Inequality in Income and Gender: New Perspectives and Evidence

Dissertation
zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft der
Rechts- und Wirtschaftswissenschaftlichen Fakultät
der Universität Bayreuth

vorgelegt von

Alexander Maximilian Stimpfle

aus

München

Dekan:	Prof. Dr. Martin Leschke
Erstberichterstatter:	Prof. Dr. David Stadelmann
Zweitberichterstatter:	Prof. Dr. Bernhard Herz
Tag der mündlichen Prüfung:	09. Februar 2017

Acknowledgments

During the course of writing this dissertation, I have benefited from the support, suggestions, and encouragements of many people. First and foremost, I would like to express my deepest gratitude to my supervisor Professor David Stadelmann for his outstanding guidance and expertise, ongoing support, and truly collaborative spirit. It has been a privilege to work with him on this dissertation project. I also thank Professor Bernhard Herz for his kind cooperation, and for agreeing to serve as second examiner.

I am indebted to the Economics faculty and research community of the University of Bayreuth. In particular I cordially thank participants of the Graduate Research Seminars of the University for helpful and encouraging discussions on all three topics. I highly appreciate the feedback received from faculty as well as from fellow doctoral students. I also thank all researchers affiliated with the University's Chair of Development Economics for their kind support and encouragement. Further helpful comments from two conferences of the German Economic Association (Verein für Socialpolitik), and from the European Public Choice Society (EPCS) are gratefully acknowledged.

The Konrad Adenauer Foundation provided crucial financial as well as intellectual support as I completed my dissertation, for which I am very thankful. I also owe special gratitude to McKinsey & Company for providing excellent research conditions and facilitating discussions with fellow doctoral colleagues. As an external student, this altogether helped me tremendously to stay on track with this work, and I could not have asked for a better overall study environment.

Finally, I would also like to extend my deepest gratitude to my family and my partner. Their encouragement and backing especially at the beginning has been invaluable.

Bayreuth, October 2016

Abstract

We provide three new empirical perspectives on inequality in the context of economic development, one focusing on income inequality, and the remaining two on gender inequality. First, we jointly analyze the causal effects of geography, trade integration, and institutional quality on different income groups for developing and developed countries from 1983 to 2012. Favorable geographic conditions tend to discriminate strongly between income groups as low incomes benefit from equator distance whereas high incomes decline. Controlling for institutional quality and geography, trade integration has a negative income effect which increases in absolute size and significance for richer percentiles. Institutional quality strongly and positively affects all income groups, however, high income groups tend to profit relatively more than low income groups. Using different instrumental variable strategies, these findings remain robust for different specification tests and they are consistent over time.

In our second perspective, we offer empirical evidence that early female marriage age significantly decreases female education with panel data from 1980 to 2010. Socio-cultural customs serve as an exogenous identification for female age at marriage, and we apply fixed effects and a quasi diff-in-diff specification to address endogeneity issues. We also show that effects of spousal age gaps between men and women significantly affect female relative to male education. Each additional year between husband and wife reduces the female secondary schooling completion rate by 10 percentage points, the time women spend at university by one month, and overall affects female education significantly more negatively than male education. We also document that marriage age and conventional measures of gender discrimination are no substitutes.

Finally, we examine whether immigrants have brought the missing women phenomenon to Germany and Switzerland. Using a range of micro data since 1990, we find no systematic gender selection of foreigners collectively, but a group of Balkan, Chinese and Indian immigrants display comparatively high sex ratios at birth. Employing different estimation methods we consistently calculate around 1,500 missing girls in Germany (2003-2014) and Switzerland (1990-2014) combined from these immigrant groups. With household survey data we attempt to identify reasons for sex selection at birth in Germany, but find no robust association for any socio-economic variable employed. However, the sex of older siblings tends to matter, and again Balkan, Chinese and Indian immigrants increase the boy-birth likelihood whereas immigrants collectively do not.

Contents

List of Tables

III. MARRIAGE AGE AFFECTS EDUCATIONAL GENDER INEQUALITY

Appendix

IV. DOES CENTRAL EUROPE IMPORT THE MISSING WOMEN PHENOMENON?

Appendix

List of Figures

I. Inequality and Economic Development: New Perspectives on Old Questions

I. 1. Introduction

Inequality is among the most widely discussed and researched themes in the world. And yet it is far from being brought under control, let alone subjugated. According to United Nations, even though the number of people living in extreme poverty has dropped by more than half since 1990, more than 75% of the global population now live in societies that have become more unequal since that year, and inequality has in fact never been higher since World War Two (UNDP, 2016). Many fear that these trends hamper, if not reverse, growth and economic development already achieved, and thus put at stake the welfare prospects of millions of people in the world. For both developing countries and advanced economies, managing inequality remains a key challenge in the 21st century.

For those who aim to grasp and tackle this phenomenon, inequality represents not only a complex intellectual task but is also more prone for personal judgment than other fields of research and policy making. Analyses on the level of socio-economic equality easily lead to normative discussions on how egalitarian a society shall be (Rawls, 1999). Inequality hence represents a field of economic research at least as much as it represents an ethical and philosophical question. Ringen (2006, p. 1) defines inequality as referring "sometimes descriptively to any distribution of goods or bads that deviates from an equal distribution, and sometimes normatively to unequal distributions that represent a problem from an egalitarian point of view". No matter which of these two aspects was in focus, mankind has pondered over inequality through millennia, and emphasis has rather been placed on advocating higher equality and more solidarity.

Yet, if humans are so heterogeneous in nature, attitudes, abilities, and many more characteristics, what is it that makes the concept of achieving more equality so attractive and timeless?[1] Sen (1995, p. 2) points out that "the judgment and measurement of inequality is thoroughly dependent on the choice of variable". Indeed, a broad notion exists that a set of fundamental dimensions shall be equal for every individual. Human rights are a nutshell term

[1] We recognize that heterogeneity does not equal inequality, although the two dimensions may coincide in reality. Blau (1967) distinguishes between heterogeneity, which he defines as a horizontal differentiation among groups in terms of a nominal parameter, and inequality, which distributes social groups vertically in terms of a graduated parameter.

capturing this idea, but even those may be interpreted differently. Some posit equal access to and universal right for employment as first human right, others may advocate equal justice under law, and again others may demand that all people be equally happy.

Economic research on inequality is similarly challenging. We find no agreement regarding what reference should be taken as "right" distribution to compare with observed inequalities, in particular with respect to income. Moreover, there is an increasing call for a more comprehensive perspective on living standards, which next to income should also include aspects such as education, health, individual freedom and more (Goodman, Johnson, & Webb, 1997; Sen, 1985b). Measuring all these individual aspects of life against a (somewhat arbitrary) benchmark adds substantial hurdles to the quest for managing inequality. Thus, in our attempt to contribute to inequality research meaningfully, we are confronted with the challenge of analyzing a highly complex, ill-defined, and interwoven array of questions that in addition tend to be emotionally charged. Economists cannot ignore these complex and challenging features of inequality research, but the hope is nonetheless that the insights presented in this dissertation prove to be meaningful by focusing on specific, well-defined aspects of this field.

In this first chapter, we begin by contrasting some of the most important forms of inequality, and we motivate our choice to focus on income inequality and gender inequality. For the latter, we specifically examine educational gender inequality and gender inequality at birth. Hence, in this introduction we also reserve additional space for familiarizing the reader with these two aspects. Furthermore, we revisit the main ideas proposed in the literature for the relationship between inequality and economic development, from classical economics to the particular contribution of development economics and some of the most recent findings. We conclude this introductory chapter with an executive summary of all main results presented in this dissertation.

I. 2. RESEARCH MOTIVATION

Economic development and income inequality are strongly linked, which explains the high relevance of this relationship in economics. In ancient subsistence economies, nobody was rich or poor – there may have been hierarchical orders within a human tribe of hunters and gatherers, but the level of inequality in terms of wealth was negligible (Nolan & Lenski, 2004). Thus, income inequality tended to arise with economic development and growth. Indeed, income inequality refers to the gap between the rich and the poor, or more formally the "unequal distribution of income and financial assets in the population" (Orton & Rowlingson, 2007, p.1). Both income and assets are variables that are fundamentally linked to the advent of organized economic activities (Smith, [1776] 1976). Hence, income inequality is a cornerstone of any inequality discussion in the economic discipline, arguably drawing most of the research attention. We, too, contribute with a new perspective to this field, as we discuss in depth the impact of fundamental development factors on different income groups in the second chapter. As we will outline below, the analysis helps answer the question if institutional quality, trade, and geographic conditions each rather increase or decrease income inequality in a country.

In addition, however, we strive to go in this dissertation beyond the income dimension of inequality by providing additional perspectives. While a large literature concentrates on the income aspect, that approach does not capture all aspects of inequality in economics. Different forms of inequality are often interconnected (Ortiz & Cummins, 2011; Warwick-Booth, 2013), but exclusively considering inequality in monetary levels fails to grasp the phenomenon in larger scope. In other words, the unequal distribution of income, assets, and wealth, does not exhaustively describe the array of inequality variables that impact individual welfare. Additional facets of inequality exist, which may either have an impact via income inequality, or may affect an economy directly through entirely separate channels. As Neckerman (2004, p. 18) notes, "a wide range of social domains" need to be considered in addition to income inequality to reckon the full cost of inequality for an economy.

These other forms of inequality are sometimes subsumed under the umbrella term social inequality, which Naidoo and Wills (2008) define as differences in resources, power and status within and between societies. In our own words, it is as a discrimination of individuals to access certain resources because of (socially defined) markers. Sernau (2013) lists race, class, and gender as the three most salient markers which societies tend to discriminate against, typically when it comes to questions like: What child features do parents wish for? Who is paid more for the job? Whom do we grant further education? These and many more situations are frequently discussed in the public discourse on unequal treatment and discrimination. Yet, such unequal treatment has in addition distinct economic implications, and therefore, chapters three and four are dedicated to the examination of such facets of inequality. We specifically tackle two aspects of gender inequality, namely educational gender inequality and prenatal gender discrimination. We offer new perspectives by exploring causes as well as the actual extent of each of these gender inequality types, and will introduce the respective topics in further detail below.

