	0.68
Hansen p–value							0.15	0.39	0.10	0.28
Diff.–in–Hansen p–value									0.63	0.86
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	553	553	553	553	553	553	447	447	553	553
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
L.Inst. Quality	0.78***	0.77***	0.56***	0.54***	0.89***	0.84***	0.75***	0.69***	0.79***	0.71***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)	(0.05)
L.Equality	0.07**	-0.06	0.00	-0.17	-0.02	-0.14	0.05	-0.15	0.09	-0.11
	(0.04)	(0.08)	(0.08)	(0.11)	(0.07)	(0.09)	(0.10)	(0.14)	(0.07)	(0.12)
L.Democracy	0.02	-0.09*	0.02	-0.18**	-0.02	-0.16*	-0.01	-0.17*	-0.01	-0.20**
	(0.01)	(0.05)	(0.02)	(0.07)	(0.02)	(0.08)	(0.03)	(0.10)	(0.02)	(0.10)
L.(Eq×Demo)		0.19**		0.38***		0.27*		0.31*		0.36**
		(0.09)		(0.13)		(0.14)		(0.18)		(0.17)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.16	0.14	0.13	0.10
Hansen p–value							0.02	0.13	0.07	0.20
Diff.–in–Hansen p–value									0.87	0.73
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	553	553	553	553	553	553	447	447	553	553

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by ($1 -$ Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.14: Robustness: Gross Gini

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Economic Freedom									
L.Inst. Quality	0.71***	0.71***	0.48***	0.48***	0.82***	0.81***	0.71***	0.67***	0.73***	0.69***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.09)	(0.09)	(0.06)	(0.05)	(0.06)	(0.04)
L.Equality	0.04	-0.01	0.06	-0.08	0.03	-0.09	0.07	-0.11	0.03	-0.12
	(0.03)	(0.07)	(0.06)	(0.10)	(0.05)	(0.07)	(0.07)	(0.13)	(0.06)	(0.10)
L.Democracy	0.02	-0.03	0.04**	-0.10	0.04***	-0.09	0.04*	-0.10	0.02	-0.11
	(0.01)	(0.05)	(0.02)	(0.06)	(0.01)	(0.06)	(0.02)	(0.08)	(0.02)	(0.07)
L.(Eq×Demo)		0.09		0.26**		0.24**		0.28*		0.27**
		(0.09)		(0.12)		(0.10)		(0.15)		(0.12)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.87	0.74	0.88	0.71
Hansen p–value							0.21	0.39	0.07	0.19
Diff.–in–Hansen p–value									0.46	0.80
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
L.Inst. Quality	0.78***	0.78***	0.56***	0.54***	0.88***	0.84***	0.75***	0.71***	0.78***	0.75***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)	(0.04)
L.Equality	0.07*	-0.03	0.08	-0.10	0.08	-0.05	0.11	-0.07	0.15*	-0.03
	(0.04)	(0.08)	(0.08)	(0.11)	(0.07)	(0.10)	(0.08)	(0.14)	(0.08)	(0.10)
L.Democracy	0.02	-0.07	0.02	-0.15**	-0.02	-0.14*	-0.02	-0.16*	-0.02	-0.18**
	(0.01)	(0.05)	(0.02)	(0.07)	(0.02)	(0.08)	(0.02)	(0.09)	(0.02)	(0.08)
L.(Eq×Demo)		0.16*		0.34**		0.26*		0.30*		0.32**
		(0.09)		(0.13)		(0.14)		(0.17)		(0.14)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.17	0.14	0.14	0.12
Hansen p–value							0.06	0.14	0.14	0.42
Diff.–in–Hansen p–value									0.96	0.93
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Civil Liberties in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by $(1 - \text{Gross Gini})$. Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.15: Robustness: Human Capital Equality

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(a) Dependent Variable: Economic Freedom										
L.Inst. Quality	0.73***	0.73***	0.54***	0.52***	0.85***	0.83***	0.72***	0.70***	0.73***	0.71***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.08)	(0.08)	(0.05)	(0.05)	(0.05)	(0.04)
L.Equality	0.01	0.00	0.10*	0.09	0.08	0.08	0.04	0.03	0.05	-0.00
	(0.03)	(0.03)	(0.06)	(0.06)	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.06)
L.Democracy	0.02*	0.01	0.04**	-0.03	0.04***	-0.01	0.04**	-0.00	0.03	0.03
	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)
L.(Eq×Demo)		0.02		0.12**		0.08		0.07		0.01
		(0.04)		(0.05)		(0.05)		(0.06)		(0.06)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.59	0.56	0.56	0.50
Hansen p–value							0.04	0.14	0.04	0.16
Diff.–in–Hansen p–value									0.15	0.64
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	639	639	639	639	639	639	543	543	639	639
(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)										
L.Inst. Quality	0.80***	0.80***	0.59***	0.58***	0.87***	0.87***	0.79***	0.77***	0.78***	0.80***
	(0.03)	(0.03)	(0.06)	(0.06)	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)
L.Equality	0.05	0.05	0.19**	0.19**	0.16*	0.16	0.27***	0.17*	0.23***	0.15***
	(0.04)	(0.04)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.09)	(0.08)	(0.06)
L.Democracy	0.00	-0.01	0.01	-0.04	-0.03	-0.03	-0.02	-0.03	-0.02	-0.00
	(0.02)	(0.03)	(0.02)	(0.05)	(0.02)	(0.05)	(0.03)	(0.05)	(0.03)	(0.05)
L.(Eq×Demo)		0.02		0.09		0.01		0.02		-0.04
		(0.05)		(0.08)		(0.07)		(0.09)		(0.08)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.26	0.24	0.25	0.22
Hansen p–value							0.09	0.11	0.11	0.18
Diff.–in–Hansen p–value									0.73	0.72
Instruments							62	82	68	89
Groups	96	96	96	96	96	96	96	96	96	96
Observations	638	638	638	638	638	638	542	542	638	638

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by $(1 - \text{Human Capital Gini})$. Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.16: Robustness: Different Equality Sub-Samples

	Dependent Variable: Economic Freedom				
	Random Effects (1)	Fixed Effects (2)	Corr. Fixed Effects (3)	Differences GMM (4)	System GMM (5)
	(a) Lowest Quintile				
L.Inst. Quality	0.62***	0.49***	1.23**	0.68***	0.72***
	(0.05)	(0.10)	(0.53)	(0.11)	(0.08)
L.Democracy	0.00	-0.00	-0.02	0.05	-0.03
	(0.02)	(0.04)	(0.10)	(0.03)	(0.02)
AR(2) test				0.82	0.78
Hansen p–value				0.50	0.42
Diff.–in–Hansen p–value					0.94
Instruments				30	39
Groups	38	38	38	35	38
Observations	126	126	126	109	126
	(b) Second to Fourth Quintile				
L.Inst. Quality	0.76***	0.44***	0.82***	0.69***	0.81***
	(0.04)	(0.07)	(0.17)	(0.08)	(0.08)
L.Democracy	0.02	0.06**	0.07**	0.07**	0.06***
	(0.01)	(0.03)	(0.03)	(0.03)	(0.02)
AR(2) test				0.99	0.94
Hansen p–value				0.13	0.01
Diff.–in–Hansen p–value					0.09
Instruments				42	47
Groups	80	80	80	77	80
Observations	324	324	324	280	324
	(c) Highest Quintile				
L.Inst. Quality	0.65***	0.45***	1.14	0.68***	0.69***
	(0.05)	(0.07)	(1.28)	(0.10)	(0.10)
L.Democracy	0.09**	0.09*	0.10	0.14***	0.10
	(0.04)	(0.05)	(0.17)	(0.04)	(0.06)
AR(2) test				0.84	0.73
Hansen p–value				0.32	0.68
Diff.–in–Hansen p–value					0.97
Instruments				22	36
Groups	29	29	29	29	29
Observations	114	114	114	99	114
Controls	✓	✓	✓	✓	✓
Oil/Socialist	✓				✓
FE & TE	✓	✓	✓	✓	✓

Notes: Time periods are five-year intervals. The sample is split with respect to the distribution of income equality proxied by (1 − Net Gini). All regressions include country-fixed and time effects. The dependent variable is Economic Freedom. Democracy is measured by the Constraints on Executive indicator. Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). For Panel (b), the instrument set corresponds to the baseline specification. For Panels (a) and (c), instruments for models are limited up to the second lag of the lagged dependent variable and up to the first lag of the explanatory variables for both, differences and system GMM. In the levels equation of SGMM in Panels (a) and (c), the lagged difference of the regressors are used. In Panel (a), the second lag of the explanatory variables is used additionally for efficiency gains. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.17: Robustness: Different Equality Sub-Samples II

	Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)				
	Random Effects (1)	Fixed Effects (2)	Corr. Fixed Effects (3)	Differences GMM (4)	System GMM (5)
	(a) Lowest Quintile				
L.Inst. Quality	0.73***	0.48***	0.84***	0.04	0.81***
	(0.04)	(0.10)	(0.32)	(0.25)	(0.15)
L.Democracy	-0.03	-0.04	-0.08	-0.08**	-0.13**
	(0.03)	(0.03)	(0.07)	(0.03)	(0.06)
AR(2) test				0.96	0.59
Hansen p–value				0.36	0.30
Diff.–in–Hansen p–value					0.90
Instruments				30	39
Groups	38	38	38	37	38
Observations	125	125	125	109	125
	(b) Second to Fourth Quintile				
L.Inst. Quality	0.77***	0.47***	0.82***	0.45***	0.62***
	(0.04)	(0.11)	(0.11)	(0.10)	(0.12)
L.Democracy	0.03	0.07**	0.03	-0.05**	-0.06**
	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)
AR(2) test				0.01	0.00
Hansen p–value				0.13	0.08
Diff.–in–Hansen p–value					0.04
Instruments				42	47
Groups	80	80	80	78	80
Observations	324	324	324		324
	(c) Highest Quintile				
L.Inst. Quality	0.66***	0.53***	1.00	0.25***	0.47***
	(0.06)	(0.08)	(0.63)	(0.09)	(0.11)
L.Democracy	0.13***	0.20***	0.05	0.26**	0.16*
	(0.05)	(0.07)	(0.15)	(0.11)	(0.09)
AR(2) test				0.36	0.21
Hansen p–value				0.11	0.83
Diff.–in–Hansen p–value					1.00
Instruments				21	36
Groups	29	29	29	28	29
Observations	114	114	114	101	114
Controls	✓	✓	✓	✓	✓
Oil/Socialist	✓				✓
FE & TE	✓	✓	✓	✓	✓

Notes: Time periods are five-year intervals. The sample is split with respect to the distribution of income equality proxied by (1 − Net Gini). All regressions include country-fixed and time effects. The dependent variable is the principal component of Economic Freedom and Civil Liberties. Democracy is measured by the Constraints on Executive indicator. Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). For Panel (b), the instrument set corresponds to the baseline specification. For Panels (a) and (c), instruments for models are limited up to the second lag of the lagged dependent variable and up to the first lag of the explanatory variables for both, differences and system GMM. In the levels equation of SGMM in Panels (a) and (c), the lagged difference of the regressors are used. In Panel (a), the second lag of the explanatory variables is used additionally for efficiency gains. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.18: Robustness: Reversed Top-10-Percent Income Share as Equality Measure

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Economic Freedom									
L.Inst. Quality	0.81***	0.77***	0.44***	0.40***	1.55**	1.56**	0.81***	0.84***	0.51**	0.74***
	(0.03)	(0.04)	(0.07)	(0.07)	(0.68)	(0.70)	(0.13)	(0.22)	(0.22)	(0.11)
L.Equality	0.06	-1.04**	0.02	-0.95**	-0.22	0.65	-0.27	0.29	0.18*	-0.83
	(0.05)	(0.41)	(0.17)	(0.38)	(2.04)	(6.30)	(0.35)	(1.83)	(0.09)	(0.69)
L.Democracy	0.00	-0.77***	0.07*	-0.70**	0.07	0.78	0.19*	0.74	0.08	-0.74*
	(0.03)	(0.27)	(0.04)	(0.29)	(0.78)	(5.38)	(0.11)	(1.31)	(0.06)	(0.42)
L.(Eq×Demo)		1.12***		1.09**		-1.02		-0.71		1.05*
		(0.40)		(0.42)		(7.96)		(1.76)		(0.58)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.11	0.16	0.47	0.28
Hansen p–value							0.74	0.72	0.99	0.99
Explained variance							0.79	0.86	0.77	0.81
Sampling adequacy							0.74	0.76	0.84	0.78
Instruments							15	17	22	24
Groups	19	19	19	19	19	19	19	19	19	19
Observations	117	117	117	117	117	117	98	98	117	117
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
L.Inst. Quality	0.68***	0.61***	0.33***	0.26**	0.75**	0.66**	0.56***	0.49***	0.56***	0.32*
	(0.04)	(0.05)	(0.11)	(0.11)	(0.34)	(0.33)	(0.10)	(0.16)	(0.17)	(0.18)
L.Equality	0.02	-1.93***	0.06	-1.92***	-0.11	-1.37	-0.45	-4.74	0.14	-3.24
	(0.09)	(0.52)	(0.23)	(0.50)	(0.34)	(1.07)	(0.42)	(5.74)	(0.21)	(2.89)
L.Democracy	0.09**	-1.25***	0.21***	-1.35***	0.07	-0.91	0.35	-2.82	0.04	-1.95
	(0.04)	(0.33)	(0.06)	(0.35)	(0.20)	(0.78)	(0.24)	(4.15)	(0.12)	(1.97)
L.(Eq×Demo)		1.98***		2.24***		1.42		4.01		3.19
		(0.51)		(0.48)		(1.12)		(5.27)		(2.80)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.39	1.00	0.31	0.71
Hansen p–value							0.73	0.52	0.50	0.70
Explained variance							0.81	0.88	0.78	0.83
Sampling adequacy							0.77	0.77	0.85	0.79
Instruments							15	17	22	25
Groups	19	19	19	19	19	19	19	19	19	19
Observations	117	117	117	117	117	117	98	98	117	117

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Top-10% Income Share). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instrument sets are collapsed and principal component analysis is used to choose the instruments that explain the largest share of the variation based on their eigenvalues. *Explained variance* reports the portion of the variation the extracted components explain. *Sampling adequacy* refers to the Kaiser-Meyer-Olkin measure of sampling adequacy. It ranges from 0 to 1 with small values indicating that variables have not enough correlation in common to warrant a PCA analysis. A widely accepted convention in judging sampling adequacy is: 0.00–0.49 unacceptable; 0.50–0.59 miserable; 0.60–0.69 mediocre; 0.70–0.79 middling; 0.80–0.89 meritorious; 0.90–1.00 marvelous. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.19: Robustness: GMM Specifications

	More Lags	Fewer Lags	Minimum Y	Collapsed Y	Minimum X & Y	Collapsed X & Y
	(1)	(2)	(3)	(4)	(5)	(6)
L.Inst. Quality	0.61***	0.66***	0.84***	0.74***	0.68***	0.73***
	(0.05)	(0.06)	(0.07)	(0.06)	(0.09)	(0.05)
L.Equality	-0.21*	-0.21**	-0.21*	-0.16	-0.29	-0.03
	(0.12)	(0.10)	(0.11)	(0.10)	(0.19)	(0.10)
L.Democracy	-0.17*	-0.17**	-0.17**	-0.13*	-0.26	-0.07
	(0.09)	(0.08)	(0.08)	(0.07)	(0.16)	(0.08)
L.(Eq×Demo)	0.40**	0.38***	0.40***	0.33**	0.59*	0.24*
	(0.17)	(0.14)	(0.14)	(0.13)	(0.31)	(0.14)
Controls	✓	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓
AR(2) p–value	0.60	0.70	0.68	0.69	0.55	0.70
Hansen p–value	0.42	0.15	0.16	0.22	0.15	0.30
Instruments	102	58	19	19	37	40
Groups	96	96	96	96	96	96
Observations	447	447	447	447	447	447

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. Results are computed using the differences GMM estimator by Arellano and Bond (1991). The dependent variable is Economic Freedom. Democracy is measured by the Constraints on Executive indicator. Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Relative to the baseline specification, the specifications in Columns (1) and (2) add/subtract one lag period for the dependent and the explanatory variables. Column (3) uses only the first lag of the lagged dependent variable as instrument. Column (4) collapses the instrument set for the lagged dependent variable. In Column (5) only the first lag of the lagged dependent variable and the explanatory variables is used for instrumentation. Column (6) collapses the instrument set for all explanatory variables including the lagged dependent variable. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.20: Robustness: Alternative Specification of GMM Estimators

	Dependent Variable: Economic Freedom							
	Baseline		More Lags		Fewer Lags		Collapsed	
	FOD	FD	FOD	FD	FOD	FD	FOD	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(a) Differences GMM							
L.Inst. Quality	0.65***	0.63***	0.61***	0.55***	0.66***	0.56***	0.73***	0.63***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)
L.Equality	-0.16	-0.20*	-0.21*	-0.20*	-0.21**	-0.23**	-0.03	-0.17
	(0.11)	(0.11)	(0.12)	(0.12)	(0.10)	(0.11)	(0.10)	(0.11)
L.Democracy	-0.14	-0.17**	-0.17*	-0.18**	-0.17**	-0.22***	-0.07	-0.14*
	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)
L.(Eq×Demo)	0.32**	0.37**	0.40**	0.41***	0.38***	0.45***	0.24*	0.37**
	(0.14)	(0.14)	(0.17)	(0.16)	(0.14)	(0.14)	(0.14)	(0.15)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist								
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value	0.71	0.71	0.60	0.62	0.70	0.67	0.70	0.68
Hansen p–value	0.39	0.25	0.42	0.44	0.15	0.25	0.30	0.13
Diff.–in–Hansen p–value								
Instruments	82	82	102	102	58	58	40	40
Groups	96	96	96	96	96	96	96	96
Observations	447	435	447	435	447	435	447	435
	(b) System GMM							
L.Inst. Quality	0.72***	0.72***	0.72***	0.73***	0.73***	0.76***	0.71***	0.72***
	(0.03)	(0.04)	(0.03)	(0.05)	(0.05)	(0.06)	(0.04)	(0.05)
L.Equality	-0.21**	-0.23	-0.19	-0.23*	-0.37**	-0.34*	-0.13	-0.15
	(0.10)	(0.17)	(0.12)	(0.13)	(0.15)	(0.20)	(0.12)	(0.12)
L.Democracy	-0.15*	-0.18	-0.14*	-0.20**	-0.24**	-0.26*	-0.07	-0.07
	(0.08)	(0.12)	(0.09)	(0.09)	(0.11)	(0.14)	(0.07)	(0.08)
L.(Eq×Demo)	0.35**	0.40*	0.32**	0.41**	0.51**	0.54**	0.22*	0.21
	(0.13)	(0.22)	(0.16)	(0.16)	(0.20)	(0.27)	(0.14)	(0.15)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓	✓	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value	0.65	0.66	0.69	0.68	0.61	0.65	0.64	0.61
Hansen p–value	0.28	0.24	0.99	1.00	0.31	0.09	0.71	0.36
Diff.–in–Hansen p–value	0.83	0.59	1.00	1.00	0.71	0.63	0.91	0.66
Instruments	89	89	137	137	68	68	73	73
Groups	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	543	543

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by $(1 - \text{Net Gini})$. Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. Panel (a) reports estimates for differences GMM and Panel (b) for system GMM. Specifications differ with respect to the transformation matrix that removes the country-fixed effect, which is either forward orthogonalized deviations (FOD) or first differences (FD). Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). For the baseline models, instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM in. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. Models with *more lags* use an additional lag (lagged difference) compared to the baseline, while models with *fewer lags* use the minimum specification of exactly one lag (lagged difference) per variable. Collapsed specifications create instruments for each variable and lag distance rather than for each time period. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.21: Robustness: Alternative Specification of GMM Estimators II

	Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)							
	Baseline		More Lags		Fewer Lags		Collapsed	
	FOD	FD	FOD	FD	FOD	FD	FOD	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(a) Differences GMM							
L.Inst. Quality	0.69***	0.54***	0.66***	0.54***	0.68***	0.51***	0.73***	0.64***
	(0.06)	(0.07)	(0.06)	(0.06)	(0.07)	(0.08)	(0.07)	(0.07)
L.Equality	-0.16	-0.38***	-0.20	-0.41***	-0.22	-0.33**	0.08	-0.21
	(0.15)	(0.10)	(0.14)	(0.12)	(0.18)	(0.14)	(0.23)	(0.17)
L.Democracy	-0.18*	-0.39***	-0.20**	-0.39***	-0.24**	-0.34***	-0.13	-0.26**
	(0.10)	(0.08)	(0.09)	(0.09)	(0.11)	(0.09)	(0.13)	(0.11)
L.(Eq×Demo)	0.33*	0.67***	0.36**	0.69***	0.42**	0.58***	0.22	0.42**
	(0.19)	(0.15)	(0.17)	(0.16)	(0.20)	(0.16)	(0.24)	(0.20)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist								
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value	0.14	0.13	0.13	0.13	0.13	0.14	0.12	0.16
Hansen p–value	0.13	0.11	0.43	0.43	0.04	0.12	0.01	0.01
Diff.–in–Hansen p–value								
Instruments	82	82	102	102	58	58	40	40
Groups	96	96	96	96	96	96	96	96
Observations	447	435	447	435	447	435	447	435
	(b) System GMM							
L.Inst. Quality	0.75***	0.73***	0.76***	0.73***	0.72***	0.76***	0.78***	0.79***
	(0.05)	(0.05)	(0.04)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)
L.Equality	-0.06	-0.12	-0.04	-0.12	-0.04	0.19	0.15	-0.00
	(0.12)	(0.15)	(0.11)	(0.13)	(0.20)	(0.24)	(0.19)	(0.15)
L.Democracy	-0.21***	-0.29***	-0.18**	-0.31***	-0.20*	-0.13	-0.03	-0.14
	(0.08)	(0.11)	(0.08)	(0.09)	(0.11)	(0.13)	(0.12)	(0.09)
L.(Eq×Demo)	0.40***	0.51***	0.31**	0.54***	0.49**	0.24	0.03	0.20
	(0.14)	(0.19)	(0.15)	(0.17)	(0.22)	(0.25)	(0.22)	(0.17)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓	✓	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value	0.09	0.09	0.10	0.09	0.07	0.08	0.12	0.12
Hansen p–value	0.23	0.26	0.99	1.00	0.25	0.22	0.02	0.03
Diff.–in–Hansen p–value	0.91	0.98	1.00	1.00	0.79	0.91	0.57	0.70
Instruments	89	89	137	137	68	68	73	73
Groups	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	543	543

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is the principal component of Economic Freedom and Civil Liberties. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. Panel (a) reports estimates for differences GMM and Panel (b) for system GMM. Specifications differ with respect to the transformation matrix that removes the country-fixed effect, which is either forward orthogonalized deviations (FOD) or first differences (FD). Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). For the baseline models, instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM in. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. Models with *more lags* use an additional lag (lagged difference) compared to the baseline while models with *fewer lags* use the minimum specification of exactly one lag (lagged difference) per variable. Collapsed specifications create instruments for each variable and lag distance rather than for each time period. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.22: Robustness: Parsimonious IV Sets

	Minimum Y		Y Collapsed		Minimum X and Y		X and Y Collapsed		PCA–Selection	
	DGMM	SGMM	DGMM	SGMM	DGMM	SGMM	DGMM	SGMM	DGMM	SGMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Economic Freedom									
L.Inst. Quality	0.84***	0.82***	0.74***	0.74***	0.68***	0.72***	0.73***	0.72***	0.77***	0.70***
	(0.07)	(0.09)	(0.06)	(0.04)	(0.09)	(0.04)	(0.05)	(0.04)	(0.09)	(0.07)
L.Equality	-0.21*	-0.11	-0.16	-0.05	-0.29	-0.22*	-0.03	-0.11	-0.18	-0.05
	(0.11)	(0.11)	(0.10)	(0.10)	(0.19)	(0.12)	(0.10)	(0.14)	(0.13)	(0.11)
L.Democracy	-0.17**	-0.08	-0.13*	-0.05	-0.26	-0.17*	-0.07	-0.06	-0.13	-0.05
	(0.08)	(0.08)	(0.07)	(0.07)	(0.16)	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)
L.(Eq×Demo)	0.40***	0.19	0.33**	0.14	0.59*	0.37**	0.24*	0.21	0.32*	0.13
	(0.14)	(0.14)	(0.13)	(0.13)	(0.31)	(0.15)	(0.14)	(0.15)	(0.18)	(0.14)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist		✓		✓		✓		✓		✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value	0.68	0.78	0.69	0.79	0.55	0.66	0.70	0.64	0.74	0.79
Hansen p–value	0.16	0.00	0.22	0.02	0.15	0.13	0.30	0.66	0.01	0.01
Diff.–in–Hansen p–value		0.01		0.02		0.41		0.88		–
Explained variance									0.91	0.92
Sampling adequacy									0.86	0.86
Instruments	19	29	19	29	37	71	40	71	18	28
Groups	96	96	96	96	96	96	96	96	96	96
Observations	447	543	447	543	447	543	447	543	447	543
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
L.Inst. Quality	0.91***	0.81***	0.79***	0.75***	0.57***	0.77***	0.73***	0.77***	0.84***	0.74***
	(0.09)	(0.10)	(0.09)	(0.07)	(0.08)	(0.05)	(0.07)	(0.05)	(0.07)	(0.06)
L.Equality	-0.18	-0.12	-0.22	-0.05	-0.26	-0.01	0.08	0.27	-0.21	-0.05
	(0.12)	(0.10)	(0.16)	(0.14)	(0.32)	(0.16)	(0.23)	(0.24)	(0.14)	(0.10)
L.Democracy	-0.14*	-0.12*	-0.19**	-0.14*	-0.21	-0.20	-0.13	0.05	-0.16**	-0.10
	(0.09)	(0.06)	(0.10)	(0.08)	(0.18)	(0.13)	(0.13)	(0.16)	(0.08)	(0.07)
L.(Eq×Demo)	0.20	0.21*	0.32*	0.24*	0.61*	0.34	0.22	-0.09	0.24*	0.21*
	(0.16)	(0.12)	(0.17)	(0.15)	(0.33)	(0.21)	(0.24)	(0.28)	(0.15)	(0.11)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist		✓		✓		✓		✓		✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value	0.17	0.13	0.15	0.12	0.08	0.10	0.12	0.11	0.15	0.11
Hansen p–value	0.03	0.02	0.00	0.00	0.08	0.13	0.01	0.02	0.01	0.00
Diff.–in–Hansen p–value		0.04		0.23		0.71		0.54		–
Explained variance									0.91	0.93
Sampling adequacy									0.87	0.87
Instruments	19	29	19	29	37	71	40	71	18	28
Groups	96	96	96	96	96	96	96	96	96	96
Observations	447	543	447	543	447	543	447	543	447	543

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Columns (1), (2), (5), and (6) employ the minimum instrument set with only one lag for the dependent and the explanatory variables, respectively. Columns (3), (4), (7), and (8) employ all possible lags as instruments but collapse the instrument set. In Columns (9) and (10), a principal component analysis is used to choose the instruments that explain the largest share of the variation based on their eigenvalues. *Explained variance* reports the portion of the variation the extracted components explain. *Sampling adequacy* refers to the Kaiser-Meyer-Olkin measure of sampling adequacy. It ranges from 0 to 1 with small values indicating that variables have not enough correlation in common to warrant a PCA analysis. A widely accepted convention in judging sampling adequacy is: 0.00–0.49 unacceptable; 0.50–0.59 miserable; 0.60–0.69 mediocre; 0.70–0.79 middling; 0.80–0.89 meritorious; 0.90–1.00 marvelous. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.23: Robustness: Interval Regressions

	(a) Democracy Indicator: Constraints on Executive						
	Economic Freedom	PC EF & CL	Civil Liberties	Regulation	Soundness of Money	Investment Profile	Protection vs. Corruption
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
L.Inst. Quality	0.47***	0.54***	0.47***	0.42***	0.39***	0.09*	0.17***
	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
L.Equality	-0.34**	-0.20*	-0.14	-0.23*	-0.31*	-0.47**	0.22
	(0.15)	(0.11)	(0.10)	(0.12)	(0.19)	(0.22)	(0.24)
L.Democracy	-0.31***	-0.20***	-0.15**	-0.24***	-0.35**	-0.31**	-0.02
	(0.11)	(0.07)	(0.07)	(0.08)	(0.17)	(0.15)	(0.18)
L.(Eq×Demo)	0.60***	0.42***	0.35***	0.42***	0.74**	0.65**	0.10
	(0.20)	(0.13)	(0.12)	(0.16)	(0.29)	(0.27)	(0.33)
Controls	✓	✓	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓
Groups	96	96	112	96	96	90	90
Observations	543	543	610	539	557	407	407
	(b) Democracy Indicator: Composite PolityIV Score						
	Economic Freedom	PC EF & CL	Civil Liberties	Regulation	Soundness of Money	Investment Profile	Protection vs. Corruption
L.Inst. Quality	0.44***	0.52***	0.48***	0.39***	0.40***	0.09	0.18***
	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
L.Equality	-0.31**	-0.12	-0.10	-0.28**	-0.20	-0.31	0.25
	(0.16)	(0.12)	(0.09)	(0.12)	(0.19)	(0.22)	(0.24)
L.Democracy	-0.25**	-0.10	-0.07	-0.27***	-0.19	-0.12	0.01
	(0.11)	(0.07)	(0.07)	(0.09)	(0.19)	(0.15)	(0.18)
L.(Eq×Demo)	0.56***	0.28**	0.25**	0.49***	0.49	0.37	0.02
	(0.20)	(0.12)	(0.11)	(0.17)	(0.31)	(0.25)	(0.31)
Controls	✓	✓	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓
Groups	96	96	112	96	96	90	90
Observations	550	550	618	555	564	410	410