To underline the motivation for this research agenda, let us revisit what makes the economic discipline, and in particular development economists, worry so much about inequality. We briefly noted above that income inequality, i.e. the difference in how assets and income are distributed, could only develop with coordinated economic activities, which explains the scholarly interest throughout time. In recent decades, globalization effects and ever growing trade integration have reinforced debates on how increasing economic activities and inequality will affect each other in the 21st century (Berman, Bound, & Griliches, 1994; Jaumotte, Lall, & Papageorgiou, 2013; Krugman, 1995). Many consider the unprecedented migration waves as a symbol of unsustainable global inequalities, and then quickly blame unprecedented trade volumes, among others, as one of the root causes.

Current controversies on trade liberalization agreements are also exemplifying. Among others, we witness heated debates about the Transatlantic Trade and Investment Partnership (TTIP), which for some "symbolizes the worst of global capitalism" (Dearden, 2016) and according to Oxfam would intensify inequality in developing countries (Sarmadi, 2015). But trade liberalization also faces strong headwind in developed countries, where more trade integration is often believed to harm middle and low income groups due to job relocation to low-wage countries. Trade policies can be quickly modified by government intervention. Hence the effects for the income distribution in a country from changes in national trade patterns carry an outstanding political relevance. In short, the impact of economic activities on inequality is mobilizing societies, and scientific inequality research may be a valuable ingredient for a more fact-based discourse in an otherwise often heavily normative discussion environment. In the next chapter we will therefore propose a new way of looking at the income distribution effects of trade, among other variables, to strengthen the empirical basis for related public policy discussions.

Economic activities have an impact on inequality, but also the flipside matters (and tends to happen in parallel), i.e. the effects of inequality on economic welfare in a society. For the sake of simplification, we structure the literature dealing with the economic effects of inequality along three broad strands. The first major argument revolves around distortions of a functioning incentive scheme. In an ideal economic system, there should always be a reward, or pay-off, to the marginal input and resources of an individual, be it labor, human capital investments, or other economic activities. However, such an efficient incentive system is distorted if individuals are rewarded not based on resources they put in, but because they are simply born in the "right" family, have the "right" sex or "right" skin color (Blanden, Gregg, & Machin, 2005). As simple example, a gifted girl may wonder at school "why try hard for good grades because even if I excel, university education is inaccessible for people with my background." Indeed, this way of thinking will be highly relevant for our chapters on gender inequality. We point out already here the well-established notion that resources are no longer allocated efficiently if people with "correct" social markers are selected instead of those who would be most apt or qualified. For instance, substantial restrictions of female access to the labor market massively hampers national labor productivity (Cain, 1986). Since an irrational economic pay-off that is simply based on luck of birth (Piketty, 1995) stands opposite to a welfare-enhancing mechanism where skills and labor are optimally allocated, such inequality is harmful to an economy and thus societal welfare. Gender inequality may be interpreted as such an irrationality, and our work in chapters 3 and 4 aims to find out more on why such a suboptimal welfare scenario may exist, and if it could even be rational after all.

However, also proponents of pronounced egalitarian economic thinking face criticism. The conventional approach in economics suggests that *some* level of inequality is again positive for a functioning incentive system (Aghion, Caroli, & Garcia-Penalos, 1999). As Mirrlees (1971) describes, rewarding individual agents with a constant wage independent of their performance would discourage additional efforts, or even the full exploitation of existing potential resources. This argument is often linked to the general economic underperformance of nations with radical left-wing or socialist economic policies.[2] One contribution by Galor and Moav (2004) adds to the question on the "right" level of inequality. The authors develop an intertemporal theory on the effect of inequality on growth, in which the replacement of physical capital accumulation by human capital accumulation as the main source of economic growth changes the qualitative effect of inequality. As a nation advances economically, the advantages of unequal physical capital accumulation are gradually replaced by the advantages of more equally distributed human capital as the latter is increasingly required. Thus modern economies rather benefit from higher levels of equality, in particular regarding education. This provides a strong motivation for our work on

---

[2] Some of the additional and frequently cited reasons on why inequality is good for growth include the higher marginal propensity of the rich than of the poor to save (Kaldor, 1961), and investment indivisibilities (Aghion et al., 1999), which require, in the absence of well-functioning markets, individual large concentration of capital to realize large-scale investments.

educational inequality in chapter 3, where we provide a new perspective on what could be persisting forces against higher female education despite the negative welfare consequences as described.

A second strand of literature finds evidence that inequality creates unhealthy volatility and has a negative effect on the socio-economic stability (Neckerman & Torche, 2007). Social solidarity, social cohesion and other factors that crucially influence successful interaction of individual economic agents in a society tend to be undermined by increasing inequality (among others, see Acemoglu & Robinson, 2000; Alesina & Perotti, 1996; Bourguignon & Verdier, 2000; Gradstein, 2007). Benabou (1996) suggests that lowering the income of the median voter or middle class relative to the national average, i.e. greater inequality, increases the pressure for redistribution policies which, in turn, discourages investment. Barro (2000) notes how inequality of wealth and income acts as motivation for the poor to engage in crime, riots, and other disruptive activities. Zak and Knack (2001) link trust with inequality and growth by showing how low-trust countries, i.e. countries with less social cohesion and higher inequality, display lower growth rates.

Finally, taking into account the development context of this dissertation, we need to consider the overarching theme of the relationship between inequality and poverty (Todaro & Smith, 2011). Economic development strives towards improvement of living standards in a society and towards eradication of poverty. As poverty is undoubtedly interconnected with – though not the same as – inequality, we need to examine how to best tackle inequality. Apart from less absolute human misery which in itself is desirable, less poverty through inequality reduction also entails macroeconomic aspects that are beneficial for development. Lower credit constraints are among the most important positive consequences, which otherwise prevent the poor from undertaking the efficient amount of investment (Benabou, 1996). In a high inequality setting, individuals with comparatively low incomes cannot borrow money, e.g. to invest in their children's education or to expand their business. Moreover, high inequality tends to depress the savings rate in a country, since the highest rate of marginal savings is typically found in the middle class (Bourguignon, 1981; Gallo, 2002). If that population segment is marginalized between few rich and many poor, insufficient savings are generated that can be used for investments.

This brief summary of the main pathways through which inequality and economic development are linked helps to recognize how various inequalities can each create substantial economic shockwaves. The picture is of course more complex, and we only touched upon the major concepts in order to illustrate the relevance of inequality in the field of economics and to link the fundamental economic ideas to our subsequent work. Our motivation is then to offer three new perspectives on inequality, namely effects of fundamental growth determinants for income distribution, explanations for gender gaps in education, and economic analyses of biased sex ratios at birth.

I. 3. Brief Overview of Economic Thinking on Inequality

In order to situate our subsequent work in the context of the economic literature, we provide a brief and non-technical overview in this section. Over the last centuries many economists have incorporated inequality in their research agenda, but despite the vast amount of thinking and research conducted so far, discussions on the relationship between inequality and economic development are ongoing. As Bigsten (1983) notes, this appears logical since in many theories income distribution usually represents the final outcome element of the economic process, which is already in itself developed and interpreted very differently depending on the strand of literature. In addition, the overview is deliberately broad to be able to point the most important contributions that relate to our research topics. We have reserved separate focused literature reviews in each of the following chapters.

Income Inequality

Income inequality represents our research focus in chapter two. The gap between rich and poor received significant attention in classical economics, beginning with Adam Smith ([1776] 1976) who famously noted that no society can be flourishing and happy, of which the far greater part of the members are poor and miserable. However, he did not provide a distinct theory on what determines a distribution that creates poverty in a society. Malthus (1798) addressed inequality more explicitly, but took a pessimistic view on the growth-inequality nexus. He expected workers' wages to remain at subsistence levels, which would increase inequality as average per capita incomes would in parallel increase with growth. For Ricardo, laws that regulate the distribution of income were of central importance (Gallo, 2002), and Marx ([1867] 2014) developed most of his contributions around exploitation and inequality, where for the first time he envisaged effects in the context of a modern industrial economy. George ([1879] 1960) regarded inequality as a main cause of poverty and argued that more equal societies have greater abilities to cooperate, which would stimulate economic growth.

Clark (1891) in return, one of the fathers of the neo-classical perspective, defended free markets, in which demand and supply determine the rent for capital and land as well as workers' wages, resulting in a Pareto optimum. Consequently, the inequality of a nation was a somewhat objective and fair outcome of individual contributions. Obviously opposing Marxian thinking, the neo-classical or marginal productivity theory regarded all factors of production as creating value and as scarce in supply, so that the share of income received by individuals reflected the general pricing process in an economy. Later in Keynesian economics, the marginal propensity to save is seen as responsible for different income levels (Kurz, 1994). All of those theories have in common a largely functional view on distribution, i.e. they are concerned with the determinants of income from three factors (labor, land and capital).

A seminal contribution to inequality research has been made by Kuznets' (1955) "inverted-U" hypothesis. While according to the author himself it consists of "perhaps 5 percent empirical information and 95 percent speculation" (ibid, p. 26), it has formed a cornerstone of the literature. Kuznets argues that the concentration of savings in the higher income groups and the industrial structure of an economy determine its level of income inequality. As a consequence, in early phases of economic growth, i.e. during the development towards industrialization, the inequality gap widens. Then, on the further path to an advanced industrialized economy, inequality first flattens out before decreasing again – creating altogether the pattern of an inverted U.

While a substantial literature further examined Kuznets' proposed pattern in the aftermath (among others, Ahluwalia, Carter, & Chenery, 1979; Papanek & Kyn, 1986; Paukert, 1973), inequality research overall rather moved away from the frontline of economic interest throughout the rest of the 20th century. But in recent decades the discipline experienced a strong revival, famously dubbed by Atkinson (1997) as "bringing income distribution in from the cold". We believe two developments significantly contributed to this renewed interest. On the one hand, the reignited large public awareness of global inequalities, together with concerns regarding the effects of international economic integration, have called for new scientific contributions. For example, some theories relate inequality to the relative use of old versus more advanced technologies in different sectors, such that the relative ability of individuals to shift between sectors determine their income levels (Galor & Tsiddon, 1997; Helpman, 1997).