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. Dependent variables in the respective columns are Economic Freedom; the principal component of Economic Freedom and Civil Liberties; Civil Liberties; Regulation of Credit, Labor, and Business; Soundness of Money; Investment Profile (Property Rights); and Protection against Corruption. Column reports results for fixed effects regressions. Democracy is measured by the Constraints on Executive indicator in Panel (a) and the Combined PolityIV indicator in Panel (b), Equality is proxied by $(1 - \text{Net Gini})$. Control variables are log GDP p.c. and average years of schooling. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.24: Robustness: Interval Regressions II

	(a) Dependent Variable: Economic Freedom						
	Executive Constraints	PC XC & VHC	PolityIV Index	Political Rights	PC PIV & PR & VH	Vanhanen Index	Vanhanen Competition
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
L.Inst. Quality	0.47***	0.48***	0.48***	0.47***	0.48***	0.49***	0.49***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
L.Equality	-0.14	-0.13	-0.10	-0.18*	-0.11	-0.03	-0.09
	(0.10)	(0.09)	(0.09)	(0.10)	(0.09)	(0.07)	(0.08)
L.Democracy	-0.15**	-0.15**	-0.07	-0.18**	-0.13*	-0.22	-0.20**
	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.19)	(0.08)
L.(Eq×Demo)	0.35***	0.38***	0.25**	0.43***	0.38***	0.64**	0.48***
	(0.12)	(0.13)	(0.11)	(0.14)	(0.14)	(0.29)	(0.15)
Controls	✓	✓	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓
Groups	96	96	96	96	96	96	96
Observations	543	543	550	550	550	550	550

	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)						
	Executive Constraints	PC XC & VHC	PolityIV Index	Political Rights	PC PIV & PR & VH	Vanhanen Index	Vanhanen Competition
L.Inst. Quality	0.54***	0.53***	0.52***	0.50***	0.51***	0.55***	0.55***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
L.Equality	-0.20*	-0.17	-0.12	-0.21*	-0.14	-0.05	-0.10
	(0.11)	(0.11)	(0.12)	(0.12)	(0.11)	(0.10)	(0.10)
L.Democracy	-0.20***	-0.18**	-0.10	-0.21***	-0.16*	-0.28	-0.19*
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.23)	(0.10)
L.(Eq×Demo)	0.42***	0.41***	0.28**	0.47***	0.41***	0.68**	0.43**
	(0.13)	(0.13)	(0.12)	(0.12)	(0.13)	(0.33)	(0.17)
Controls	✓	✓	✓	✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓
Groups	96	96	96	96	96	96	96
Observations	543	543	550	550	550	550	550

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy in the respective columns is proxied by Constraints on Executive; an artificial indicator based on the principal component of Constraints and Executive and the Vanhanen Competition indicator; the combined PolityIV indicator; Political Rights; an artificial indicator based on the principal components of the combined PolityIV indicator, Political Rights, and the composite Vanhanen Democracy indicator; the composite Vanhanen indicator; and the Vanhanen Competition indicator. Columns report results for fixed effects regressions. Equality is proxied by $(1 - \text{Net Gini})$. Control variables are log GDP p.c. and average years of schooling. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.25: Robustness: Different Specification of Controls

	Dependent Variable: Economic Freedom									
	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) No Controls									
L.Inst. Quality	0.74***	0.74***	0.47***	0.46***	0.86***	0.83***	0.70***	0.65***	0.76***	0.68***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.10)	(0.10)	(0.06)	(0.06)	(0.05)	(0.04)
L.Equality	0.06***	-0.03	0.06	-0.10	0.02	-0.10	0.08	-0.11	0.06	-0.19**
	(0.02)	(0.07)	(0.05)	(0.09)	(0.06)	(0.07)	(0.07)	(0.11)	(0.06)	(0.10)
L.Democracy	0.03***	-0.04	0.05***	-0.13*	0.04***	-0.11*	0.05**	-0.13	0.04**	-0.17**
	(0.01)	(0.05)	(0.02)	(0.08)	(0.01)	(0.06)	(0.02)	(0.09)	(0.02)	(0.07)
L.(Eq×Demo)		0.12		0.32**		0.27**		0.32**		0.38***
		(0.09)		(0.13)		(0.11)		(0.15)		(0.12)
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.72	0.62	0.77	0.59
Hansen p–value							0.10	0.30	0.05	0.23
Diff.–in–Hansen p–value									0.60	0.70
Instruments							60	80	64	85
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543
	(b) Full Controls									
L.Inst. Quality	0.71***	0.71***	0.49***	0.48***	0.85***	0.82***	0.71***	0.66***	0.72***	0.70***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.10)	(0.09)	(0.07)	(0.06)	(0.06)	(0.05)
L.Equality	0.04*	-0.01	0.03	-0.13	-0.00	-0.15**	0.02	-0.16	0.01	-0.14
	(0.02)	(0.08)	(0.06)	(0.10)	(0.06)	(0.08)	(0.08)	(0.11)	(0.06)	(0.10)
L.Democracy	0.01	-0.03	0.05***	-0.14**	0.04***	-0.13**	0.05**	-0.14	0.02	-0.09
	(0.01)	(0.06)	(0.02)	(0.07)	(0.01)	(0.06)	(0.02)	(0.09)	(0.02)	(0.07)
L.(Eq×Demo)		0.08		0.34***		0.31***		0.33**		0.21*
		(0.11)		(0.12)		(0.11)		(0.14)		(0.12)
Full Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
Colonial History	✓	✓							✓	✓
Ethnic Polarization	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.82	0.71	0.96	0.96
Hansen p–value							0.11	0.33	0.15	0.72
Diff.–in–Hansen p–value									0.45	1.00
Instruments							65	85	73	94
Groups	83	83	96	96	96	96	96	96	83	83
Observations	502	502	543	543	543	543	447	447	502	502

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by $(1 - \text{Net Gini})$. Control variables in Panel (b) are log GDP p.c., average years of schooling, log population size, an inflation dummy that takes a value of 1, if the rate exceeds a threshold of 4 percentage points, and a deflation dummy that takes a value of 1, if prices decrease. Level equations additionally employ dummies for oil export, former socialist countries, colonial history and ethnic polarization. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.26: Robustness: Different Specification of Controls II

	Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) No Controls									
L.Inst. Quality	0.79***	0.77***	0.55***	0.53***	0.92***	0.86***	0.74***	0.68***	0.81***	0.74***
	(0.02)	(0.03)	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)
L.Equality	0.17***	-0.01	0.04	-0.15	0.02	-0.09	0.10	-0.10	0.18*	-0.10
	(0.04)	(0.09)	(0.08)	(0.11)	(0.08)	(0.10)	(0.08)	(0.14)	(0.09)	(0.13)
L.Democracy	0.04**	-0.11*	0.03	-0.18**	-0.02	-0.15*	-0.01	-0.17	-0.00	-0.26***
	(0.02)	(0.06)	(0.02)	(0.07)	(0.02)	(0.09)	(0.02)	(0.10)	(0.02)	(0.09)
L.(Eq×Demo)		0.26***		0.40***		0.26*		0.31		0.47***
		(0.10)		(0.13)		(0.16)		(0.19)		(0.16)
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.13	0.11	0.11	0.08
Hansen p–value							0.04	0.18	0.10	0.25
Diff.–in–Hansen p–value									0.88	0.84
Instruments							60	80	64	85
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543
	(b) Full Controls									
L.Inst. Quality	0.77***	0.75***	0.57***	0.54***	0.91***	0.85***	0.77***	0.71***	0.79***	0.77***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.08)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)
L.Equality	0.06	-0.07	-0.00	-0.22*	-0.04	-0.20*	0.06	-0.18	0.09	-0.11
	(0.04)	(0.09)	(0.09)	(0.12)	(0.08)	(0.10)	(0.09)	(0.14)	(0.08)	(0.14)
L.Democracy	0.01	-0.10*	0.02	-0.22***	-0.03	-0.20**	-0.02	-0.22**	-0.03	-0.18*
	(0.02)	(0.06)	(0.02)	(0.07)	(0.02)	(0.09)	(0.02)	(0.10)	(0.02)	(0.09)
L.(Eq×Demo)		0.21**		0.45***		0.34**		0.37**		0.29*
		(0.10)		(0.13)		(0.15)		(0.18)		(0.17)
Full Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
Colonial History	✓	✓							✓	✓
Ethnic Polarization	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.15	0.14	0.07	0.08
Hansen p–value							0.03	0.16	0.14	0.38
Diff.–in–Hansen p–value									0.65	0.90
Instruments							65	85	73	94
Groups	83	83	96	96	96	96	96	96	83	83
Observations	502	502	543	543	543	543	447	447	502	502

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is the principal component of Economic Freedom and Civil Liberties. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables in Panel (b) are log GDP p.c., average years of schooling, log population size, an inflation dummy that takes a value of 1, if the rate exceeds a threshold of 4 percentage points, and a deflation dummy that takes a value of 1, if prices decrease. Level equations additionally employ dummies for oil export, former socialist countries, colonial history and ethnic polarization. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.27: Robustness: Controlling for Growth

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Economic Freedom									
L.Inst. Quality	0.71***	0.71***	0.48***	0.47***	0.87***	0.84***	0.71***	0.65***	0.77***	0.68***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.10)	(0.10)	(0.06)	(0.05)	(0.06)	(0.04)
L.Equality	0.03	-0.05	0.03	-0.14	-0.00	-0.14**	0.03	-0.16	0.01	-0.21**
	(0.02)	(0.08)	(0.06)	(0.10)	(0.06)	(0.07)	(0.08)	(0.12)	(0.06)	(0.10)
L.Democracy	0.02	-0.05	0.04***	-0.15**	0.04***	-0.13**	0.04**	-0.14	0.03	-0.16**
	(0.01)	(0.06)	(0.02)	(0.07)	(0.01)	(0.06)	(0.02)	(0.09)	(0.02)	(0.08)
L.(Eq×Demo)		0.11		0.35***		0.30***		0.33**		0.36***
		(0.10)		(0.13)		(0.11)		(0.15)		(0.13)
L.Income	-0.00	-0.00	-0.03**	-0.04**	-0.03**	-0.03***	-0.03**	-0.04**	-0.01	-0.01
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
L.Growth	-0.01	-0.01	0.01	0.01	-0.02	-0.02	-0.01	-0.00	-0.03*	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.84	0.71	0.94	0.67
Hansen p–value							0.15	0.35	0.10	0.21
Diff.–in–Hansen p–value									0.73	0.82
Instruments							63	83	69	90
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
L.Inst. Quality	0.78***	0.77***	0.56***	0.53***	0.89***	0.84***	0.74***	0.68***	0.80***	0.71***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.05)
L.Equality	0.08**	-0.06	-0.00	-0.20*	-0.02	-0.16	0.07	-0.16	0.13*	-0.13
	(0.04)	(0.09)	(0.09)	(0.12)	(0.08)	(0.10)	(0.10)	(0.16)	(0.07)	(0.12)
L.Democracy	0.01	-0.09*	0.02	-0.20***	-0.02	-0.17**	-0.01	-0.19*	-0.01	-0.23***
	(0.01)	(0.05)	(0.02)	(0.07)	(0.02)	(0.09)	(0.03)	(0.11)	(0.02)	(0.08)
L.(Eq×Demo)		0.20**		0.42***		0.29*		0.34*		0.42***
		(0.10)		(0.13)		(0.15)		(0.20)		(0.14)
L.Income	0.00	-0.00	-0.04	-0.04*	-0.05***	-0.05***	-0.02	-0.04	0.00	-0.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)	(0.01)
L.Growth	-0.01	-0.01	0.01	0.01	-0.01	-0.01	0.00	0.01	-0.02	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.16	0.12	0.16	0.11
Hansen p–value							0.02	0.13	0.07	0.20
Diff.–in–Hansen p–value									0.86	0.72
Instruments							63	83	69	90
Groups	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	543	543	447	447	543	543

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c., growth of GDP p.c., and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.28: Robustness: Testing Inequality Interaction Against Other Interactions

	Corr. Fixed Effects				Differences GMM				System GMM			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	(a) Dependent Variable: Economic Freedom											
L.Inst. Quality	0.81***	0.81***	0.81***	0.80***	0.65***	0.61***	0.63***	0.62***	0.68***	0.68***	0.70***	0.66***
	(0.09)	(0.09)	(0.09)	(0.09)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.06)	(0.06)	(0.05)
L.Equality	-0.15**	-0.17**	-0.17**	-0.16**	-0.16	-0.21**	-0.22**	-0.22*	-0.21**	-0.37**	-0.36**	-0.39***
	(0.07)	(0.07)	(0.08)	(0.08)	(0.11)	(0.09)	(0.10)	(0.11)	(0.09)	(0.17)	(0.15)	(0.15)
L.Democracy	-0.13**	-0.01	-0.13**	0.03	-0.14	0.01	-0.15**	0.09	-0.16**	-0.07	-0.26**	0.10
	(0.06)	(0.10)	(0.06)	(0.14)	(0.08)	(0.12)	(0.07)	(0.21)	(0.08)	(0.17)	(0.11)	(0.15)
L.Income	-0.04***	-0.02	-0.04***	-0.02	-0.04**	-0.03	-0.04***	-0.02	-0.01*	0.01	-0.02*	0.03*
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)	(0.02)
L.Human Capital	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01**	0.01**	0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
L.(Eq×Demo)	0.31***	0.34***	0.33***	0.33***	0.32**	0.37***	0.40***	0.37**	0.35***	0.66***	0.57***	0.71***
	(0.11)	(0.11)	(0.12)	(0.11)	(0.14)	(0.12)	(0.13)	(0.15)	(0.13)	(0.22)	(0.20)	(0.20)
L.(GDP×Demo)		-0.02		-0.02		-0.02		-0.03		-0.03*		-0.07**
		(0.01)		(0.02)		(0.02)		(0.03)		(0.02)		(0.03)
L.(HC×Demo)			-0.00	0.00			-0.00	0.00			-0.00	0.02**
			(0.00)	(0.01)			(0.01)	(0.01)			(0.01)	(0.01)
AR(2) p–value					0.71	0.62	0.66	0.60	0.66	0.55	0.54	0.46
Hansen p–value					0.39	0.35	0.32	0.46	0.24	0.35	0.42	0.56
Diff.–in–Hansen p–value									0.85	0.61	0.67	0.76
Instruments					82	78	78	92	89	82	82	96
Groups	96	96	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	447	447	447	447	543	543	543	543
	(b) Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)											
L.Inst. Quality	0.83***	0.83***	0.84***	0.85***	0.69***	0.68***	0.68***	0.68***	0.71***	0.68***	0.71***	0.67***
	(0.06)	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)
L.Equality	-0.16	-0.18*	-0.20*	-0.20*	-0.16	-0.23	-0.29	-0.29*	-0.12	-0.09	-0.15	-0.18
	(0.10)	(0.10)	(0.11)	(0.11)	(0.15)	(0.18)	(0.18)	(0.15)	(0.12)	(0.17)	(0.26)	(0.14)
L.Democracy	-0.17**	-0.03	-0.15*	-0.15	-0.18*	-0.11	-0.18*	-0.33	-0.22***	-0.25	-0.23	-0.40
	(0.09)	(0.14)	(0.08)	(0.20)	(0.10)	(0.18)	(0.09)	(0.24)	(0.09)	(0.26)	(0.14)	(0.28)
L.Income	-0.05***	-0.04	-0.05***	-0.05*	-0.04	-0.03	-0.03	-0.05	0.00	-0.02	-0.01	-0.03
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.04)	(0.01)	(0.02)	(0.01)	(0.03)
L.Human Capital	0.01*	0.01*	0.02**	0.02*	0.01*	0.01*	0.02**	0.03**	0.01***	0.01**	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
L.(Eq×Demo)	0.30*	0.32**	0.35**	0.35**	0.33*	0.41**	0.45**	0.47**	0.39**	0.43**	0.51*	0.54***
	(0.15)	(0.16)	(0.16)	(0.16)	(0.19)	(0.20)	(0.20)	(0.19)	(0.15)	(0.21)	(0.28)	(0.19)
L.(GDP×Demo)		-0.02		-0.00		-0.01		0.02		0.01		0.03
		(0.02)		(0.03)		(0.03)		(0.03)		(0.03)		(0.04)
L.(HC×Demo)			-0.01	-0.01			-0.01	-0.02			-0.00	-0.01
			(0.01)	(0.01)			(0.01)	(0.01)			(0.01)	(0.01)
AR(2) p–value					0.14	0.13	0.14	0.14	0.10	0.08	0.09	0.08
Hansen p–value					0.13	0.14	0.12	0.32	0.20	0.28	0.32	0.42
Diff.–in–Hansen p–value									0.73	0.79	0.75	0.71
Instruments					82	78	78	92	89	82	82	96
Groups	96	96	96	96	96	96	96	96	96	96	96	96
Observations	543	543	543	543	447	447	447	447	543	543	543	543
Oil/Socialist									✓	✓	✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is Economic Freedom in Panel (a) and the principal component of Economic Freedom and Civil Liberties in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1−Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments for models with one interaction are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. For models with more than one interaction, instruments are limited to the second lag of the lagged dependent variable and up to the second lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation the lagged difference of the regressors is used. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.29: Effect of Democracy on Equality on Redistribution

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Dependent Variable: Expenditures of Social Security Funds (Percent of GDP)									
L.Inst. Quality	0.49***	0.48***	0.01	0.05	0.81	0.73	0.06	0.08*	0.50**	0.35***
	(0.14)	(0.13)	(0.08)	(0.08)	(2.18)	(1.70)	(0.05)	(0.05)	(0.24)	(0.14)
L.Equality	0.10	0.51*	-0.02	0.58	0.20	0.89	0.02	0.67*	0.12	0.54
	(0.06)	(0.31)	(0.11)	(0.41)	(0.88)	(1.17)	(0.14)	(0.37)	(0.15)	(0.34)
L.Democracy	-0.03	0.25	-0.09	0.32	-0.12	0.39	-0.06	0.36*	0.00	0.26
	(0.04)	(0.18)	(0.07)	(0.20)	(0.10)	(0.38)	(0.08)	(0.21)	(0.02)	(0.19)
L.(Eq×Demo)		-0.46		-0.70*		-0.84		-0.78*		-0.44
		(0.33)		(0.41)		(0.69)		(0.43)		(0.34)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.94	0.88	0.16	0.50
Hansen p–value							0.05	0.10	0.11	0.13
Diff.–in–Hansen p–value									0.38	0.42
Instruments							28	37	34	44
Groups	51	51	51	51	51	51	46	46	51	51
Observations	148	148	148	148	148	148	97	97	148	148
	(b) Dependent Variable: Social Expenditures (Percent of GDP)									
L.Inst. Quality	0.76***	0.76***	0.33*	0.33*	2.93***	2.98***	0.16	0.11	0.79***	0.81***
	(0.09)	(0.10)	(0.17)	(0.17)	(0.01)	(0.01)	(0.18)	(0.35)	(0.17)	(0.17)
L.Equality	0.12*	0.23*	-0.00	0.10	0.20*	0.61***	-0.04	0.05	0.08	0.19
	(0.07)	(0.12)	(0.14)	(0.15)	(0.10)	(0.12)	(0.10)	(0.19)	(0.12)	(0.16)
L.Democracy	-0.00	0.07	0.01	0.09	-0.05**	0.24***	0.02	0.11	0.01	0.04
	(0.02)	(0.05)	(0.03)	(0.07)	(0.03)	(0.09)	(0.02)	(0.09)	(0.02)	(0.08)
L.(Eq×Demo)		-0.13		-0.16		-0.56***		-0.16		-0.10
		(0.11)		(0.14)		(0.16)		(0.16)		(0.16)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.59	0.98	0.07	0.13
Hansen p–value							0.95	0.99	0.96	1.00
Diff.–in–Hansen p–value									1.00	1.00
Instruments							45	57	50	66
Groups	33	33	33	33	33	33	33	33	33	33
Observations	167	167	167	167	167	167	134	134	167	167

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is expenditures of social security funds (in percent of GDP) in Panel (a) and social expenditures (in percent of GDP) in Panel (b). Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c., average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first two lags of the explanatory variables. In the level equation, the lagged difference of the regressors is used so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.30: Robustness: Political Stability 1970–2010

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
L.Democracy	0.77***	0.72***	0.49***	0.37*	0.69***	0.37***	0.73***	0.58	0.78***	1.02***
	(0.04)	(0.11)	(0.06)	(0.19)	(0.06)	(0.06)	(0.08)	(0.40)	(0.07)	(0.38)
L.Equality	0.03	-0.03	0.03	-0.07	0.02	-0.06	0.27	0.07	-0.10	0.60
	(0.07)	(0.18)	(0.16)	(0.22)	(0.16)	(0.27)	(0.20)	(0.43)	(0.21)	(0.46)
L.(Eq×Demo)		0.09		0.21		0.19		-0.37		-1.01
		(0.19)		(0.30)		(0.15)		(0.70)		(0.69)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.63	0.86	0.58	0.80
Hansen p–value							0.02	0.12	0.02	0.01
Diff.–in–Hansen p–value									0.09	0.03
Instruments							42	62	47	68
Groups	112	112	112	112	112	112	112	112	112	112
Observations	618	618	618	618	618	618	506	506	618	618

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The dependent variable is the composite PolityIV indicator. Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for existing oil reserves and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM in. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.31: Robustness: Political Stability 1870–2010 (10-Year Intervals)

	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
L.Democracy	0.55**	-1.85	-0.07	-1.31**	0.16	-0.18	0.09	-0.74	0.44	0.57
	(0.23)	(1.29)	(0.14)	(0.57)	(0.14)	(0.17)	(0.32)	(2.34)	(0.96)	(3.76)
L.Equality	0.23	-3.31*	-0.13	-2.03**	0.64	-0.22	-0.10	-1.11	-1.15	-0.98
	(0.19)	(1.73)	(0.23)	(0.96)	(0.74)	(0.75)	(0.43)	(3.24)	(2.83)	(5.20)
L.(Eq×Demo)		3.52**		1.93**		0.71		0.87		0.02
		(1.66)		(0.92)		(0.46)		(3.62)		(5.88)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.35	0.34	0.49	0.48
Hansen p–value							0.73	1.00	0.91	0.97
Explained variance							0.70	0.84	0.74	0.81
Sampling adequacy							0.80	0.67	0.86	0.80
Instruments							17	21	24	28
Groups	19	19	19	19	19	19	17	17	19	19
Observations	82	82	82	82	82	82	63	63	82	82

Notes: Time periods are ten-year intervals. The dependent variable is the composite PolityIV indicator. Equality is proxied by (1 − Top-10% Income Share). All regressions include country-fixed and time effects and a control variable for log GDP p.c. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instrument sets are collapsed and principal component analysis is used to choose the instruments that explain the largest share of the variation based on their eigenvalues. *Explained variance* reports the portion of the variation the extracted components explain. *Sampling adequacy* refers to the Kaiser-Meyer-Olkin measure of sampling adequacy. It ranges from 0 to 1 with small values indicating that variables have not enough correlation in common to warrant a PCA analysis. A widely accepted convention in judging sampling adequacy is: 0.00–0.49 unacceptable; 0.50–0.59 miserable; 0.60–0.69 mediocre; 0.70–0.79 middling; 0.80–0.89 meritorious; 0.90–1.00 marvelous. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.32: Robustness: Third Wave of Democratization

	Dependent Variable: Economic Freedom									
	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Countries that democratized after 1974									
L.Inst. Quality	0.60***	0.61***	0.37***	0.37***	0.98***	0.94***	0.56***	0.51***	0.65***	0.58***
	(0.04)	(0.04)	(0.06)	(0.05)	(0.19)	(0.19)	(0.08)	(0.09)	(0.08)	(0.08)
L.Equality	0.02	-0.08	0.00	-0.19	-0.07	-0.25*	-0.07	-0.33	-0.02	-0.16
	(0.04)	(0.11)	(0.07)	(0.12)	(0.12)	(0.14)	(0.12)	(0.26)	(0.09)	(0.14)
L.Democracy	0.01	-0.09	0.02	-0.21**	0.01	-0.22*	0.02	-0.23	0.02	-0.12
	(0.02)	(0.08)	(0.02)	(0.08)	(0.02)	(0.12)	(0.02)	(0.16)	(0.02)	(0.11)
L.(Eq×Demo)		0.17		0.41***		0.42**		0.44*		0.23
		(0.15)		(0.15)		(0.21)		(0.27)		(0.19)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.29	0.25	0.33	0.30
Hansen p–value							0.88	1.00	1.00	1.00
Diff.–in–Hansen p–value									1.00	1.00
Instruments							58	74	67	85
Groups	43	43	43	43	43	43	43	43	43	43
Observations	223	223	223	223	223	223	180	180	223	223
	(b) Countries that did not democratize after 1974									
L.Inst. Quality	0.77***	0.76***	0.55***	0.54***	0.93***	0.91***	0.68***	0.63***	0.76***	0.75***
	(0.04)	(0.04)	(0.05)	(0.05)	(0.25)	(0.24)	(0.07)	(0.10)	(0.07)	(0.07)
L.Equality	0.03	-0.00	0.05	-0.09	0.01	-0.08	0.03	-0.10	-0.01	-0.21
	(0.03)	(0.09)	(0.09)	(0.14)	(0.11)	(0.12)	(0.08)	(0.18)	(0.07)	(0.23)
L.Democracy	0.01	-0.02	0.10*	-0.06	0.06	-0.04	0.06	-0.07	0.03	-0.15
	(0.02)	(0.07)	(0.05)	(0.10)	(0.04)	(0.08)	(0.04)	(0.13)	(0.03)	(0.12)
L.(Eq×Demo)		0.04		0.29		0.21		0.28		0.33
		(0.12)		(0.18)		(0.16)		(0.28)		(0.25)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.62	0.61	0.59	0.57
Hansen p–value							0.54	0.97	0.70	0.98
Diff.–in–Hansen p–value									1.00	1.00
Instruments							61	81	67	88
Groups	53	53	53	53	53	53	53	53	53	53
Observations	320	320	320	320	320	320	267	267	320	320

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The sample in Panel (a) includes countries that have democratized after 1974 according to changes in the binary Democracy-Dictatorship indicator by Cheibub, Gandhi, and Vreeland (2010). All other countries are included in Panel (b). The dependent variable is Economic Freedom. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by $(1 - \text{Net Gini})$. Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.33: Robustness: Third Wave of Democratization II

	Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) Countries that democratized after 1974									
L.Inst. Quality	0.73***	0.71***	0.52***	0.49***	0.94***	0.87***	0.76***	0.68***	0.71***	0.69***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.11)	(0.11)	(0.09)	(0.08)	(0.08)	(0.10)
L.Equality	0.08	-0.06	-0.06	-0.27	-0.13	-0.25	-0.01	-0.25	0.09	-0.14
	(0.06)	(0.11)	(0.12)	(0.16)	(0.15)	(0.17)	(0.15)	(0.16)	(0.12)	(0.21)
L.Democracy	-0.05**	-0.17**	-0.03	-0.27***	-0.08**	-0.23	-0.06**	-0.18	-0.06*	-0.29
	(0.02)	(0.07)	(0.02)	(0.09)	(0.03)	(0.15)	(0.03)	(0.13)	(0.03)	(0.19)
L.(Eq×Demo)		0.23*		0.45***		0.30		0.23		0.41
		(0.13)		(0.15)		(0.27)		(0.23)		(0.34)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.16	0.15	0.12	0.07
Hansen p–value							0.94	1.00	0.99	1.00
Diff.–in–Hansen p–value									1.00	1.00
Instruments							58	74	67	85
Groups	43	43	43	43	43	43	43	43	43	43
Observations	223	223	223	223	223	223	180	180	223	223
	(b) Countries that did not democratize after 1974									
L.Inst. Quality	0.79***	0.78***	0.56***	0.55***	0.86***	0.81***	0.65***	0.63***	0.82***	0.75***
	(0.04)	(0.04)	(0.07)	(0.08)	(0.10)	(0.09)	(0.06)	(0.07)	(0.06)	(0.07)
L.Equality	0.07	-0.09	0.11	-0.14	0.07	-0.09	0.14	-0.10	0.06	-0.23
	(0.05)	(0.12)	(0.13)	(0.18)	(0.10)	(0.14)	(0.11)	(0.22)	(0.11)	(0.22)
L.Democracy	0.04*	-0.07	0.04	-0.22	0.01	-0.16	-0.02	-0.22	0.03	-0.23
	(0.02)	(0.08)	(0.07)	(0.16)	(0.05)	(0.11)	(0.05)	(0.19)	(0.04)	(0.16)
L.(Eq×Demo)		0.20		0.50**		0.36*		0.43		0.48*
		(0.14)		(0.24)		(0.21)		(0.33)		(0.26)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.53	0.59	0.63	0.62
Hansen p–value							0.76	0.99	0.82	0.99
Diff.–in–Hansen p–value									1.00	1.00
Instruments							61	81	67	88
Groups	53	53	53	53	53	53	53	53	53	53
Observations	320	320	320	320	320	320	267	267	320	320

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The sample in Panel (a) includes countries that have democratized after 1974 according to changes in the binary Democracy-Dictatorship indicator by Cheibub, Gandhi, and Vreeland (2010). All other countries are included in Panel (b). The dependent variable is the principal component of Economic Freedom and Civil Liberties. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.34: Robustness: OECD and Non-OECD Countries