Secondly, only in the last decades has inequality research also been able to extensively incorporate empirical work based on improved data availability. For instance the Lorenz Curve, depicting the cumulative share of total income against the cumulative proportion of the population owning that income, and the widely used Gini coefficient which builds on the Lorenz curve distribution, are by now commonly used economic indicators. Employing these indicators with enhanced data from many world regions has significantly enriched the picture on different aspects of inequality. In this general spirit, this contribution has a heavy empirical focus as well, drawing all main findings and conclusions from econometric analyses. Given the various ongoing controversies of our discipline, we believe it is essential to have empirical results accompany the conceptual considerations.

Indeed, using empirical findings economists concluded in the 1980s that, contrary to Kuznets' theory, inequality in the most advanced economies had begun to rise again. Harrison and Bluestone (1988) call this the "great U-turn", reversing previous trends of falling inequality in industrialized countries. Different arguments have been proposed to explain this phenomenon, such as increased female labor force participation (Thurow, 1987) or globalization effects (Alderson & Nielsen, 2002; Goldberg & Pavcnik, 2007). Also Kuznets' overall hypothesis was again repeatedly put under scrutiny using new data sets, yielding mixed results regarding the real existence of the "inverted-U" (see for example Barro, 2000, 2008; Deininger & Squire, 1998; Li, Squire, & Zou, 1998).

Additional quantitative work provided new findings of inequality effects on growth, moving from anecdotal evidence of single countries to more systematic cross-country regressions. Among many studies, Alesina and Rodrik (1994) and Easterly (2007) find a negative effect of inequality for growth, but Barro (2000) finds no robust evidence for such a relationship. Better data availability also allowed for an updated view on the extent of inequality in the world. Bourguignon and Morrisson (2002) report global Gini coefficients over two centuries which display a steadily rising trend. Findings by Milanovic (2009; 2011) also suggest that the worldwide income inequality has been constantly increasing between 1820 and 2002. Quah (2002, p. 19), however, refutes what he calls "anti-globalization" claims based on a positive trend he observes over recent decades. He specifically finds that improvements in living standards due to aggregate economic growth outweigh negative counter-effects from national inequality divergence. Indeed, opinions on the state of global inequality in the world are far from a consensus, with Pritchett (1997) finding a "big divergence" of incomes, whereas Sala-I-Martin (2006) concludes with a "convergence, period".

Recently, literature obtained fundamental new insights from analyzing top incomes (Atkinson, Piketty, & Saez, 2011; Piketty, 2003). The widely discussed book "Capital in the 21st century" (Piketty, 2014) not only represents a further great contribution to the field, it may also be the most visible proof of the dominant role of inequality research in current economics. In his work, Piketty emphasizes the distinction between capital and labor income, explaining why the rise in inequality has taken place due to the role of capital and its effective functioning in capitalist societies. Capital generally accumulates at a faster rate than economic growth occurs. Hence, Piketty argues that capital (and not labor income) tends to become more and more of an influential, even a dominating factor in wealth accumulation and concentration at the top. He proposes that capitalism on its own tends to produce a relatively high degree of inequality. While this could be witnessed in the 19th century and the decades preceding 1914, the major events of the first half of the twentieth century (including the two World Wars and the Great Depression) created economic shocks that destroyed large stocks of capital and thereby to some extent re-balanced the unequal distribution of wealth.

However, as the underlying dynamics remained unchanged (i.e. capital returns surpass economic growth rates), the natural tendency of capital to accumulate resumed after World War Two, and even accelerated with privatization and tax breaks in the Western world in recent decades. According to the author, the majority of wealth nowadays comes via inheritance, with no meritocratic justification for the current wealth owner. Piketty considers this as a long-term threat for societal stability and the acceptance of democratic and capitalist regimes. Despite the controversies around his theories and suggested solutions, the strand of literature he represents remains among the most widely discussed at the moment.

The brief literature review of income inequality research throughout time has shown the many facets of economics affected by this topic. As income inequality assumes such a large role both in inequality research and in economics overall, we aim to contribute to this literature as well in our dissertation. Specifically, chapter two examines the impact of fundamental development factors on different income groups. It adds an international empirical perspective by analyzing how determinants that are widely associated with economic development affect income inequality.

Gender Inequality

Chapters three and four are set out to improve the understanding of socially induced inequality effects on economic development, i.e. we analyze the differential economic impact of individual heterogeneities. Both sections focus on arguably the most extensive form of inequality, namely inequality in gender, thus this paragraph provides some more background introduction. Half of the world's population is female, so that by sheer size any form of discrimination against women has enormous distorting and negative effects on global economic development. McKinsey & Company (2015), a global consultancy, attempts to quantify the economic effects of gender inequality. It estimates a potential of up to 2.2 percentage points of incremental global annual growth if gender inequality could be fully eliminated. According to the study, women currently generate about 37 percent of global GDP, considerably less than what their share of the working population suggests is possible. In India and the Middle East the share even drops to less than 20 percent. The global economy simply cannot operate at its full potential with constraints holding back such a large proportion of the world's population.

While this may sound trivial, it is perhaps surprising how briefly ago economists have adopted this notion and addressed the issue. Early economic research had simply considered women – just like children – as dependents of the male wage earners, which also reflected the traditional societal gender norms (Elson & Cagatay, 2000). Only when the role of women dramatically emancipated in the second half of the 20th century, literature also modified its perspective regarding gender roles and their economic potential. The link between economic development and gender received first major attention through the work of Boserup (1970). She did not merely depict females as mothers and responsible for the household, but emphasized the productive capabilities of women and the associated opportunities for economic development. Moreover, her hypothesis of a marginalization of women in the process of economic modernization triggered a voluminous research interest. A new discipline of feminist economics developed, which placed gender relations at the core of welfare and development research (see for example Elson, 1991; Elson, 1999; Benería, 1995; Cagatay & Ertürk, 2004).

Furthermore, employing development economics has become essential for analyzing this topic, as this field is most apt to frame inequality discussions in a broader yet still in a distinct economic context. However, only in recent years the necessity for such an approach has been increasingly recognized.[3] As Atkinson (2011, p. 159) notes, the economic discipline "needs to take account of the alternatives to utilitarianism that have been advanced in the past half century, such as […] the concept of capabilities introduced by Sen". Sen's concept of income and wealth not representing ends in themselves but serving other individual purposes reaches back to Aristotle (Todaro & Smith, 2011).

Sen (1999) postulates that development should pay close attention to enhancing the lives individuals lead and the freedoms they enjoy. This cannot be achieved by looking at income levels only. He describes how personal heterogeneities such as gender, age, or disabilities may cause a gap between available income levels and actual capabilities of an individual, i.e. the freedoms he or she has to choose over available resources (Sen, 1985a). In short, only an increase in capabilities creates real development for a nation. In case these capabilities are restricted, for instance because of gender discrimination in many parts of the world, we need to obtain a better understanding regarding the extent of such a phenomenon and its underlying causes in order to mitigate the negative development effects. This is precisely what we aim to do in chapters 3 and 4.

In order to create a link back to "more traditional" economics, Cagatay, Elson, and Grown (1995) develop three propositions on the specifics of gender-aware economic research, which conventional approaches tend to overlook: First, economic institutions tend to bear and transmit gender biases. Second, the true cost and productivity of the labor force remains inaccurate as long as unpaid domestic, or "reproductive" labor is not included in economic analyses. This form of labor is typically performed by women, so that errors in economic research tend to be linked to unnoticed gender biases. Third, and most relevant for our research motivation, the authors reaffirm that gender relations have strong macroeconomic implications, i.e. they matter for the distribution of employment and income, and for overall development.

In parallel, demands for policy adjustments were raised, for instance by Benería and Sen (1981), who call for strategies to tackle the negative effects for women from ill-health and overwork. Such propositions were taken up in the following decades by academia as well as international organizations such as the U.N., World Bank, and IMF (Benería, Berik, & Floro, 2015). Finally, Galor (2009) associates a decline in gender inequality not only with higher

[3] Among other, Naqvi (1996, p. 982) provides a strong plea for the raison d'être of development economics in stating that "the many contrary assertions […] do not diminish the subject's (exclusive) claim to understanding the economic (and social) reality in the developing countries better than can be done with the help of, say, the neoclassical economics".

economic development, but also with a more egalitarian distribution of income. He finds that lower female discrimination contributed in particular to the onset of the demographic transition, i.e. lower fertility rates, as well as to a rise in female labor force participation. These factors prove to be essential for reducing women's economic marginalization and for narrowing the gender wage gap. Tzannatos (1999) concludes with similar results. Overall, the majority of studies suggests the importance of considering both the direct and indirect pathways through which gender inequality impacts growth and development (Kabeer & Natali, 2013).

In chapters three and four of this dissertation, we look at gender inequality from an economic perspective, but we also incorporate sociocultural aspects. We consider such a more holistic approach to be the most suitable for our purposes, being most apt to understand what causes gender inequality in societies. As Todaro and Smith (2011, p. 8) note, "development economics, to a greater extent than traditional neoclassical economics or even political economy, must be concerned with the economic, cultural, and political requirements for effecting rapid […] transformations […] that will most efficiently bring the fruits of economic progress to the broadest segments of their populations." In the following, we outline the background and research motivation for our two specific dimensions of gender inequality, namely educational gender inequality and gender inequality at birth.

Educational Gender Inequality

It is well established in the literature that educational gender inequality and economic development are linked, i.e. this form of inequality has direct economic effects. Education plays a key role in development through its ability to absorb modern technology in a country and to develop the capacity for self-sustaining growth (Todaro & Smith, 2011). Moreover, education is an objective of development in itself, enabling social inclusion and enhancing individual freedoms (Sen, 1999). Thus education is a pivotal determinant for economic development, and consequently differences in the level of education among individuals, for instance an educational gap between males and females, have a strong direct impact. Education is essential for building human capital, i.e. for instilling productive investments in individuals. Mincer's (1958) seminal article on the relationship between endowments of human capital and personal income distribution laid a basis on how inequality arises because of individual differences in human capital. Furthermore, investment in human capital has been increasingly recognized as important ingredient to economic development (Ben-Porath, 1967; Heckman, Lochner, & Todd, 2006; Schultz, 1961; Sweetland, 1996).