	Dependent Variable: Economic Freedom									
	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) OECD countries									
L.Inst. Quality	0.75***	0.75***	0.59***	0.58***	1.31**	1.14**	0.70***	0.70***	0.74***	0.73***
	(0.04)	(0.04)	(0.07)	(0.07)	(0.56)	(0.53)	(0.09)	(0.10)	(0.08)	(0.08)
L.Equality	-0.03	-0.26***	0.02	-0.27**	0.27	-0.07	0.05	-0.20	-0.04	-0.25*
	(0.03)	(0.08)	(0.15)	(0.12)	(0.43)	(0.21)	(0.22)	(0.15)	(0.13)	(0.13)
L.Democracy	0.04**	-0.15***	0.06*	-0.25***	-0.03	-0.28*	0.04	-0.31	0.05*	-0.24
	(0.02)	(0.06)	(0.03)	(0.06)	(0.10)	(0.17)	(0.03)	(0.22)	(0.03)	(0.16)
L.(Eq×Demo)		0.31***		0.53***		0.47**		0.58		0.47
		(0.09)		(0.11)		(0.23)		(0.37)		(0.29)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.94	0.78	1.00	0.77
Hansen p–value							1.00	1.00	1.00	1.00
Diff.–in–Hansen p–value									1.00	1.00
Instruments							59	75	64	85
Groups	32	32	32	32	32	32	32	32	32	32
Observations	208	208	208	208	208	208	176	176	208	208
	(b) Non-OECD Countries									
L.Inst. Quality	0.68***	0.68***	0.42***	0.42***	0.89***	0.88***	0.62***	0.55***	0.75***	0.63***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.13)	(0.13)	(0.07)	(0.07)	(0.05)	(0.06)
L.Equality	0.03	-0.02	0.03	-0.09	-0.05	-0.14	-0.04	-0.15	-0.05	-0.23**
	(0.03)	(0.09)	(0.06)	(0.12)	(0.07)	(0.10)	(0.09)	(0.15)	(0.06)	(0.11)
L.Democracy	0.01	-0.03	0.04*	-0.10	0.04**	-0.07	0.04	-0.10	0.03	-0.16*
	(0.01)	(0.06)	(0.02)	(0.09)	(0.02)	(0.09)	(0.02)	(0.13)	(0.02)	(0.09)
L.(Eq×Demo)		0.08		0.25		0.21		0.26		0.32**
		(0.12)		(0.16)		(0.16)		(0.22)		(0.14)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.76	0.68	0.79	0.66
Hansen p–value							0.39	0.96	0.55	0.99
Diff.–in–Hansen p–value									0.88	1.00
Instruments							62	82	68	89
Groups	64	64	64	64	64	64	64	64	64	64
Observations	335	335	335	335	335	335	271	271	335	335

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The sample in Panel (a) includes OECD countries, all non-OECD countries are included in Panel (b). The dependent variable is Economic Freedom. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.35: Robustness: OECD and Non-OECD Countries II

	Dependent Variable: Composite Index (Principal Component of Economic Freedom and Civil Liberties)									
	Random Effects		Fixed Effects		Corr. Fixed Effects		Differences GMM		System GMM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(a) OECD countries									
L.Inst. Quality	0.80***	0.79***	0.64***	0.59***	1.09***	1.00***	0.66***	0.59***	0.69***	0.72***
	(0.04)	(0.04)	(0.06)	(0.06)	(0.26)	(0.23)	(0.11)	(0.11)	(0.08)	(0.09)
L.Equality	0.09	-0.03	0.08	-0.22	0.33	0.14	-0.06	-0.37	0.15	0.01
	(0.06)	(0.14)	(0.17)	(0.14)	(0.22)	(0.25)	(0.33)	(0.37)	(0.14)	(0.26)
L.Democracy	0.01	-0.08	0.05	-0.23***	-0.08	-0.18	0.06	-0.24	0.04	-0.04
	(0.03)	(0.09)	(0.04)	(0.06)	(0.09)	(0.13)	(0.06)	(0.27)	(0.04)	(0.17)
L.(Eq×Demo)		0.16		0.50***		0.22		0.53		0.16
		(0.13)		(0.11)		(0.24)		(0.46)		(0.30)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.02	0.02	0.03	0.02
Hansen p–value							1.00	1.00	1.00	1.00
Diff.–in–Hansen p–value									1.00	1.00
Instruments							59	75	64	85
Groups	32	32	32	32	32	32	32	32	32	32
Observations	208	208	208	208	208	208	176	176	208	208
	(b) Non-OECD Countries									
L.Inst. Quality	0.77***	0.76***	0.52***	0.50***	0.89***	0.85***	0.63***	0.62***	0.67***	0.64***
	(0.04)	(0.04)	(0.08)	(0.08)	(0.09)	(0.09)	(0.07)	(0.08)	(0.07)	(0.08)
L.Equality	0.06	-0.11	-0.02	-0.20	-0.09	-0.20	-0.03	-0.15	0.03	-0.23
	(0.05)	(0.09)	(0.09)	(0.14)	(0.09)	(0.13)	(0.09)	(0.17)	(0.08)	(0.20)
L.Democracy	0.01	-0.13**	0.00	-0.21**	-0.03	-0.18	-0.02	-0.18	-0.01	-0.30**
	(0.02)	(0.06)	(0.02)	(0.10)	(0.02)	(0.12)	(0.02)	(0.16)	(0.02)	(0.14)
L.(Eq×Demo)		0.27**		0.39**		0.30		0.29		0.55**
		(0.11)		(0.17)		(0.22)		(0.28)		(0.24)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Oil/Socialist	✓	✓							✓	✓
FE & TE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AR(2) p–value							0.74	0.65	0.67	0.56
Hansen p–value							0.27	0.82	0.49	0.95
Diff.–in–Hansen p–value									0.90	1.00
Instruments							62	82	68	89
Groups	64	64	64	64	64	64	64	64	64	64
Observations	335	335	335	335	335	335	271	271	335	335

Notes: Time periods are five-year intervals. All regressions include country-fixed and time effects. The sample in Panel (a) includes OECD countries, all non-OECD countries are included in Panel (b). The dependent variable is the principal component of Economic Freedom and Civil Liberties. Democracy is measured by the Constraints on Executive indicator, Equality is proxied by (1 − Net Gini). Control variables are log GDP p.c. and average years of schooling. Level equations additionally employ dummies for oil export and former socialist countries. The variance-covariance matrix of CFE is estimated using bootstrap procedures with 100 repetitions. Preliminary estimates for the autoregressive parameter are obtained from DGMM. Standard errors in GMM are estimated with the two-step procedure and corrected with respect to finite-sample size (Windmeijer, 2005). Instruments are limited up to the second lag of the lagged dependent variable and up to the third lag of the explanatory variables for differences GMM. The differenced equation of SGMM uses the first lag of the lagged dependent variable and the first lag of the explanatory variables. In the level equation, the lagged difference of the regressors is used additionally so that for the variables of interest the same time periods are employed in both equations. The number of observations refers to the untransformed data for system GMM and the transformed data for difference GMM. Asterisks indicate significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table F.36: Estimation Sample

ID	Code	Country	EF	CL	ICRG	ID	Code	Country	EF	CL	ICRG
1	AGO	Angola	0	0	0	75	KOR	South Korea	1	1	1
2	ALB	Albania	1	1	1	76	LAO	Laos	0	1	0
3	ARG	Argentina	1	1	1	77	LBN	Lebanon	0	0	0
4	ARM	Armenia	0	1	0	78	LCA	Saint Lucia	0	0	0
5	AUS	Australia	1	1	1	79	LKA	Sri Lanka	1	1	1
6	AUT	Austria	1	1	1	80	LSO	Lesotho	0	1	0
7	AZE	Azerbaijan	0	0	0	81	LTU	Lithuania	1	1	0
8	BDI	Burundi	1	1	0	82	LUX	Luxembourg	1	1	1
9	BEL	Belgium	1	1	1	83	LVA	Latvia	1	1	0
10	BFA	Burkina Faso	0	0	0	84	MAR	Morocco	1	1	1
11	BGD	Bangladesh	1	1	1	85	MDA	Moldova	0	1	0
12	BGR	Bulgaria	1	1	1	86	MDG	Madagascar	0	0	0
13	BHS	Bahamas	0	0	0	87	MDV	Maldives	0	0	0
14	BIH	Bosnia and Hercegovina	0	0	0	88	MEX	Mexico	1	1	1
15	BLR	Belarus	0	0	0	89	MKD	Macedonia	0	0	0
16	BLZ	Belize	0	0	0	90	MLI	Mali	1	1	1
17	BOL	Bolivia	1	1	1	91	MLT	Malta	0	0	0
18	BRA	Brazil	1	1	1	92	MNG	Mongolia	0	1	1
19	BRB	Barbados	0	0	0	93	MOZ	Mozambique	0	1	1
20	BTN	Bhutan	0	0	0	94	MRT	Mauritania	0	1	0
21	BWA	Botswana	1	1	1	95	MUS	Mauritius	1	1	0
22	CAF	Central African Republic	1	1	0	96	MWI	Malawi	1	1	1
23	CAN	Canada	1	1	1	97	MYS	Malaysia	1	1	1
24	CHE	Switzerland	1	1	1	98	NAM	Namibia	1	1	1
25	CHL	Chile	1	1	1	99	NER	Niger	1	1	1
26	CHN	China	1	1	1	100	NGA	Nigeria	0	0	0
27	CIV	Ivory Coast	1	1	1	101	NIC	Nicaragua	0	0	0
28	CMR	Cameroon	1	1	1	102	NLD	Netherlands	1	1	1
29	COL	Colombia	1	1	1	103	NOR	Norway	1	1	1
30	CPV	Cape Verde	1	1	0	104	NPL	Nepal	1	1	0
31	CRI	Costa Rica	1	1	1	105	NZL	New Zealand	1	1	1
32	CYP	Cyprus	1	1	1	106	PAK	Pakistan	1	1	1
33	CZE	Czech Republic	1	1	1	107	PAN	Panama	1	1	1
34	DEU	Germany	1	1	1	108	PER	Peru	1	1	1
36	DJI	Djibouti	0	0	0	109	PHL	Philippines	1	1	1
37	DNK	Denmark	1	1	1	110	PNG	Papua New Guinea	0	0	0
38	DOM	Dominican Republic	1	1	1	111	POL	Poland	1	1	1
39	DZA	Algeria	0	0	0	112	PRT	Portugal	1	1	1
40	ECU	Ecuador	1	1	1	113	PRY	Paraguay	1	1	1
41	EGY	Egypt	1	1	1	114	PUR	Puerto Rico	0	0	0
42	ESP	Spain	1	1	1	115	ROM	Romania	1	1	1
43	EST	Estonia	1	1	0	116	RUS	Russia	1	1	1
44	ETH	Ethiopia	0	0	0	117	RWA	Rwanda	1	1	0
45	FIN	Finland	1	1	1	118	SDN	Sudan	0	0	0
46	FJI	Fiji	1	1	0	119	SEN	Senegal	1	1	1
47	FRA	France	1	1	1	120	SGP	Singapore	1	1	1
48	GBR	United Kingdom	1	1	1	121	SLE	Sierra Leone	1	1	1
49	GEO	Georgia	0	0	0	122	SLV	El Salvador	1	1	1
50	GHA	Ghana	1	1	1	123	SRB	Serbia	0	1	1
51	GIN	Guinea	0	0	0	124	SUR	Suriname	0	0	0
52	GMB	Gambia	0	1	1	125	SVK	Slovakia	1	1	1
53	GNB	Guinea-Bissau	0	0	0	126	SVN	Slovenia	1	1	0
53	GRC	Greece	1	1	1	127	SWE	Sweden	1	1	1
54	GRD	Grenada	0	0	0	128	SWZ	Swaziland	0	1	0
55	GUY	Guyana	0	0	0	129	SYC	Seychelles	0	0	0
56	HKG	Hong Kong	0	0	0	130	THA	Thailand	1	1	1
57	HND	Honduras	1	1	1	131	TJK	Tajikistan	0	1	0
58	HRV	Croatia	1	1	0	132	TKM	Turkmenistan	0	0	0
59	HTI	Haiti	0	0	0	133	TTO	Trinidad and Tobago	1	1	1
60	HUN	Hungary	1	1	1	134	TUN	Tunisia	1	1	1
61	IDN	Indonesia	1	1	1	135	TUR	Turkey	1	1	1
62	IND	India	1	1	1	136	TWN	Taiwan	1	1	1
63	IRL	Ireland	1	1	1	137	TZA	Tanzania	1	1	1
64	IRN	Iran	1	1	1	138	UGA	Uganda	1	1	1
65	ISL	Iceland	0	0	0	139	UKR	Ukraine	1	1	0
66	ISR	Israel	1	1	1	140	URY	Uruguay	1	1	1
67	ITA	Italy	1	1	1	141	USA	United States	1	1	1
68	JAM	Jamaica	1	1	1	142	UZB	Uzbekistan	0	0	0
69	JOR	Jordan	1	1	1	143	VEN	Venezuela	1	1	1
70	JPN	Japan	1	1	1	144	VNM	Vietnam	0	1	1
71	KAZ	Kazakhstan	0	1	0	145	YEM	Yemen	0	1	1
72	KEN	Kenya	1	1	1	146	ZAF	South Africa	1	1	1
73	KGZ	Kyrgyzstan	0	1	0	147	ZMB	Zambia	1	1	1
74	KHM	Cambodia	0	1	0	148	ZWE	Zimbabwe	1	1	1

Bibliography

Acemoglu, D. (2008): "Oligarchic Versus Democratic Societies," *Journal of the European Economic Association*, 6(1), 1–44.

Acemoglu, D., and D. Autor (2011): "Skills, Tasks, and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 4B, chap. 12, pp. 1043–1171. North Holland, Amsterdam.

Acemoglu, D., and S. Johnson (2005): "Unbundling Institutions," *Journal of Political Economy*, 113(5), 949–995.

——— (2007): "Disease and Development: The Effect of Life Expectancy on Economic Growth," *Journal of Political Economy*, 115(6), 925–985.

——— (2014): "Disease and Development: A Reply to Bloom, Canning, and Fink," *Journal of Political Economy*, 122(6), 1367–1375.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91(5), 1369–1401.

——— (2005): "Institutions as a Fundamental Cause of Long-Run Growth," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, vol. 1A, chap. 6, pp. 385–472. North Holland, Amsterdam.

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2008): "Income and Democracy," *American Economic Review*, 98(3), 808–842.

——— (2009): "Reevaluating the Modernization Hypothesis," *Journal of Monetary Economics*, 56(8), 1043–1058.

Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2014): "Democracy, Redistribution, and Inequality," in *Handbook of Income Distribution*, ed. by A. Atkinson, vol. 2B, chap. 21, pp. 1885–1966. North Holland, Amsterdam.

——— (2016): "Democracy Does Cause Growth," *Journal of Political Economy*, forthcoming.

Acemoglu, D., and P. Restrepo (2017): “Secular Stagnation? The Effect of Aging on Economic Growth in the Age of Automation,” *American Economic Review: Papers & Proceedings*, 107(5), 174–179.

Acemoglu, D., and J. A. Robinson (2000): “Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective,” *Quarterly Journal of Economics*, 115(4), 1167–1199.

——— (2005): *Economic Origins of Dictatorship and Democracy*. Cambridge University Press, New York, NY.

——— (2008): “Persistence of Power, Elites and Institutions,” *American Economic Review*, 98(1), 267–293.

Acemoglu, D., J. A. Robinson, and R. Torvik (2013): “Why Do Voters Dismantle Checks and Balances?,” *Review of Economic Studies*, 80(3), 845–875.

Ager, P., C. W. Hansen, and P. S. Jensen (2018): “Fertility and Early-Life Mortality: Evidence from Smallpox Vaccination in Sweden,” *Journal of the European Economic Association*, 16(2), 487–521.

Aghion, P., and P. Howitt (1992): “A Model of Growth Through Creative Destruction,” *Econometrica*, 60(2), 323–351.

Aghion, P., P. Howitt, and F. Murtin (2011): “The Relationship Between Health and Growth: When Lucas Meets Nelson-Phelps,” *Review of Economics and Institutions*, 2(1), 1–24.

Aidt, T. S., and R. Franck (2015): “Democratization Under the Threat of Revolution: Evidence from the Great Reform Act of 1832,” *Econometrica*, 83(2), 505–547.

Aiyar, S., C. Ebeke, and X. Shao (2016): “The Impact of Workforce Aging on European Productivity,” *IMF Working Paper*, 16/238.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003): “Fractionalization,” *Journal of Economic Growth*, 8(2), 155–194.

Almås, I., T. Havnes, and M. Mogstad (2011): “Baby Booming Inequality? Demographic Change and Earnings Inequality in Norway, 1967–2000,” *Journal of Economic Inequality*, 9(4), 629–650.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools,” *Journal of Political Economy*, 113(1), 151–184.

Anderson, T. W., and C. Hsiao (1981): "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, 76(375), 598–606.

Angrist, J. D., and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.

Arbetman-Rabinowitz, M., A. Fisunoglu, J. Kugler, M. Abdollahian, K. Johnson, K. Kang, and Z. Yang (2013): *Replication Data For: Relative Political Capacity Dataset*. Transresearch Consortium [Distributor] V4 [Version], http://hdl.handle.net/1902.1/16845.

Arellano, M., and S. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58(2), 277–297.

Arellano, M., and O. Bover (1995): "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68(1), 29–51.

Ashenfelter, O. C., and C. E. Rouse (1999): "The Payoff to Education," *mimeo, Princeton University*.

Autor, D., and D. Dorn (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103(5), 1553–1597.

Barro, R. J. (1991): "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106(2), 407–443.

——— (1997): *Determinants of Economic Growth: A Cross-Country Empirical Study*. MIT Press, Cambridge, MA.

——— (1999): "Determinants of Democracy," *Journal of Political Economy*, 107(6), 158–183.

Barro, R. J., and J. W. Lee (2013): "A New Data Set of Educational Attainment in the World, 1950–2010," *Journal of Development Economics*, 104, 184–198.

Barro, R. J., and X. Sala-i-Martin (2004): *Economic Growth*. MIT Press, Cambridge, MA.

Baumol, W. J. (1986): "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show," *American Economic Review*, 76(5), 1072–1085.

Becker, G. S. (1960): "An Economic Analysis of Fertility," in *Demographic and Economic Change in Developed Countries*, ed. by Universities-National Bureau Committee for Economic Research, George B. Roberts, Chairman, chap. 7, pp. 209–240. National Bureau of Economic Research, Cambridge, MA.

——— (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70(5), 9–49.

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75(4), 352–365.

Benhabib, J., A. Corvalan, and M. M. Spiegel (2013): "Income and Democracy: Evidence from Nonlinear Estimations," *Economics Letters*, 118(3), 489–492.

Benhabib, J., and M. M. Spiegel (1994): "The Role of Human Capital in Economic Development Evidence from Aggregate Cross-Country Data," *Journal of Monetary Economics*, 34(2), 143–173.

Bhalotra, S. R., and A. Venkataramani (2015): "Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions," *Working Paper*, available online at SSRN: https://ssrn.com/abstract=1940725.

Bhargava, A., D. T. Jamison, L. J. Lau, and C. J. Murray (2001): "Modeling the Effects of Health on Economic Growth," *Journal of Health Economics*, 20(3), 423–440.

Bils, M., and P. J. Klenow (2000): "Does Schooling Cause Growth?," *American Economic Review*, 90(5), 1160–1183.

Bleakley, H. (2007): "Disease and Development: Evidence from Hookworm Eradication in the American South," *Quarterly Journal of Economics*, 122(1), 73–117.

Bloom, D. E., and D. Canning (2000): "The Health and Wealth of Nations," *Science*, 287(5456), 1207–1209.

Bloom, D. E., D. Canning, and G. Fink (2010): "Implications of Population Ageing for Economic Growth," *Oxford Review of Economic Policy*, 26(4), 583–612.

——— (2014): "Disease and Development Revisited," *Journal of Political Economy*, 122(6), 1355–1366.

Bloom, D. E., D. Canning, and B. Graham (2003): "Longevity and Life-Cycle Savings," *Scandinavian Journal of Economics*, 105(3), 319–338.

Bloom, D. E., D. Canning, R. K. Mansfield, and M. Moore (2007): "Demographic Change, Social Security Systems, and Savings," *Journal of Monetary Economics*, 54(1), 92–114.

Bloom, D. E., D. Canning, and M. Moore (2014): "Optimal Retirement with Increasing Longevity," *Scandinavian Journal of Economics*, 116(3), 838–858.

Bloom, D. E., D. Canning, and J. Sevilla (2003): *The Demographic Dividend: A New Perspective on the Economic Consequences of Population Change*. Rand Corporation, Santa Monica, CA.

——— (2004): "The Effect of Health on Economic Growth: A Production Function Approach," *World Development*, 32(1), 1–13.

Bloom, D. E., M. Kuhn, and K. Prettner (2015): "The Contribution of Female Health to Economic Development," *NBER Working Paper*, 21411.

Bloom, D. E., and J. G. Williamson (1998): "Demographic Transitions and Economic Miracles in Emerging Asia," *World Bank Economic Review*, 12(3), 419–455.

Blundell, R., and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87(1), 115–143.

——— (2000): "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19(3), 321–340.

Börsch-Supan, A., A. Ludwig, and J. Winter (2006): "Ageing, Pension Reform, and Capital Flows: A Multi-Country Simulation Model," *Economica*, 73(292), 625–658.

Börsch-Supan, A., and M. Weiss (2016): "Productivity and Age: Evidence from Work Teams at the Assembly Line," *Journal of the Economics of Ageing*, 7, 30–42.

Brückner, M. (2013): "On the Simultaneity Problem in the Aid and Growth Debate," *Journal of Applied Econometrics*, 28(1), 126–150.

Brückner, M., and A. Ciccone (2011): "Rain and the Democratic Window of Opportunity," *Econometrica*, 79(3), 923–947.

Brückner, M., A. Ciccone, and A. Tesei (2012): "Oil Price Schocks, Income, and Democracy," *Review of Economics and Statistics*, 94(2), 389–399.

Bruno, G. (2005): "Estimation and Inference in Dynamic Unbalanced Panel-Data Models with a Small Number of Individuals," *Stata Journal*, 5(4), 473–500.

Bun, M. J. G., and M. A. Carree (2005): "Bias-Corrected Esitmation in Dynamic Panel Data Models," *Journal of Business and Economic Statistics*, 23(2), 200–210.

Bun, M. J. G., and J. F. Kiviet (2003): "On the Diminishing Returns of Higher-Order Terms in Asymptotic Expansions of Bias," *Economics Letters*, 79(2), 145–152.

——— (2006): "The Effects of Dynamic Feedbacks on LS and MM Estimator Accuracy in Panel Data Models," *Journal of Econometrics*, 132, 409–444.

Bun, M. J. G., and F. Windmeijer (2010): "The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models," *Econometrics Journal*, 13(1), 95–126.

Cameron, A. C., and P. K. Trivedi (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press, New York, NY.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 116(2), 705–746.

Carnevale, A. P., J. Strohl, and A. Gulish (2015): *College Is Just the Beginning: Employers' Role in the $1.1 Trillion Postsecondary Education and Training System*. Center on Education and the Workforce, Georgetown University.

Caselli, F. (2015): "Experience-Biased Technical Change," *CEPR Discussion Paper*, 10752.

Caselli, F., G. Esquivel, and F. Lefort (1996): "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics," *Journal of Economic Growth*, 1(3), 363–389.

Cass, D. (1965): "Optimum Growth in an Aggregative Model of Capital Accumulation," *Review of Economic Studies*, 32(3), 233–240.

Castelló, A., and R. Doménech (2002): "Human Capital Inequality and Economic Growth: Some New Evidence," *Economic Journal*, 112(478), 187–200.

Castelló-Climent, A., and R. Doménech (2014): "Human Capital and Income Inequality: Some Facts and Some Puzzles," *BBVA Research Working Paper*, 12–28.

Catlin, A. C., and C. A. Cowan (2015): *History of Health Spending in the United States, 1960–2013*. Centers for Medicare and Medicaid Services, Baltimore, MD.

CDC (1999a): "Achievements in Public Health, 1900–1999: Tobacco Use – United States, 1900–1999," *Department of Health and Human Services, Centers for Disease Control and Prevention, Morbidity and Mortality Weekly Report*, 48(43), 986–993.

——— (1999b): "Decline in Deaths from Heart Disease and Stroke – United States, 1900–1999," *Department of Health and Human Services, Centers for Disease Control and Prevention, Morbidity and Mortality Weekly Report*, 48(30), 649–656.

——— (2001): "Physical Activity Trends–United States, 1990–1998.," *Department of Health and Human Services, Centers for Disease Control and Prevention, Morbidity and Mortality Weekly Report*, 50(9), 166–169.

——— (2017): "Health-Related Behaviors by Urban-Rural County Classification – United States, 2013," *Department of Health and Human Services, Centers for Disease Control and Prevention, Morbidity and Mortality Weekly Report*, 66(5), 1–8.

Cervellati, M., P. Fortunato, and U. Sunde (2008): "Hobbes to Rousseau: Inequality, Institutions, and Development," *Economic Journal*, 118(531), 1354–1384.

——— (2012): "Consensual and Conflictual Democratization," *B.E. Journal of Theoretical Economics*, 12(1), 1–51.

——— (2014): "Violence During Democratization and the Quality of Democratic Institutions," *European Economic Review*, 66, 226–247.

Cervellati, M., F. Jung, U. Sunde, and T. Vischer (2014): "Income and Democracy: Comment," *American Economic Review*, 104(2), 707–719.

Cervellati, M., and U. Sunde (2005): "Human Capital Formation, Life Expectancy, and the Process of Development," *American Economic Review*, 95(5), 1653–1672.

——— (2011): "Life Expectancy and Economic Growth: The Role of the Demographic Transition," *Journal of Economic Growth*, 16(2), 99–133.

——— (2013): "Life Expectancy, Schooling, and Lifetime Labor Supply: Theory and Evidence Revisited," *Econometrica*, 81(5), 2055–2086.

Cervellati, M., U. Sunde, and K. F. Zimmermann (2017): "Demographic Dynamics and Long-Run Development: Insights for the Secular Stagnation Debate," *Journal of Population Economics*, 30(2), 401–432.

Che, Y., Y. Lu, Z. Tao, and P. Wang (2013): "The Impact of Income on Democracy Revisited," *Journal of Comparative Economics*, 41(1), 159–169.

Cheibub, J. A., J. Gandhi, and J. Vreeland (2010): "Dictatorship and Democracy Revisited," *Public Choice*, 143(1–2), 67–101.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016): "The Association between Income and Life Expectancy in the United States, 2001–2014," *Journal of the American Medical Association*, 315(16), 1750–1766.

Chong, A., and C. Calderon (2000): "Institutional Quality and Income Distribution," *Economic Development and Cultural Change*, 48(4), 761–786.

Chong, A., and M. Gradstein (2007a): "Inequality and Informality," *Journal of Public Economics*, 91(1–2), 159–179.

——— (2007b): "Inequality and Institutions," *Review of Economics and Statistics*, 89(3), 454–465.

Commission on Macroeconomics and Health (2001): *Macroeconomics and Health: Investing in Health for Economic Development*. World Health Organization, Geneva.

Cooper, R., J. Cutler, P. Desvigne-Nickens, S. P. Fortmann, L. Friedman, R. Havlik, G. Hogelin, J. Marler, P. McGovern, and G. Morosco (2000): "Trends and Disparities in Coronary Heart Disease, Stroke, and Other Cardiovascular Diseases in the United States," *Circulation*, 102(25), 3137–3147.

Cuaresma, J. C., E. Loichinger, and G. A. Vincelette (2016): "Aging and Income Convergence in Europe: A Survey of the Literature and Insights from a Demographic Projection Exercise," *Economic Systems*, 40(1), 4–17.

Cuaresma, J. C., W. Lutz, and W. Sanderson (2014): "Is the Demographic Dividend an Education Dividend?," *Demography*, 51(1), 299–315.

Cutler, D., A. Deaton, and A. Lleras-Muney (2006): "The Determinants of Mortality," *Journal of Economic Perspectives*, 20(3), 97–120.

d'Albis, H., S.-H. P. Lau, and M. Sánchez-Romero (2012): "Mortality Transition and Differential Incentives for Early Retirement," *Journal of Economic Theory*, 147(1), 261–283.

Dawber, T. R. (1980): *The Framingham Study: The Epidemiology of Atherosclerotic Disease*. Harvard University Press, Cambridge, MA.

De la Fuente, A., and R. Domenech (2001): "Schooling Data, Technological Diffusion, and the Neoclassical Model," *American Economic Review, Papers & Proceedings*, 91(2), 323–327.

De Nardi, M., E. French, and J. B. Jones (2009): "Life Expectancy and Old Age Savings," *American Economic Review*, 99(2), 110–115.

De Tocqueville, A. (1835): *Democracy in America*. Gosselin, Paris.

Deaton, A. S., and C. H. Paxson (1994): "Intertemporal Choice and Inequality," *Journal of Political Economy*, 102(3), 437–467.

——— (1997): "The Effects of Economic and Population Growth on National Saving and Inequality," *Demography*, 34(1), 97–114.

——— (1998a): "Aging and Inequality in Income and Health," *American Economic Review: Papers & Proceedings*, 88(2), 248–253.

——— (1998b): "Health, Income, and Inequality over the Life Cycle," in *Frontiers in the Economics of Aging*, ed. by D. A. Wise, chap. 10, pp. 431–462. University of Chicago Press, Chicago, IL.

Deininger, K., and L. Squire (1996): "A New Data Set Measuring Income Inequality," *World Bank Economic Review*, 10(3), 565–591.

Diamond, J. (1997): *Guns, Germs, and Steel: The Fates of Human Societies*. W. W. Norton & Company, New York, NY.

Diamond, P. A. (1965): "National Debt in a Neoclassical Growth Model," *American Economic Review*, 55(5), 1126–1150.