Human capital enters also as distinct variable in the new growth theories, also referred to as "augmented Solow models" or "endogenous growth theories" (Ray, 1998).[4] In these conceptualizations, technological change is not exogenously given but endogenized (Lucas, 1988; Romer, 1986, 1994). Equally important, returns to capital are not diminishing, but the capital/output ratio is still assumed to remain constant. This is possible since the new growth theories explicitly recognize human capital to be as important as physical capital. The amount of individual human capital, often proxied empirically by levels of education, is regarded as the relative availability of skilled labor. Put differently, nations with higher human capital are more skilled in production, they are able to operate sophisticated machinery and to improve methods of production, and they display overall higher innovation. An important empirical contribution by Mankiw, Romer, and Weil (1992) lends further support to endogenous growth theories. The authors examine human capital (measured by educational attainment), income, and growth in a cross-country setting, and find that when human capital is accounted for as well as physical capital, the Solow model is well suited for explaining incomes and growth across countries.

Work by Glaeser, LaPorta, Lopez-de-Silanes, and Shleifer (2004) confirms that human capital represents a fundamental source for growth, building on earlier ideas by Lipset (1960). We take up their considerations in a dedicated section in chapter 2 (p. 55) where we will examine in a refinement analysis how human capital impacts different income groups. Overall, theory and empirics that incorporate human capital as determinant for economic development have received strong support in the literature. In other words, a lack of education, restricted access to schooling, or forms of educational inequality have a direct negative effect on growth prospects. As gender gaps in education affect about half of the population, i.e. half of the potential labor force in a country, the negative impact from this form of gender inequality is particularly salient. Even more so, since the economic value of education has risen dramatically within the last decades (Neckerman, 2004).

[4] These theories build on Solow's (1957) propositions, especially the idea that technological progress causes long-run economic growth. However, the nature of this variable, also referred to as Solow residual, is reexamined. Neoclassical theory simply treats technological progress as exogenous or completely independent process, even though around 50 percent of historical growth in industrialized nations are associated with that variable (Blanchard & Fischer, 1989).

A stream of literature has examined economic effects of educational gender inequality (see chapter 3 for a more detailed literature survey). Among others, Knowles, Lorgelly, and Owen (2002) augment the Solow model in the spirit of Mankiw et al. (1992) to incorporate female and male human capital separately. In doing so, they estimate the effect of these types of human capital and of the gender gap on incomes between 1960 and 1990. Their findings indicate a negative correlation between the size of the gap and income, i.e. lower levels of female education are associated with lower incomes, also when controlling for male educational attainment. One report estimates the economic cost to 65 low and middle income countries from educational gender gaps at USD 92 billion each year (Plan International, 2008). Further studies have been, among others, conducted by Dollar and Gatti (1999) Barro and Lee (1994), Klasen (2002), and Klasen and Lamanna (2009).

In addition to the significant economic impact of educational inequality, a literature review by Kabeer and Natali (2013) also concludes that findings on gender inequality in education are more reliable than other measures of gender inequality, which is good news for our research. The authors argue that the positive impact of gender-equal levels of education on economic growth appears to be robust in a variety of econometric specifications, data, time periods and country groupings. Yet, they also report that studies vary with regard to what level of education matters empirically for economic development. Some papers consider only primary levels, while others scrutinize secondary schooling. In our work, we will consider different stages of education, namely measures of primary, secondary, and tertiary education in order to address these concerns.

Our research contribution focuses on what actually causes educational gender inequality, a question that received substantially less attention in the literature. In fact, Bandiera and Natraj (2013), among others, call for a better understanding of the root causes of gender inequality. To our knowledge, economic contributions on the reasons of educational gaps between men and women are particularly scarce. In chapter 3 we seek to provide evidence that one root cause that has been neglected so far is marriage age. Marriage is a universal institution, which may also be analyzed from an economic perspective. In fact, the timing of marriage, i.e. the individual marriage age, can be related to a person's level of education through economic rationale. As a consequence, we will argue that the female marriage age, and the age gap between husband and wife impacts the level of educational gender inequality.

Gender Inequality at Birth

Gender inequality at birth broadly refers to an unnatural ratio of male to female births, due to a preference of parents to have a son rather than a daughter. The phenomenon can be observed in different regions of the world, and despite limited data availability, can also be traced back in history (UNFPA, 2012a). With the rise of modern medical examinations that allow gender detection before birth, it became much easier to implement the bias towards boys, which led to cases of enormous distortions in the gender ratios at birth. Specifically, sex-selective abortion is a rather recent yet alarming trend, in which parents decide to abort an embryo in case it is found to be female. Sen (1990) was the first in giving an estimate of the number of all women that are globally "missing" due to sex-selective abortion as well as excessive female death rates at later stages in life. He calculated around 100 million women who are not alive solely due to parental son preference. Later estimates concluded with a similar magnitude (Coale, 1991; Klasen, 1994, 2002; UNFPA, 2012b). Apart from ethical concerns, these imbalances also lead to socio-economic distortions, affecting the marriage market and labor market outcomes (Angrist, 2002).

The World Bank (2011) finds that out of the total number of missing women, about two-fifths are never born, one-fifth goes missing in infancy and childhood, and the remaining two-fifths do so between the ages of 15 and 59. Hence, gender inequality at birth, i.e. the children who are never born, comprise a substantial amount of the total excess female mortality, and we focus our research on this aspect, also referred to as prenatal gender inequality. Specifically, we examine whether immigrants from countries, which are known for prenatal gender inequality, "import" this phenomenon to their new environment. We analyze immigrant groups in Germany and Switzerland as two case studies. The question under examination becomes ever more relevant due to an increase in global migration, as well as the spread of technologies that allow parents to know a child's sex before birth and consequently implement sex-selective practices relatively easily.

While documenting the extent of prenatal gender inequality in a population and among different sub-groups is a worthwhile end in itself, understanding the rationale behind sex preference remains the real challenge to understand and potentially mitigate the dynamics. The specific contribution of economics lies in examining the incentives that may lead parents to exhibit son preference as *rational* choice. Qian (2008), among others, provides compelling evidence that economic conditions may have a causal effect on the extent of son preference. In other words, the low demand for girls has been interpreted as a rational response to economic constraints. Often these constraints cannot be considered in isolation, but require a broader analysis of the socioeconomic and cultural factors that in turn affect household economics (UNFPA, 2012b, p. 31). Eswaran (2014) provides a comprehensive overview of the economics of gender bias in the sex composition of newborns.

The most frequently advanced economic arguments on why parents may prefer a son over a daughter focus on the cost and benefit of raising boys and girls. In many of the migrants' countries of origin, a boy provides future economic benefits, such as support of parents in their old age and possible receipt of a dowry upon marriage, and often continues to work in his parent's household. A girl, in contrast, may represent a significant financial burden because of a dowry upon marriage (Das Gupta, 2000). Moreover, the daughter will move to the husband's family, becoming responsible for the welfare of her husband's parents rather than her own (Bhasin, 1993; Ebenstein, 2014; Geeta, 2007; Sun, 2002). In such an economic setting, sons represent new (subordinate) members to a household while daughters, who ultimately join their husband's family, are lost "investments" to their parents. This creates economic incentives for household heads who decide on resources to minimize the costs associated with children while maximizing their individual (parental) self-interest. Both may indeed be attained more optimally by engaging in sex selection before birth, i.e. by having sons instead of daughters. While this economic rationale may seem reasonable, the situation for migrants is more complex. When moving to a high income country such as Germany or Switzerland, these groups typically do not experience any longer the dynamics of their countries of origin, such as dowry payments or reliance on male children for old-age support. Nonetheless, economic variables have been identified that may affect the extent of prenatal gender inequality among migrant groups (Abrevaya, 2009; Almond, Edlund, & Milligan, 2013; a detailed overview is given at the end of section 2 in chapter 4). Combining these with additional key socio-cultural variables, we will in summary examine the differential effects of family income, education levels, health, and religiosity on the extent of sex selection practices.

While for countries with distorted sex ratios at birth an economic rationale could already be documented, we hence contribute by examining whether there might also be economic reasons among migrants for engaging in prenatal sex selection. As migrants are exposed to a new and de facto exogenously given societal environment, we can also differentiate better between economic reasons for sex selection, and other socio-cultural reasons. In other words, our empirical strategy leads to better identification of the variables examined, which are otherwise rather difficult to disentangle. We also document the overall extent of sex selection in Germany and Switzerland, leading to an estimate of the number of "missing women" at birth in the two countries.

Further Forms of Inequality

We emphasized that a significant part of our contribution lies in an inequality analysis that goes beyond the narrow perspective of inequality in income terms. Apart from gender inequality, economic contributions in this field also tend to focus on inequalities based on race and ethnicity (McCrate, 1999 provides a very good overview of the literature). Following Becker's (1971) work on "The Economics of Discrimination" with a mostly neo-classical approach to race discrimination (Arrow, 1973), a shift towards investigating the phenomenon more broadly occurred. Reflections by Arrow (1998) recognize that "market-based explanations will tend to predict that racial discrimination will be eliminated. Since they are not, we must seek elsewhere for non-market factors influencing economic behavior". Loury (1977) emphasizes the social relations between racial groups and the social setting in which economic activities take place, both of which traditional economics tends to overlook.

In more recent empirical work, inequalities and tensions between ethnicities have been found to affect growth rates, implying that increasing societal fractionalization has negative effects for development (Alesina, Devleeschauwer, Easterly, Kurlat, & Wacziarg, 2003). In particular for sub-Saharan Africa there is some evidence that many of the factors associated with poor economic development, such as low education, political instability, underdeveloped financial systems, and lack of infrastructure, can be empirically associated with ethnic fractionalization (Easterly & Levine, 1997). For Latin America, indigenous populations are performing significantly worse than other social groups on almost every measure of economic and social progress. In return, some of the most successful development stories in Asia come from societies which are highly homogeneous (Todaro & Smith, 2011). We adopt the notion that societal fractionalization matters for inequality in chapter 2, when we perform a robustness check on our main results to test if findings also hold when controlling for ethno-linguistic fractionalization.

Further social markers that establish social inequality but are not based on physical traits such as race or ethnicity have been increasingly investigated as well. Among others, effects on economic development from inequalizing religious beliefs have been widely studied, in particular the Indian caste system (Banerjee & Knight, 1985; Deshpande, 2011; van de Walle &

Gunewardena, 2001).[5] Jeffrey (2002), and Lanjouw and Stern (1991) report that higher poverty among the most discriminated casts is not only due to a lack of productive assets, but also stems from low educational standards and a highly restricted job market access. Borooah, Diwakar, Mishra, Naik, and Sabharwal (2014) examine nearly 20,000 Indian households and attribute their position on the income distribution ladder as well as their poverty risk to a large extent to caste affiliation. Similarly, Kijima (2006) finds that there are ongoing high levels of inequality between castes, i.e. differences in human and physical capital, different returns to education and unequal opportunities for obtaining well-paid jobs. Religion plays also a major role in all of our empirical analyses, as we regularly test effects from religious variables on inequalities in both income and gender. This is important to avoid omitted variable bias and to ensure that the new perspectives we propose are valid.

I. 4. Non-Technical Summary of Main Dissertation Findings

As already noted at the beginning, this dissertation contains two distinct research foci, namely *income inequality* and *gender inequality*. The next chapter 2 covers *income inequality,* and explores how the fundamental development factors geographic conditions, trade integration and institutional quality affect different income groups. While related literature has come up with different fundamental development factors and intensely examined average growth and inequality, the effect of fundamental factors on different income groups has received relatively little attention. We add to this research a more refined perspective by analyzing income deciles, so that observed changes in overall inequality can be traced down further. A further contribution is the focus on variables that are widely recognized as most important determinants for economic development, i.e. we provide a more systematic picture of the impact of key growth regressors. Estimates are repeatedly conducted for several time periods and in various specifications to identify potential outliers and ensure robustness.

We base our findings on a newly constructed dataset of income deciles for 138 countries over 30 years, which incorporates income distribution data from the latest World Income Inequality Database, and using the established instruments for trade integration and institutional

[5] Some scholars question the negative role of the caste system at least in historical context. Bardhan (1996) describes how caste-based mercantile associations and courts provided credible mechanisms of coordination and enforcement which facilitated trade and economic activities.

quality for causal interpretation, we find the following: geographic conditions seem to discriminate between income groups, which is evidenced in a consistent pattern of decreasing coefficients as we move from low to high income groups. The influence of favorable geographic conditions turns even negative once we passed the mean income group, indicating that the poor are affected most by equator proximity. However, geography also loses significance once institutions enter the equation simultaneously, hence the latter seem to be a "deeper" cause than geography.

Trade integration has a negative but frequently insignificant effect on all income groups, and we tend to find that negative effects as well as the significant levels increase for higher income groups. Hence, trade seems to have an equalizing effect across income groups. Institutional quality is associated with systematic and large income gains for all groups at high statistical significance levels, but high income groups seem to benefit even more than the poor. Overall, results are consistent over time, and we observe that the model is relatively better in explaining lower incomes.

Chapters 3 and 4 then each cover a distinct aspect of *gender inequality*. Chapter 3 addresses educational gender inequality, and discusses the causal effect of marriage age. It thereby adds to the literature which proposes that cultural customs and traditions may explain gender gaps in education, but which neglected international evidence on effects of age of marriage on educational achievements. We argue that the marriage age of women and the phenomenon of wives being on average younger than husbands impacts educational investments. This is because the earlier a woman gets married, the shorter her anticipated pay-off to educational investments such that educational investments are lower for younger marriage ages than for older marriage ages. We then provide empirical evidence that early female marriage age significantly decreases female education with cross-country panel data from 1980 to 2010. Socio-cultural customs serve as an exogenous identification for female age at marriage, i.e. our empirical contribution explicitly addresses the theory that causal links run from marriage age due to societal conventions to female education outcomes. Each year of marriage postponement for women is associated with a 3 percentage points higher female completion rate in secondary schooling, and to about three weeks, or 13 percent longer female tertiary education.

We also confirm that effects of spousal age gaps between men and women, i.e. the relative female marriage age, significantly affect female education relative to male education. We adopt this methodology because in case couples in a certain region habitually marry at comparatively younger ages than international average, one might draw incorrect conclusions from examining absolute age levels only. Each additional year between husband and wife reduces the female secondary schooling completion rate by 10 percentage points, the time women spend at university by one month, and overall affects female education significantly more negatively than male education.

Further evidence rests on a quasi difference-in-difference strategy, which focuses on differences between women and men regarding marriage age and educational achievement, i.e. we specifically examine spousal age gaps and educational gaps. This approach helps to eliminate potential confounding factors that affect the level of educational achievements jointly for women and men as we focus only on the differences between the two sexes. We show that spousal age gaps affect female education significantly more negatively than male education. A number of robustness tests confirm our empirical strategy. In order to ensure our proposed link from female marriage age to female education is valid, we explicitly examine the impact of other "regular" gender discrimination variables. Results document that marriage age and conventional measures of gender discrimination do not act as substitutes, so that female marriage age affects educational achievements of women independently of existing levels of other gender discrimination in society.

The second contribution to gender inequality, chapter 4, explores the extent of gender inequality at birth, i.e. biased sex ratios at birth, among migration groups in Germany and Switzerland. The research contributes to the economics of migration and gender, specifically it adds to the empirical literature of sex selection practices among immigrants for the first time a deeper explorative analysis of Balkan migrant groups that moved to Central Europe. It also adds to the literature quantifying excess female mortality rates, and employs for the first time wider individual-level data to examine which socio-economic variables might affect sex selection practices. We resort to three distinct micro data sources since 1990 to analyze the sex ratios at birth of different population groups, and to potentially identify underlying reasons for continued sex selection practices.

Abortion data and birth registries provide no evidence for systematic gender selection at birth among foreigners collectively. Their average sex ratio at birth is slightly elevated compared to natives, but still within biologically normal ranges. However, immigrants from Balkan countries as well as China and India, which are all countries known for strong son preference, indeed display comparatively high sex ratios at birth. Yet we also note that there are substantial fluctuations, depending on whether we inspect birth records in Switzerland or in Germany. Only in Germany we find that sex ratios between Balkan immigrants and natives (Germans) differ on a one percent significance level, i.e. the probability of a newborn being a son is significantly higher if the parents are from the Balkan region. Employing different estimation methods we consistently calculate around 1,500 missing girls in Germany (2003-2014) and Switzerland (1990-2014) combined from these selected Balkan and Asian immigrant groups.

Further German-specific results indicate that sex ratios at birth do not vary much if the father is German or from the same country as the mother, and also the number of years the parents spent in Germany has no substantial differentiating effect. Swiss birth registries indicate a skewed ratio at statistically significant levels for all higher parity births among Chinese and Indians, whereas other migrants have no elevated sex ratio at higher parities, except spiking ratios observed at fourth parity among Balkan families. Lastly, using household survey data we attempt to identify underlying reasons for sex selection practices in Germany, but find no robust associations for any socio-economic variable employed. We conclude from the findings that the gender of children of households in Germany is primarily determined by nature, i.e. through a random outcome, even for migrant sub-groups However, the sex of older siblings tends to matter, and again Balkan, Chinese and Indian immigrants increase the boy-birth likelihood whereas immigrants collectively do not.

Finally, following the core chapters of this book, chapter 5 offers overarching concluding remarks. The subsequent three chapters are all self-contained and have their own introduction and appendix, thus all can be read independently. Each chapter also exists as a working paper version and may be retrieved online as stated in the bibliography. The working paper for chapter 2 has been presented at the 2015 Annual Conference of the German Economic Association (VfS), and the working paper for Chapter 3 at the 2015 Annual Meeting of the European Public Choice Society (EPCS) as well as the 2016 Annual Conference of the German Economic Association (VfS). Chapters 2 – 4 have all been also presented at the Graduate Research Seminars of the University of Bayreuth.

II. The Impact of Fundamental Development Factors on Different Income Groups: International Evidence*

ABSTRACT

This chapter jointly analyzes the causal effects of geography, trade integration, and institutional quality on different income groups for developing and developed countries from 1983 to 2012. Favorable geographic conditions tend to discriminate strongly between income groups as low incomes benefit whereas high incomes decline. Controlling for institutional quality and geography, trade integration has a negative effect which increases in absolute size and significance for higher income groups. Institutional quality strongly and positively affects all income groups, however, high income groups tend to profit relatively more than low income groups. These findings are robust for different specification tests and they are consistent over time.

II. 1. INTRODUCTION

Geographic conditions, trade integration and institutional quality are frequently advanced as causal factors for economic development and growth (see, e.g., Acemoglu, Johnson, & Robinson, 2001; Diamond, 1997; Dollar & Kraay, 2004; Frankel & Romer, 1999; Rodrik, Subramanian, & Trebbi, 2004; Sachs, 2001). As differences in average income levels between developed and developing countries are enormous, the identification of fundamental drivers for economic development has received central attention in economic debates. At the same time, there are ongoing discussions on a widening of the income gap between the rich and the poor in economies worldwide. After Kuznets' (1955) seminal work, a voluminous literature has emerged which analyzes the link between income inequality and growth (see e.g., Barro, 2000; Easterly, 2007; Milanovic, 2000). Looking at political debates, many fear that the rich may benefit disproportionally from a nation's overall economic advancement. While the received literature has come up with different fundamental development factors and intensely explored average growth

---

* A working paper of this chapter is circulating (Stimpfle & Stadelmann, 2015) and has been presented at the 2015 Annual Conference of the German Economic Association (Verein für Socialpolitik) and at the Graduate Research Seminar of the University of Bayreuth. We cordially thank Sabrina Studer, Hartmut Egger and Benedikt Heid for helpful comments, which have been incorporated in this chapter.

and inequality, the effect of such fundamental factors on different income groups has received relatively little attention. This chapter aims to fill that gap.