Dohmen, T., B. Enke, A. Falk, D. Huffman, and U. Sunde (2015): "Patience and the Wealth of Nations," *mimeo, University of Bonn.*

Domeij, D., and M. Floden (2006): "Population Aging and International Capital Flows," *International Economic Review*, 47(3), 1013–1032.

Doppelhofer, G., R. I. Miller, and X. Sala-i-Martin (2004): "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review*, 94(4), 813–835.

Dorsch, M. T., and P. Maarek (2014): "A Note on Economic Inequality and Democratization," *Peace Economics, Peace Science and Public Policy*, 20(4), 599–610.

——— (2014b): "Recessions, Inequality, and Democratization," *Thema Working Paper*, 2014–19.

Dowrick, S., and M. Rogers (2002): "Classical and Technological Convergence: Beyond the Solow-Swan Growth Model," *Oxford Economic Papers*, 54(3), 369–385.

Durlauf, S. N., P. A. Johnson, and J. R. Temple (2005): "Growth Econometrics," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, vol. 1A, chap. 8, pp. 555–677. North Holland, Amsterdam.

Easterly, W., and R. Levine (1997): "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 112(4), 1203–1250.

Eberhardt, M., and F. Teal (2011): "Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics," *Journal of Economic Surveys*, 25(1), 109–155.

Ettner, S. L. (1996): "New Evidence on the Relationship Between Income and Health," *Journal of Health Economics*, 15(1), 67–85.

Ezzati, M., and E. Riboli (2012): "Can Noncommunicable Diseases Be Prevented? Lessons from Studies of Populations and Individuals," *Science*, 337(6101), 1482–1487.

Fagerberg, J. (1994): "Technology and International Differences in Growth Rates," *Journal of Economic Literature*, 32(3), 1147–1175.

Falck, O., A. Heimisch, and S. Wiederhold (2016): "Returns to ICT Skills," *CESifo Working Paper*, 5720.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. B. Huffman, and U. Sunde (2017): "Global Evidence on Economic Preferences," *Quarterly Journal of Economics*, forthcoming.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015): "The Next Generation of the Penn World Table," *American Economic Review*, 105(10), 3150–3182.

Feyrer, J. (2007): "Demographics and Productivity," *Review of Economics and Statistics*, 89(1), 100–109.

Feyrer, J., D. Politi, and D. N. Weil (2017): "The Cognitive Effects of Micronutrient Deficiency: Evidence from Salt Iodization in the United States," *Journal of the European Economic Association*, 15(2), 355–387.

Field, E., O. Robles, and M. Torero (2009): "Iodine Deficiency and Schooling Attainment in Tanzania," *American Economic Journal: Applied Economics*, 1(4), 140–169.

Fitzenberger, B., and K. Kohn (2006): "Skill Wage Premia, Employment, and Cohort Effects: Are Workers in Germany All of the Same Type?," *IZA Discussion Paper*, 2185.

Flegal, K. M., M. D. Carroll, R. J. Kuczmarski, and C. L. Johnson (1998): "Overweight and Obesity in the United States: Prevalence and Trends, 1960–1994," *International Journal of Obesity*, 22(1), 39–47.

Foege, W. H. (1987): "Public Health: Moving from Debt to Legacy.," *American Journal of Public Health*, 77(10), 1276–1278.

Fogel, R. W. (1994): "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy," *American Economic Review*, 84(3), 369–395.

——— (1997): "New Findings on Secular Trends in Nutrition and Mortality: Some Implications for Population Theory," in *Handbook of Population and Family Economics*, ed. by M. R. Rosenzweig, and O. Stark, vol. 1A, chap. 9, pp. 433–481. North Holland, Amsterdam.

Fortunato, P., and U. Panizza (2015): "Democracy, Education, and the Quality of Government," *Journal of Economic Growth*, 20(4), 333–363.

Freedom House (2014): *Freedom in the World 2014*. Freedom House, New York, NY.

Frijters, P., J. P. Haisken-DeNew, and M. A. Shields (2005): "The Causal Effect of Income on Health: Evidence from German Reunification," *Journal of Health Economics*, 24(5), 997–1017.

Fuller, G. (1995): "The Demographic Backdrop to Ethnic Conflict: A Geographic Overview," in *The Challenge of Ethnic Conflict to National and International Order in the 1990s: Geographic Perspectives*. Central Intelligence Agency, Washington, DC.

Fürnkranz-Prskawetz, A., B. Hammer, and E. Loichinger (2016): *Projections of the Labour Force by Age, Gender, and Highest Level of Educational Attainment until 2050*. Österreichische Akademie der Wissenschaft and Vienna Institute of Demography (OEAW-VID), Vienna.

Gallup, J. L., J. D. Sachs, and A. D. Mellinger (1999): "Geography and Economic Development," *International Regional Science Review*, 22(2), 179–232.

Galor, O. (2005): "From Stagnation to Growth: Unified Growth Theory," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, vol. 1A, chap. 4, pp. 171–293. North Holland, Amsterdam.

——— (2011): *Unified Growth Theory*. Princeton University Press, Princeton, NJ.

Galor, O., and O. Moav (2002): "Natural Selection and the Origin of Economic Growth," *Quarterly Journal of Economics*, 117(4), 1133–1191.

Galor, O., and D. N. Weil (2000): "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review*, 90(4), 806–828.

Glaeser, E. L., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004): "Do Institutions Cause Growth?," *Journal of Economic Growth*, 9(3), 271–303.

Göbel, C., and T. Zwick (2013): "Are Personnel Measures Effective in Increasing Productivity of Old Workers?," *Labour Economics*, 22, 80–93.

Granger, C. W. (1969): "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37(3), 424–438.

——— (1980): "Testing for Causality: A Personal Viewpoint," *Journal of Economic Dynamics and Control*, 2, 329–352.

GRILICHES, Z., AND J. MAIRESSE (1998): “Production Functions: The Search for Identification,” in *Econometrics and Economic Theory in the 20th Century*, ed. by S. Strøm, chap. 6, pp. 169–203. Cambridge University Press, Cambridge.

GROSSMAN, M., AND R. KAESTNER (1997): “Effects of Education on Health,” in *The Social Benefits of Education*, ed. by J. Behrman, and S. Nevzer, vol. 12, chap. 4, pp. 69–107. University of Michigan Press, Ann Arbor, MI.

GROVE, R., AND A. HETZEL (1968): *Vital Statistics Rates in the United States 1940–1960*. U.S. Department of Health, Education, and Welfare, Public Health Service, Washington, DC.

GUVENEN, F., G. KAPLAN, J. SONG, AND J. WEIDNER (2017): “Lifetime Incomes in the United States over Six Decades,” *mimeo, University of Minnesota*.

GWARTNEY, J., R. LAWSON, AND J. HALL (2013): *Economic Freedom of the World: 2013 Annual Report*. Fraser Institute, Vancouver, BC.

——— (2017): *Economic Freedom of the World: 2017 Annual Report*. Fraser Institute, Vancouver, BC.

HALL, R. E., AND C. I. JONES (1999): “Why Do Some Countries Produce So Much More Output Per Worker than Others?,” *Quarterly Journal of Economics*, 114(1), 83–116.

HANSEN, C. W. (2014): “Cause of Death and Development in the U.S.,” *Journal of Development Economics*, 109, 143–153.

HANSEN, C. W., P. S. JENSEN, AND L. LØNSTRUP (2018): “The Fertility Decline in the United States: Schooling and Income,” *Macroeconomic Dynamics*, 22(6), 1584–1612.

HANSEN, C. W., AND L. LØNSTRUP (2015): “The Rise in Life Expectancy and Economic Growth in the 20th Century,” *Economic Journal*, 125(584), 838–852.

HANSEN, C. W., AND H. STRULIK (2017): “Life Expectancy and Education: Evidence from the Cardiovascular Revolution,” *Journal of Economic Growth*, 22(4), 421–450.

HANUSHEK, E. A., J. RUHOSE, AND L. WOESSMANN (2017): “Knowledge Capital and Aggregate Income Differences: Development Accounting for US States,” *American Economic Journal: Macroeconomics*, 9(4), 184–224.

HANUSHEK, E. A., AND L. WOESSMANN (2008): “The Role of Cognitive Skills in Economic Development,” *Journal of Economic Literature*, 46(3), 607–668.

——— (2012): “Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation,” *Journal of Economic Growth*, 17(4), 267–321.

Haughwout, S. P., and M. E. Slater (2017): *Apparent Per Capita Alcohol Consumption: National, State and Regional Trends, 1977–2015*. U.S. Department of Health and Human Services, National Institute on Alcohol Abuse and Alcoholism, Arlington, VA.

Hayakawa, K. (2009): "First-Difference or Forward Orthogonal Deviation: Which Transformation to Use in Dynamic Panel Data Models? A Simulation Study," *Economics Bulletin*, 29(3), 2014–2023.

Hazan, M. (2009): "Longevity and Lifetime Labor Supply: Evidence and Implications," *Econometrica*, 77(6), 1829–1863.

Heckman, J. J., and P. J. Klenow (1997): *Human Capital Policy*. University of Chicago Press, Chicago, IL.

Heid, B., J. Langer, and M. Larch (2012): "Income and Democracy: Evidence from System GMM," *Economics Letters*, 116(2), 166–169.

Hellerstein, J. K., and D. Neumark (1995): "Are Earnings Profiles Steeper than Productivity Profiles? Evidence from Israeli Firm-Level Data," *Journal of Human Resources*, 30(1), 89–112.

Hendrix, C. S. (2010): "Measuring State Capacity: Theoretical and Empirical Implications for the Study of Civil Conflict," *Journal of Peace Research*, 47(3), 273–285.

Hobbs, F., and N. Stoops (2002): *Demographic Trends in the 20th Century*. U.S. Census Bureau, Census 2000 Special Reports, Series CENSR-4, U.S. Government Printing Office, Washington, DC.

HRS (2017): "Health and Retirement Study, (HRS Cross-Wave Geographic Information) Restricted Use Dataset," produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI.

Huntington, S. P. (1993): *The Third Wave: Democratization in the Late Twentieth Century*. University of Oklahoma Press, Oklahoma, OK.

Hyclak, T. J., C. L. Skeels, and L. W. Taylor (2016): "The Cardiovascular Revolution and Economic Performance in the OECD Countries," *Journal of Macroeconomics*, 50, 114–125.

International Labour Organization (2011): *ILO Estimates and Projections of the Economically Active Population: 1990–2020 (Sixth Edition)*. International Labour Organization, Geneva.

Islam, N. (1995): "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110(4), 1127–1170.

Jayachandran, S., and A. Lleras-Muney (2009): "Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines," *Quarterly Journal of Economics*, 124(1), 349–397.

Jenkins, S. P. (2015): "World Income Inequality Databases: An Assessment of WIID and SWIID," *Journal of Economic Inequality*, 13(4), 629–671.

Jones, B. F., and B. A. Weinberg (2011): "Age Dynamics in Scientific Creativity," *Proceedings of the National Academy of Sciences*, 108(47), 18910–18914.

Judson, R. A., and A. L. Owen (1999): "Estimating Dynamic Panel Data Models: A Guide For Macroeconomists," *Economics Letters*, 65(1), 9–15.

Katz, L. F. (1986): "Efficiency Wage Theories: A Partial Evaluation," *NBER Macroeconomics Annual*, 1, 235–276.

Keys, A. (1980): *Seven Countries. A Multivariate Analysis of Death and Coronary Heart Disease.* Harvard University Press, Cambridge, MA.

Keys, A., H. L. Taylor, H. Blackburn, J. Brozek, J. T. Anderson, and E. Simonson (1963): "Coronary Heart Disease Among Minnesota Business and Professional Men Followed Fifteen Years," *Circulation*, 28(3), 381–395.

Kirkwood, T. (2001): *Time of Our Lives: The Science of Human Aging.* Oxford University Press, New York, NY.

Kline, P., and E. Moretti (2013): "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 129(1), 275–331.

Kohl, H. W., and H. D. Cook (2013): *Educating the Student Body: Taking Physical Activity and Physical Education to School.* National Academies Press, Washington, DC.

Kotschy, R. (2018): "Life Expectancy and Life-Cycle Wages: Evidence from the Cardiovascular Revolution in U.S. States," *PGDA Working Paper*, 156.2018.

Kotschy, R., and U. Sunde (2017): "Democracy, Inequality, and Institutional Quality," *European Economic Review*, 91, 209–228.

——— (2018): "Can Education Compensate the Effect of Population Aging on Macroeconomic Performance?," *Economic Policy*, forthcoming.

Krueger, A. B., and M. Lindahl (2001): "Education for Growth: Why and for Whom?," *Journal of Economic Literature*, 39(4), 1101–1136.

Krueger, A. B., and L. H. Summers (1988): "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, 56(2), 259–293.

Kulshreshtha, A., A. Goyal, K. Dabhadkar, E. Veledar, and V. Vaccarino (2014): "Urban-Rural Differences in Coronary Heart Disease Mortality in the United States: 1999–2009," *Public Health Reports*, 129(1), 19–29.

Kuznets, S. (1955): "Economic Growth and Income Inequality," *American Economic Review*, 45(1), 1–28.

Landes, D. S. (1990): "Why Are We So Rich and They So Poor?," *American Economic Review: Papers & Proceedings*, 80(2), 1–13.

Law, M. R., N. J. Wald, and S. Thompson (1994): "By How Much and How Quickly Does Reduction in Serum Cholesterol Concentration Lower Risk of Ischaemic Heart Disease?," *BMJ*, 308(6925), 367–372.

Levine, R., and D. Renelt (1992): "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 82(4), 942–963.

Lindahl, M. (2005): "Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income," *Journal of Human Resources*, 40(1), 144–168.

Lindh, T., and B. Malmberg (1999): "Age Structure Effects and Growth in the OECD, 1950–1990," *Journal of Population Economics*, 12(3), 431–449.

Lipset, S. M. (1959): "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *American Political Science Review*, 53(3), 69–105.

Lizzeri, A., and N. Persico (2004): "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform'," *Quarterly Journal of Economics*, 119(2), 707–765.

Lleras-Muney, A. (2005): "The Relationship between Education and Adult Mortality in the United States," *Review of Economic Studies*, 72(1), 189–221.

Lorentzen, P., J. McMillan, and R. Wacziarg (2008): "Death and Development," *Journal of Economic Growth*, 13(2), 81–124.

Luca, D. L., J. Iversen, A. S. Lubet, E. Mitgang, K. H. Onarheim, K. Prettner, and D. E. Bloom (2014): "Benefits and Costs of the Women's Health Targets for the Post-2015 Development Agenda," *Copenhagen Consensus Center Working Paper*.

Lucas, R. E. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1), 3–42.

Lutz, W., A. Goujon, S. KC, and W. Sanderson (2007): "Reconstruction of Populations by Age, Sex, and Level of Educational Attainment for 120 Countries for 1970–2000," *Vienna Yearbook of Population Research*, 5, 193–235.

Lynch, J. W., G. D. Smith, G. A. Kaplan, and J. S. House (2000): "Income Inequality and Mortality: Importance to Health of Individual Income, Psychosocial Environment, or Material Conditions," *BMJ*, 320(7243), 1200–1204.