We analyze whether exogenous changes in geographic conditions, trade integration and institutional quality favor or disfavor specific income groups relatively more than others. Thereby, we advance the literature which studies the effects of fundamental factors of economic development on average incomes. Instead of analyzing whether we can attribute different average incomes across countries to differences in geographic conditions, trade and institutions, we examine whether and how these variables causally affect low and high income groups within countries.

To analyze this question we take a deliberately detailed perspective that systematically looks at the effects of the fundamental factors established in the literature on different income groups over 30 years. We construct a dataset of income deciles for 138 countries which incorporates income distribution data from the latest World Income Inequality Database. We then apply the established empirical cross-country growth methodology on our dataset and we employ the development factors which are analyzed in the recent literature for our econometric analysis. In particular, we use the established instruments for trade integration and institutional quality to ensure that our results can be causally interpreted.

The findings generally confirm the related literature results for average income levels. However, we find important differential effect of the variables on low versus high income groups. Geographic conditions seem to discriminate between income groups, which is evidenced in a consistent pattern of decreasing coefficients as we move from low to high income groups. The influence of favorable geographic conditions turns even negative once we passed the mean income group, indicating that the poor are affected most by equator proximity. This pattern of results is broadly consistent with views proposed by Sachs (2001). Trade integration has a negative but often insignificant effect on all income groups which is similar to the negative average impact shown by Rodrik et al. (2004). However, we tend to find that negative effects as well as significant levels increase for higher income groups. Hence, trade seems in fact to have an equalizing effect across income groups. Institutional quality is associated with systematic and large income gains for all groups at high statistical significance levels. However, the effect of good institutional quality displays an increasing coefficient so that high income groups seem to benefit more than the poor from institutional improvements. Overall, results are consistent over time, and we observe that the model is relatively better in explaining lower incomes. We test the effect of additional control variables, discuss methodological concerns, and perform a number of validity tests. All robustness tests confirm the central results.

The remainder of this chapter is structured as follows: Section 2 provides a detailed literature review. We present the data and the estimation strategy in Section 3. Empirical estimation results for different income groups are presented in Section 4, and we perform refinement analyses and robustness tests in Section 5. Section 6 offers concluding remarks.

II. 2. LITERATURE REVIEW

Development Factors and Overall Income Inequality Effects

A number of papers in the development literature have looked at the effects of development factors on inequality within a country (for a literature survey in this field see also Benabou, 1996; Heshmati, 2006; Lopez, 2004). The influences of trade, sectorial composition of an economy, and public policies on income inequality are among the most thoroughly studied.

On the effects of trade, Wade (2001) points out that there is somewhat of a theoretical economic dilemma: put simply, the neoclassical theory predicts convergence (equality) while the endogenous theory predicts divergence (inequality).[6] Fischer (2001) provides an alternative theoretical framework on how income distribution changes following trade liberalization. According to his theory, the type of the export good determines the effects of trade on inequality: In land-abundant countries, inequalities increase as a country opens up – in labor-abundant countries, the opposite happens. Yet another model by Feenstra and Hanson (1996) suggests that trade between advanced and developing countries could lead to higher inequality in both nations. The authors incorporate the notion that countries increasingly engage not only in final goods trade, but also in trade with intermediate goods. Goldberg and Pavcnik (2007) review a number of papers in this field since the 1980s and list theoretical considerations (such as trade in intermediate products, international flows of capital, trade-induced skilled biased technological change, short-run factor immobility, and firm heterogeneity) that may explain why experiences of countries did

[6] The neoclassical growth theory predicts that economies will converge in productivity and income levels because of higher capital mobility. However, the endogenous growth theory argues that diminishing returns to capital are offset by increasing returns to technological innovation in the developed countries.

not conform to conventional thinking of globalization effects. Their paper calls for a careful consideration of individual country circumstances when evaluating trade integration effects.

Barro (2000, p. 27) similarly notes how effects of trade on inequality following standard trade theory differs from the real-world observations: “From the perspective of standard trade theory, the effect […] depends on factor endowments […] Greater international openness would raise inequality in rich countries and lower it in poor countries […] The standard theory seems to conflict with the concerns expressed in the ongoing popular debate about globalization. The general notion is that an expansion of international openness—including access to foreign technology and culture—will benefit most the domestic residents who are already well of.” His supporting data analysis suggests that there is indeed a positive relationship between trade and inequality, which is most pronounced in poor countries.

A number of additional empirical studies have also attempted to shed more light on this debate. For a panel dataset over 28 years, Spilimbergo, Londoño, and Székely (1999) argue that trade openness has different effects on the wage distribution depending on the national endowments of production factors. Land and capital can be accumulated endlessly by an individual, whereas factors such as education have a natural upper limit of accumulation. This alone leads to different inequality levels of countries, independent of trade patterns, as each country has a different composition of endowments. The authors then examine effects when including trade openness and find that it reduces inequality in capital-abundant countries, but increases inequality in skill-abundant countries. Results by Alderson and Nielsen (2002) for advanced industrial countries indicate that trade openness is better suited to explain the longitudinal trend of increasing inequality than cross-sectional inequality differences among countries. For a regional study on Africa, Odedokun, and Round (2004) find that regional dummies, overall size of the government, and lack of skilled manpower are significant inequalizing variables. However, their results when testing effects of openness to trade on inequality are non-significant. The authors conclude that “the recent fear that increased globalization would aggravate the existing inequality is not supported by this evidence and neither is the standard prediction of trade theory” (ibid., p. 305).

On the impact of sectorial composition on inequality, Ravallion and Datt (2002) analyze household surveys for 15 major states in India over a 30 year period and observe that higher non-agricultural output reduces income inequality. They also find that rural growth is a stronger factor for reducing poverty than urban growth, and that initial development conditions largely explain differences between states in poverty reduction. Lopez (2005), in contrast, finds for his sample of 14 country case studies that inequality is mainly driven by the people employed in the non-agricultural sector. Hence, when this sector does relatively better than other sectors in an economy, inequality tends to increase. When this sector underperforms (as it did during the 1980s when non-agricultural growth was below agricultural growth), ceteris paribus inequality tends to decline despite overall growth. The low skills of the labor force employed in the agricultural sector, which are insufficient for upward sector mobility, are described as critical barrier in preventing rising inequality during economic upswings. A similar observation has already been made by Viner (1953). However, Acemoglu (2003) does not consider the argument on supply of skilled labor to be sufficient to explain observed inequality trends. Devroye and Freeman (2001) also report that skill inequality explains only a small fraction of the observed cross-country inequality differences.

Easterly (2007) confirms that agricultural endowments predict inequality, building on similar arguments that have been brought forward by Engerman and Sokoloff (1997, 2000). These authors use as instrument the relative abundance of land suitable for wheat to that suitable for sugarcane to establish causality between different agricultural endowments and inequality: "In the Americas, the land endowments of Latin America lent themselves to commodities featuring economies of scale and the use of slave labor (sugar cane [as] premier example) and thus were historically associated with high inequality. In contrast, the endowments of North America lent themselves to commodities grown on family farms and thus promoted the growth of a large middle class" (Easterly, 2007, p. 757). In a related study, Acemoglu, Johnson, and Robinson (2002) link income inequality to population density, labor endowment, and resulting colonization patterns by European powers. They argue that inequality increased in colonial areas that were densely populated, as the European colonizers and supporting local elites could exploit such existing structures more easily.

The realm of public policy as determinant for inequality has received renewed attention. Adelman and Robinson (1989) note that, overall, one observes an inevitable initial deterioration in the distribution of income which reflects the uneven, disequilibrium nature of the first phase of the development process. Second, however, the persistence of this deterioration is a matter of policy choice. Milanovic (2000) provides evidence that inequality can be steered by social choice variables (social transfers and state sector employment), which decrease inequality on average by some 13 Gini points, or about one quarter of "given" inequality. He argues that the preference for social equality is income-elastic so that social choice variables play a more prominent role as the nation gets wealthier. Atkinson, Piketty, and Saez (2011) devote a major part of their attempt to explain income distribution to politics and historical episodes of the political economy. Loss of capital income due to political turmoil and wars, which leads to physical destruction as well as nationalization, and the levelling of earned incomes during warfare heavily impact the distribution of wealth in a society. Results by Checchi and García-Peñalosa (2008) indicate that stronger labor market institutions are correlated with lower inequality, with the exception of the tax wedge that exhibits a positive correlation with the Gini coefficient.

According to a study by Bourguignon and Morrisson (2002) on inequality in the world from 1820-1992, a remarkable finding is that population growth rates do not seem to have a big impact on inequality dynamics. Barro (2000) finds mixed results for effects of education on inequality, but notes that taxation is generally equalizing: the Gini value is lower by roughly 0.05 when taking data of income after tax rather than gross income. Piketty (2001, 2003) highlights the role of taxation for his inequality studies on France, too. Li, Squire, & Zou (1998) draw on the argumentation of Bertola (1993) that the rich may have the resources to lobby for policies which are good for them but rather harmful to the rest of the economy, so that inequality is expected to persist. They test this hypothesis empirically and indeed find support in the data, using a measure of political freedom and initial secondary schooling as proxy variables. Furthermore, capital market imperfections and associated constraints of the poor to access credit are significant inequalizing determinants in their study, while higher union density and relative success of social-democratic programs are correlated with lower inequality.

Development Factors and Income Group-Specific Effects

Being important contributions, most of the cited literature is, however, relatively mute on factors driving incomes of specific income groups in a country, so that changes in overall inequality cannot be traced down further. In contrast, few papers analyze the effect of fundamental factors for economic growth on different income percentiles within a country.[7] Often such research focuses on the bottom income groups (for a literature review of pro-poor growth, see for example

[7] Grossmann and Stadelmann (2013) contribute by examining the wage effects for specific income groups (80th and 90th percentile) migrating from developing countries to advanced economies. This study, however, has a within-country focus and disregards effects from international mobility.

Ravallion & Datt, 2002). White and Anderson (2000) report that growth associated with progressive distributional changes will have a greater impact in raising poor incomes than "general" growth which leaves distribution unchanged. Using data from 143 growth episodes they find that the effects from "general" growth generally dominate, but that redistributional strategies matter since for around every fourth case distribution can be shown to be equally important as growth for explaining income growth of the poor.

Lundberg and Squire (2003) emphasize that changes in trade openness show strongly diverging effects on aggregate equality versus on the poor only: While aggregate inequality effects are small, the welfare of the poor is significantly negatively affected by trade, which suggests greater vulnerability of the bottom income groups. The authors' estimates suggest that more trade is negatively associated with growth among the poorest 40 percent, but strongly and positively with growth among the middle 60 percent and richest 40 percent. Jaumotte, Lall, and Papageorgiou (2013), using a panel of 51 countries from 1981 to 2003, conclude that inequality is more affected by technological progress than globalization (which includes trade). Globalization effects tend to offset one another because trade integration is associated with a reduction in inequality, but foreign direct investment leads to an increase in inequality. A further break-down by income group reveals that trade growth is associated with higher incomes of the bottom four quintiles, but decreasing incomes of the richest quintile. In contrast, foreign direct investment and technological progress benefit mostly the rich.

A study by Weil (2007) considers health as key determinant for reducing income inequality. He finds that eliminating health gaps would reduce the ratio of the 90th income percentile to the 10th percentile by 12.7 percent of its initial value, with most of the impact taking place in the median/10th percentile ratio. While he concludes that health is an important determinant of income variation, he recognizes that "it is also much smaller than existing estimates derived from cross-country regressions" (ibid, p. 1301). Roine, Vlachos, and Waldenström (2009) study economic determinants that are particularly pro-rich. They find that periods of high economic growth, and financial development measured as the relative share of the banking and stock market sectors, benefit the top income bracket disproportionally. In contrast, government spending and openness to trade have no clear effects on the rich, with the latter even tilting towards a negative effect.

A different stream of literature takes the stand that most of the variation in changes of bottom incomes can be attributed to the growth rate of average incomes. Dollar and Kraay (2002) focus on the effects for the bottom 20 percent of the income distribution, applying the regressors openness to international trade, macroeconomic stability, moderate size of government, financial development, and strong property rights and rule of law. They do not find a systematic relationship between any of these variables and the poorest quintile and conclude that the poor benefit equi-proportionately from growth determinants like all other income groups. Bruno, Ravallion, and Squire (1998) report similar findings as they summarize that distributional changes of income are generally uncorrelated with economic growth, i.e. growth has no systematic impact on inequality. Since distribution does not alter, growth will reduce absolute poverty. Quah (2001) also documents that improvements in income levels from general growth prevail over any deterioration due to increases in inequality. Work by Kray (2006) and Dollar, Kleineberg, and Kraay (2014) echoes these findings. The latter find, through a Bayesian Model Averaging, that there is little empirical evidence that any of their 13 growth variables[8] are robustly correlated with the income share of the bottom 40 percent. In conclusion, they underscore the pivotal role of rapid growth in average incomes because thereby the poor benefit most as well.

However, Balakrishnan, Steinberg, and Syed (2013) report deviating findings when applying the same methodology to only Asian and Latin American countries, but instrumenting the dependent income variable.[9] In a rare research specification which analyzes both poor and rich income groups, they find that the bottom quintile participated less than proportional in average income growth while the top quintile participated over-proportionally. The authors also emphasize significant differences across regions. Overall, education, industry employment, and financial inclusion reforms appear as pro-poor and inclusive growth variables. On the other hand, financial openness seems to be negative for the bottom income brackets. Ravallion (2007) takes general issue with this strand of research, since the finding of relative inequality remaining unchanged with growth is perfectly consistent with large increases in absolute income disparities. Hence,

---

[8] These are a measure of financial development (M2 as percentage of GDP), the Sachs-Warner indicator of trade openness, the Chinn-Ito Index of financial openness, the inflation rate, the general government budget balance, life expectancy, population growth, Freedom House measure of civil liberties and political rights, the frequency of revolutions, a dummy variable indicating whether the country was party to a civil or international war in a given year, primary school enrollment rates, a measure of educational inequality, and the share of agriculture in GDP.

[9] Specifically, they use lags of real per capita income as measured in the Penn World Tables (PWT) to instrument the (household-survey-based) average income variable. The authors argue that "the lagged variables help correct for endogeneity bias by identifying the component of income that is predetermined, and the PWT measure of income corrects for measurement error by identifying the component of income as measured by the household survey that is also consistent with this secondary measure of income" (Balakrishnan et al, 2013, p.9).

measuring relative income distribution shares may be a deceiving exercise overall. In addition, he points at considerable measurement errors that are likely in the context of such research.

Dollar, Kleineberg, and Kraay (2016) re-visit and update their research in a dataset spanning 121 countries and four decades. They re-confirm the essential outcome that variation of the incomes of the bottom 20 percent and bottom 40 percent of the population are not significantly different from the variations in average incomes. Specifically, "a standard variance decomposition indicates that 61 percent (77 percent) of the cross-country variation in growth in incomes of the poorest 20 percent (40 percent) is due to growth in average incomes." (ibid., p.69). Moreover, none of their country-level variables that are typically associated with economic growth display significant correlation with income growth of the poor. This pattern is robust over time, leading to the authors' summary of being unable to identify evidence of "pro-poor" growth determinants.

The differential analysis of growth variable effects on poor and rich uses a different angle for examining the growth-inequality nexus. While there has already been valuable work in this field, the literature review identified a set of interesting research gaps. So far, there has not been a detailed global effort to systematically analyze the effect of fundamental growth factors on both the rich and the poor. Most of the empirical work presents itself as rather scattered, with key growth regressors and/or income groups missing, and with explanatory variables employed that make it difficult to identify a common consensus. In addition, there is limited knowledge on whether the role of development factors changes for specific income groups over time. Effects have been mostly estimated for only one point in time, and hence, results are susceptible to time-variant effects.

In this chapter, we offer a detailed perspective on all key income percentiles to determine how growth variables affect different parts of society, from the lowest 20% to the highest 10% of the income distribution. Cross-section estimates are also repeatedly conducted for several time periods to address potential outliers, and in addition we employ a panel dataset. This research design recognizes the need to go beyond a narrow view definition of development, measured through average incomes only. It incorporates the aspect that certain development factors may be considered preferable if they favor the poor, or at least lead to higher incomes throughout all parts of society. Related literature tends to issue policy recommendations which are based on empirical evidence for solely average incomes. However, inclusive growth is only possible if we can confirm that income growth accrues in particular to the lower-income groups in society. Unprecedented migration in search for better economic prospects reflects that many are excluded from an overall growing world economy. Yet, the debate on this topic tends to be occupied by popular speculation rather than scientific insights. We provide new empirical evidence and thereby help design more targeted policy recommendations.

II. 3. DATA AND IDENTIFICATION STRATEGY

Data

We construct income deciles for 138 countries[10] by combining information on average national income per capita reported by the Penn World Tables 8.0 (Feenstra, Inklaar, & Timmer, 2015) with the most recent data on income dispersion from the UNU-WIDER (2014a) database. The literature on cross-country growth regressions warns of pitfalls in "just merging" data from different sources (Atkinson & Piketty, 2007). We follow the argumentation of Dollar and Kraay (2002) who point at the pragmatic advantages of incorporating per capita GDP data for income distribution data, namely better data availability and enhanced comparability with existing literature. Therefore, for the average income level measurement, we apply the real GDP per capita data at current PPPs from the Penn World Tables. Sala-i-Martin (2006) also advocates such an approach of merging national account (Penn World Tables) and survey (UNU-Wider) inequality data. Roine et al. (2009) base their income measurement on personal income tax returns (for a similar methodology see also Piketty & Saez, 2003; Atkinson, Piketty, & Saez, 2011). Individual income tax data would be valuable in order to construct income deciles for this research effort. However, even such data may suffer from tax avoidance and evasion (Atkinson & Piketty, 2007; Davies, Shorrocks, Sandstrom, & Wolff, 2007; Leigh, 2007) and, more importantly, data are not available for a sufficiently large number of countries, in particular less-developed countries.

The UNU-WIDER database on income dispersion by the United Nations University builds on previous work by Deininger and Squire (1996). The revision WIID3b used here contains data for developed, developing, and transition countries. It represents an updated and enhanced level of data availability with the latest observations now reaching the year 2012 (UNU-WIDER 2014b). It also responds to earlier criticism regarding quality and consistency (Atkinson & Brandolini,

[10] As many of these 138 countries have only selected data entries over the entire timespan examined, no time period sample has all 138 countries included simultaneously. The maximum sample size is 117.

2001, 2009), for example by closely following the recommendations of the Canberra Group (2001) on international standards for income data. The break-down of the UNU-WIDER income distribution data is generally limited to the decile level. As the heterogeneity of the top decile has frequently been pointed out (Atkinson et al., 2011; Roine et al., 2009), data on the top one percent or top five percent would have potentially provided additional valuable insight. However, as the focus here is to examine the macro-effect of development factors across various income groups from poor to rich in lieu of an exclusive top income study, we regard the given dataset as sufficient.[11]

A set of five-year timespans will be the subject of analysis. If there is at least one data point available per timespan and income bracket under examination, the respective country is included in the dataset. No data points are constructed if they are not available. In case of several data points per period, we apply a simple average of the years with available data. A detailed overview of the countries which form the respective sample per time period can be found in the appendix (see table 2.34). If there were several sources for the same single year and country available in the UNU-Wider database, we used the one with the most data points across all percentiles to enter the average calculation for the respective time period. Then, for calculating the dependent variable Income D_i (income of a population decile D for country *i*), we follow Dollar and Kraay (2002) and multiply the average national income per capita $yavg_i$ with the given decile share D_i divided by respective decile d:

(2.1) $$\text{Income } D_i = (D_i/d) \cdot yavg_i$$

At the country level, we attempt to explain income levels with three variables. GEO_i, $TRADE_i$, and $INST_i$ are respectively country measures for geography, trade integration, and institutions. This core regression specification is closely aligned with the choice of variables by Rodrik et al. (2004) who employ these three fundamental development factors, which they refer to as the "three strands of thought [that] stand out" (p. 132) for determining whether economic development takes place.[12] The three explanatory variables hence represent development factors which are widely regarded as most fundamental for development (see for example Acemoglu, Johnson, & Robinson, 2001, 2002; Barro, 1991; Diamond, 1997; Frankel & Romer, 1999; Gallup,

---

[11] We nonetheless tackle top incomes separately in the robustness section later in this study.