Maestas, N., K. J. Mullen, and D. Powell (2016): "The Effect of Population Aging on Economic Growth, the Labor Force, and Productivity," *NBER Working Paper*, 22452.

Mahlberg, B., I. Freund, J. Crespo Cuaresma, and A. Prskawetz (2013): "Ageing, Productivity, and Wages in Austria," *Labour Economics*, 22, 5–15.

Mankiw, N. G., D. Romer, and D. N. Weil (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107(2), 407–437.

Marmot, M., and E. Brunner (1991): "Alcohol and Cardiovascular Disease: The Status of the U-Shaped Curve," *BMJ*, 303(6802), 565–568.

Marshall, M. G., K. Jaggers, and T. R. Gurr (2013): *Polity IV project: Political Regime Characteristics and Transitions, 1800–2013.* Center for Systemic Peace, Vienna, VA.

Mason, A. (2007): "Demographic Dividends: The Past, the Present, and the Future," in *Population Change, Labor Markets, and Sustainable Growth: Towards a New Economic Paradigm*, ed. by A. Mason, and M. Yamaguchi, chap. 4, pp. 75–98. Emerald Group Publishing Limited, Bingley.

Mayer, T., and S. Zignago (2011): "Notes on CEPII's Distances Measures: The GeoDist Database," *CEPII Working Papers*, 2011–25.

Mincer, J. (1958): "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 66(4), 281–302.

Mincer, J. (1974): "Schooling and Earnings," in *Schooling, Experience, and Earnings*, ed. by J. Mincer, chap. 2, pp. 41–63. National Bureau of Economic Research, New York, NY.

Molloy, R., C. L. Smith, and A. Wozniak (2011): "Internal Migration in the United States," *Journal of Economic Perspectives*, 25(3), 173–196.

Moral-Benito, E., and C. Bartolucci (2012): "Income and Democracy: Revisiting the Evidence," *Economics Letters*, 117(3), 844–847.

Moretti, E. (2010): "Local Multipliers," *American Economic Review: Papers & Proceedings*, 100(2), 373–377.

Murray, R. P., J. E. Connett, S. L. Tyas, R. Bond, O. Ekuma, C. K. Silversides, and G. E. Barnes (2002): "Alcohol Volume, Drinking Pattern, and Cardiovascular Disease Morbidity and Mortality: Is There a U-Shaped Function?," *American Journal of Epidemiology*, 155(3), 242–248.

Murtin, F., and R. Wacziarg (2014): "The Democratic Transition," *Journal of Economic Growth*, 19(2), 141–181.

National Center for Health Statistics (1963): *Vital Statistics of the United States 1960*, vol. 2A,. U.S. Department of Health, Education, and Welfare, Public Health Service, Washington, DC.

——— (2017a): *Health, United States, 2016: With Chartbook on Long-term Trends in Health.* U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Hyattsville, MD.

——— (2017b): *Vital Statistics of the United States.* U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, Hyattsville, MD, available online at: https://www.cdc.gov/nchs/products/life_tables.htm.

National Heart, Lung, and Blood Institute (2012a): *Morbidity and Mortality: 2012 Chart Book on Cardiovascular, Lung, and Blood Diseases.* National Institutes of Health, Bethesda, MD.

——— (2012b): *Unpublished Tabulations of National Vital Statistics System Mortality Data.* Available online at: http://www.cdc.gov/nchs/nvss/mortality_public_use_data.htm.

Nelson, R. R., and E. S. Phelps (1966): "Investment in Humans, Technological Diffusion, and Economic Growth," *American Economic Review: Papers & Proceedings*, 56(1), 69–75.

Nickell, S. (1981): "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49(6), 1417–1426.

Nordås, R., and C. Davenport (2013): "Fight the Youth: Youth Bulges and State Repression," *American Journal of Political Science*, 57(4), 926–940.

NORTH, D. C. (1990): *Institutions, Institutional Change, and Economic Performance.* Cambridge University Press, Cambridge.

——— (1991): "Institutions," *Journal of Economic Perspectives*, 5(1), 97–112.

OECD (2017): *Preventing Ageing Unequally.* OECD Publishing, Paris.

——— (2018a): *Public Spending on Education (Indicator).* Available online at http://data.oecd.org/eduresource/public-spending-on-education.htm.

——— (2018b): *Public Spending on Labour Markets (Indicator).* Available online at http://data.oecd.org/socialexp/public-spending-on-labour-markets.htm.

OLSHANSKY, S. J., AND A. B. AULT (1986): "The Fourth Stage of the Epidemiologic Transition: The Age of Delayed Degenerative Diseases," *Milbank Quarterly*, 64(3), 355–391.

OLSON, M. (1982): *The Rise and Decline of Nations: Economic Growth, Stagnation, and Social Rigidities.* Yale University Press, New Haven, CT.

OMRAN, A. R. (1971): "The Epidemiologic Transition: A Theory of the Epidemiology of Population Change," *Milbank Quarterly*, 49(4), 509–538.

——— (1998): "The Epidemiologic Transition Theory Revisited Thirty Years Later," *World Health Statistics Quarterly*, 51(2–4), 99–119.

OSTER, E. (2017): "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, forthcoming.

OZA, S., M. J. THUN, S. J. HENLEY, A. D. LOPEZ, AND M. EZZATI (2011): "How Many Deaths Are Attributable to Smoking in the United States? Comparison of Methods for Estimating Smoking-Attributable Mortality when Smoking Prevalence Changes," *Preventive Medicine*, 52(6), 428–433.

PAPAIOANNOU, E., AND G. SIOUROUNIS (2008): "Democratisation and Growth," *Economic Journal*, 118(532), 1520–1551.

PERSSON, T., AND G. TABELLINI (2006): "Democracy and Development: The Devil in the Details," *American Economic Review*, 96(2), 25–45.

PICKLE, L. W., AND R. F. GILLUM (1999): "Geographic Variation in Cardiovascular Disease Mortality in US Blacks and Whites," *Journal of the National Medical Association*, 91(10), 545–556.

PIKETTY, T. (2014): *Capital in the 21st Century.* Harvard University Press, Cambridge, MA.

Preston, S. H. (1975): "The Changing Relation between Mortality and Level of Economic Development," *Population Studies*, 29(2), 231–248.

Prettner, K. (2013): "Population Aging and Endogenous Economic Growth," *Journal of Population Economics*, 26(2), 811–834.

Pritchett, L. H. (1994): "Desired Fertility and the Impact of Population Policies," *Population and Development Review*, 20(1), 1–55.

——— (2001): "Where Has All the Education Gone?," *World Bank Economic Review*, 15(3), 367–391.

Psacharopoulos, G. (1994): "Returns to Investment in Education: A Global Update," *World Development*, 22(9), 1325–1343.

Psacharopoulos, G., and H. A. Patrinos (2004): "Returns to Investment in Education: A Further Update," *Education Economics*, 12(2), 111–134.

Ravn, M. O., and H. Uhlig (2002): "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics*, 84(2), 371–376.

Reynal-Querol, M., and J. G. Montalvo (2005): "Ethnic Polarization, Potential Conflict, and Civil War," *American Economic Review*, 95(3), 796–816.

Rodrik, D., A. Subramanian, and F. Trebbi (2004): "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development," *Journal of Economic Growth*, 9(2), 131–165.

Rogers, A., and J. Watkins (1987): "General Versus Elderly Interstate Migration and Population Redistribution in the United States," *Research on Aging*, 9(4), 483–529.

Romer, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98(5), 71–102.

Roodman, D. (2009): "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.

Rubin, D. B. (1996): "Multiple Imputation After 18+ Years," *Journal of the American Statistical Association*, 91(343), 473–489.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015): *Integrated Public Use Microdata Series: Version 6.0 [Dataset]*. University of Minnesota, Minneapolis, MN.

Sachs, J. D., A. Warner, A. Åslund, and S. Fischer (1995): "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity*, 1995(1), 1–118.

Sachs, J. D., and A. M. Warner (1997): "Sources of Slow Growth in African Economies," *Journal of African Economies*, 6(3), 335–376.

Sala-i-Martin, X. (1997): "I Just Ran Two Million Regressions," *American Economic Review: Papers & Proceedings*, 87(2), 178–183.

Sánchez-Romero, M. (2013): "The Role of Demography on Per Capita Output Growth and Saving Rates," *Journal of Population Economics*, 26(4), 1347–1377.

Sánchez-Romero, M., and A. Prskawetz (2017): "Redistributive Effects of the US Pension System among Individuals with Different Life Expectancy," *Journal of the Economics of Ageing*, 10, 51–74.

Sato, K. (1967): "A Two-Level Constant-Elasticity-of-Substitution Production Function," *Review of Economic Studies*, 34(2), 201–218.

Schultz, T. P. (2002): "Wage Gains Associated with Height as a Form of Health Human Capital," *American Economic Review*, 92(2), 349–353.

Shastry, G. K., and D. N. Weil (2003): "How Much of Cross-Country Income Variation is Explained by Health?," *Journal of the European Economic Association*, 1(2–3), 387–396.

Smith, A. (1776): *An Inquiry into the Nature and Causes of the Wealth of Nations*. W. Strahan and T. Cadell, London.

Solow, R. M. (1956): "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70(1), 65–94.

Solt, F. (2009): "Standardizing the World Income Inequality Database," *Social Science Quarterly*, 90(2), 231–242.

——— (2015): "On the Assessment and Use of Cross-National Income Inequality Datasets," *Journal of Economic Inequality*, 13(4), 683–691.

——— (2016a): "The Standardized World Income Inequality Database," *Social Science Quarterly*, 97(5), 1267–1281, SWIID Version 6.1, October 2017.

——— (2016b): "The Standardized World Income Inequality Database," *Social Science Quarterly*, 97(5), 1267–1281.

STIGLITZ, J. E. (1986): “Theories of Wage Rigidities,” in *Keynes’ Economic Legacy: Contemporary Economic Theories*, ed. by J. L. Butkiewicz, K. J. Koford, and J. B. Miller, chap. 4, pp. 153–206. Praeger, New York, NY.

STRAUSS, J., AND D. THOMAS (1998): “Health, Nutrition, and Economic Development,” *Journal of Economic Literature*, 36(2), 766–817.

STRITTMATTER, A., AND U. SUNDE (2013): “Health and Economic Development – Evidence from the Introduction of Public Health Care,” *Journal of Population Economics*, 26(4), 1549–1584.

STRULIK, H., K. PRETTNER, AND A. PRSKAWETZ (2013): “The Past and Future of Knowledge-Based Growth,” *Journal of Economic Growth*, 18(4), 411–437.

STRULIK, H., AND K. WERNER (2016): “50 is the New 30 – Long-Run Trends of Schooling and Retirement Explained by Human Aging,” *Journal of Economic Growth*, 21(2), 165–187.

SUNDE, U., M. CERVELLATI, AND P. FORTUNATO (2008): “Are All Democracies Equally Good? The Role of Interactions between Political Environment and Inequality for Rule of Law,” *Economics Letters*, 99(3), 552–556.

SUNDE, U., AND T. DOHMEN (2016): “Aging and Preferences,” *Journal of the Economics of Ageing*, 7, 64–68.

SUNDE, U., AND T. VISCHER (2015): “Human Capital and Growth: Specification Matters,” *Economica*, 82(326), 368–390.

TAMURA, R. (2006): “Human Capital and Economic Development,” *Journal of Development Economics*, 79(1), 26–72.

TRAINING MAGAZINE (2015): “Training Industry Report,” *Training Magazine*, 52(6), 20–33.

UNESCO (1997): *Statistical Yearbooks 1963–1997*. United Nations Educational, Scientific, and Cultural Organization, Paris.

——— (2017): *Data for Sustained Development Goals*. United Nations Educational, Scientific, and Cultural Organization, Institute for Statistics, Paris.

UNITED NATIONS (2015): *World Population Prospects: The 2015 Revision*. United Nations, Department of Economic and Social Affairs, Population Division, New York, NY.

——— (2017): *World Population Prospects: The 2017 Revision*. United Nations, Department of Economic and Social Affairs, Population Division, New York, NY.

Urdal, H. (2006): "A Clash of Generations? Youth Bulges and Political Violence," *International Studies Quarterly*, 50(3), 607–629.

U.S. Bureau of the Census (1972): *Public Use Samples of the Basic Record from the 1970 Census: Description and Technical Documentation.* U.S. Department of Commerce, Washington, DC.

U.S. Department of Health and Human Services (1998): *Tobacco Use Among U.S. Racial/Ethnic Minority Groups – African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, and Hispanics: A Report of the Surgeon General.* U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, GA.

——— (2000): *Reducing Tobacco Use: A Report of the Surgeon General.* U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, GA.

U.S. Department of Health, Education, and Welfare (1964): *Smoking and Health – Report of the Advisory Committee to the Surgeon General of the Public Health Service.* U.S. Department of Health, Education, and Welfare, Public Health Service, Washington, DC.

Vallin, J., and F. Meslé (2009): "The Segmented Trend Line of Highest Life Expectancies," *Population and Development Review*, 35(1), 159–187.

Vandenberghe, V. (2017): "The Productivity Challenge. What to Expect from Better-Quality Labour and Capital Inputs?," *Applied Economics*, 49(40), 4013–4025.

Vanhanen, T., and K. Lundell (2014): "Measures of Democracy 1810–2012 [Computer File]," in *FSD1289, Version 6.0 (2014–01–31) [Data Collection].* Tampere: Finnish Social Science Data Archive [Distributor].

Voigt, S. (2013): "How (Not) to Measure Institutions," *Journal of Institutional Econmics*, 9(1), 1–26.

Weil, D. N. (2007): "Accounting for the Effect of Health on Economic Growth," *Quarterly Journal of Economics*, 122(3), 1265–1306.

——— (2014): "Health and Economic Growth," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, vol. 2, chap. 3, pp. 623–682. North Holland, Amsterdam.

WEIL, D. N. (2017): "Health Improvement and Income Growth in the Long Run," in *Demographic Change and Long-Run Development*, ed. by M. Cervellati, and U. Sunde, chap. 3, pp. 43–68. MIT Press, Cambridge, MA.

WHO (2008): *The Global Burden of Disease: 2004 Update*. World Health Organization, Geneva.

WILKINSON, R. G., AND K. E. PICKETT (2006): "Income Inequality and Population Health: A Review and Explanation of the Evidence," *Social Science and Medicine*, 62(7), 1768–1784.

WINDMEIJER, F. (2005): "A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators," *Journal of Econometrics*, 126(1), 25–51.

WORLD BANK (1993): *World Development Report 1993: Investing in Health*. World Bank, Washington, DC.

——— (2014): *World Development Indicators 1960–2014*. World Bank, Washington, DC.

——— (2017): *World Development Indicators 1960–2017*. World Bank, Washington, DC.

ZHANG, J., AND J. ZHANG (2005): "The Effect of Life Expectancy on Fertility, Saving, Schooling, and Economic Growth: Theory and Evidence," *Scandinavian Journal of Economics*, 107(1), 45–66.