[12] Rodrik et al. (2004) call them "deeper determinants" as opposed to the term "fundamental development factors" used in this context, but these are both equivalent concepts.

Sachs, & Mellinger, 1999; Hall & Jones, 1999; Sachs, 2001; Sachs & Warner, 1995; Sala-i-Martin, Doppelhofer, & Miller, 2004).

The concrete choice of variables to measure each fundamental factor is based on the acceptance in the literature as well as the level of data availability for our specific set of countries. Institutional quality is measured by World Bank data on "Rule of Law" which reflects perceptions of confidence in rules of the society, including quality of contract enforcement and property rights. This measure can take values from -2.5 (weakest institutions) to +2.5 (strongest institutions) (Kaufmann, Kraay, & Zoido-Lobatón, 2002). We instrument this endogenous variable by Hall and Jones' (1999) fraction of the population speaking English and other European languages.[13] Rodrik et al. (2004) proceed with a very similar methodology by using the Hall and Jones language data as instrument for institutions in order to expand their sample.

Trade integration is measured by the combined share of exports and imports of national GDP, using World Bank data and taking logs.[14] The variable is instrumented by Frankel and Romer's (1999) constructed trade shares, an established method to tackle endogeneity issues (Dollar & Kraay, 2002; Easterly & Levine, 2003; Grossmann & Stadelmann, 2013). The authors compute predicted values of bilateral trade based on geographical features, and allocate these bilateral trade flow coefficients also for country pairs which are not included in their original sample. There has been criticism as to the weakness of their instrument and a call for more explicit geography controls (Noguer & Siscart, 2005). Thus, we separately account for the variable geography in our regression equations. Geography is expressed through "distance from equator" in our base specification, but we will also include alternative geography variables as robustness tests.

Table 2.1 provides descriptive statistics for the key variables per period of examination.[15] Improvements in data availability are reflected in an increasing sample size over time, from 56 countries in the 1985 set (measured as average of 1983-1987) to 117 countries in the 2005 set (measured as 2003-2007). The average real GDP per capita income in our sample has risen by 64 percent (from $5,454 to $8,968) between 1985 and 2010, which corresponds to a compound annual growth rate of 2.0 percent. This falls somewhat short of the actual reported global per capita

[13] These language data were originally used by the authors to construct an aggregate social infrastructure index.
[14] Here we also closely follow Frankel and Romer (1999) who use the current price, local currency trade-GDP ratio reported in the Penn World Table, although there are other methods proposed. Alcalá and Ciccone (2004), for example, provide a theoretical reasoning for PPP-adjusted trade ratio as a measure of trade openness.
[15] Detailed variable sources and descriptions are given in table 2.33 in the appendix

income growth of 2.9 percent p.a. during that period (World Bank, 2014). This is due to the limited data availability in early periods which biases our first samples towards higher income countries. Indeed, the 56 countries' average income level is about twice the average global income for 1985. This higher jump-off point leads to a smaller subsequent growth rate until 2010, where the annual per capita income of our gradually increased sample and the actual global income levels converge at around $9,000 in PPP terms.

A granular view of the sample data reveals that, over the 25 years, the bottom 20 percent of the income distribution have actually grown disproportionally by 2.7 percent per annum, while the wealthiest ten percent saw their incomes increase by only 1.8 percent per annum. Mere income level dynamics hence suggest a converging trend, albeit at a slow pace. In absolute numbers, in 2010 the average global top 10 percent income of $26,140 was still nearly ten times the $2,788 reported for the bottom 20 percent; in 1985, this top-to-bottom ratio had even been close to twelve. For the regressors, we can observe a reduction of institutional quality over the timeframe by 16 percent, whereas the sample's average geographic dispersion remained constant. Trade volumes have seen a sizeable hike over the years, growing at almost two percent annually for our dataset.

Table 2.1: Descriptives

	1985	1990	1995	2000	2005	2010
Sample size	56	71	107	97	117	91
Log Income First Quintile	7.27	7.15	7.08	7.26	7.42	7.93
	(1.32)	(1.56)	(1.52)	(1.52)	(1.5)	(1.42)
Log Income Median	8.28	8.16	8.10	8.25	8.33	8.83
	(1.22)	(1.36)	(1.35)	(1.37)	(1.43)	(1.34)
Log Income Average Population	8.60	8.51	8.46	8.58	8.63	9.10
	(1.08)	(1.21)	(1.21)	(1.27)	(1.35)	(1.25)
Log Income Top Quintile	9.44	9.37	9.35	9.43	9.47	9.89
	(0.98)	(1.10)	(1.12)	(1.20)	(1.28)	(1.18)
Log Income Top Decile	9.72	9.67	9.65	9.71	9.74	10.1
	(0.94)	(1.06)	(1.08)	(1.18)	(1.27)	(1.15)
Geography (GEO_disteq)	0.34	0.30	0.33	0.32	0.31	0.35
	(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)
Log Trade Openness (LN_Trade_WB)	3.92	3.98	4.14	4.22	4.35	4.37
	(0.60)	(0.51)	(0.53)	(0.51)	(0.52)	(0.55)
Institutions (Inst_Rule_of_Law)	0.39	0.19	0.02	0.04	0.04	0.22
	(0.98)	(1.01)	(0.95)	(0.98)	(1.00)	(1.04)

Identification and Estimation Strategy

The first step of the estimation procedure is to analyze a series of regressions in which the logs of income deciles are linked to fundamental development factors. Then we inspect the respective variable coefficients and their variation depending on the income group examined. The basic econometric model is as follows:

(2.2) $\ln[\text{Income } D_i] = \mu + \alpha\, GEO_i + \beta\, TRADE_i + \gamma\, INST_i + \varepsilon_i$

We address the challenge of measuring the variables institutions and trade as exogenous regressors through a two-stage least squares estimation approach. Here, we resort to the well-established instrumental variables (IV) as introduced before (amongst others see Alcala & Ciccone, 2004; Dollar & Kraay, 2003; Rodrik et al., 2004).[16]

In the first-stage, institutions $INST_i$ and trade integration $TRADE_i$ are regressed on all exogenous variables and instruments, which yields:

(2.3) $INST_i = \theta + \sigma\, LANG_i + \pi\, CONSTRA_i + \omega\, GEO_i + \varepsilon_{INST_i}$,

(2.4) $TRADE_i = \lambda + \phi\, LANG_i + \xi\, CONSTRA_i + \nu\, GEO_i + \varepsilon_{TRADE_i}$,

where $LANG_i$ refers to language data of Hall and Jones (1999), and $CONSTRA_i$ to constructed trade shares by Frankel and Romer (1999). Consequently, the estimated slope coefficients capture the partial correlations between the set of regressors and the different income groups and can be causally interpreted. We analyze the average GDP per capita, the bottom 20 percent, the median, the top 20 percent, and the top ten percent of the income distribution per country. This allows an understanding of dynamics along the entire distribution. For example, let us assume that the variable trade integration has a positive coefficient for the bottom 20 percent, but a negative coefficient for the top ten percent. The coefficient sign of the average GDP per capita would then help us understand the overall "mean" effect of the trade variable for a country.

We run the regression on a set of different 5-year timespan averages, i.e. in order to put the findings on a broader basis we aim to identify robust patterns independent of a potential economic cycle. This represents a step beyond single cross-section analyses conducted in the past, as there is only selected literature using panel data in this field of research (Irwin & Terviö, 2002; Dollar et al., 2014). We start with the 5-year average around 1985 (1983-1987) as this represents the

[16] There is a literature which discusses shortcomings of these standard instruments (amongst others, see Eberhardt and Teal, 2011; Deaton, 2010; Bazzi and Clemens, 2013). We will deal with some of the major issues when analyzing the robustness of our results.

earliest sensible data set available. We repeat estimates analogously from 1988-1992, 1993-1997, 1998-2002, 2003-2007, and finally 2008-2012.

The regressor GEO_i as well as the instrumental variables $LANG_i$ and $CONSTRA_i$ remain constant over the various periods. While trade shares on the basis of Frankel and Romer's (1999) methodology could also be constructed for other time periods than their original base year 1985, there are a number of reasons to refrain from doing so. First, the constructed trade shares are calculated using geographical variables which remain generally constant, in particular over our limited timeframe of 30 years. Hence, there will be very little data variation when re-creating trade shares for other years. Second, this rational consideration is underpinned by empirical work from Feyrer (2009), who introduces a dynamic instrument for trade on the basis of Frankel and Romer. His results are a close confirmation of the original instrument, so Frankel and Romer's constructed variable is quite robust over time. Third, Rodrik et al. (2004) also decide to keep trade shares as a constant instrumental variable with the original 1985 values, even when using them to estimate 1995 GDP per capita values. Hence, rationale, empirics, and recognized literature point us towards using fixed values for $CONSTRA_i$. For all except the first period of our data set, the 1985 values are consequently lags which may be even considered preferable from an exogeneity perspective. $INST_i$ and $TRADE_i$, in contrast, will be dynamically adjusted to the respective period.

We examine samples with the respectively largest number of country data available per time period, which results in larger sample sizes as we move towards the present. Unfortunately, a number of countries from the former Eastern bloc are not included. This is due to missing income distribution data as well as territorial re-organizations which affect comparability over time. Nonetheless, the larger countries in that region such as Russia, Poland or Romania can be included in the analysis.

II. 4. Empirical Results

Baseline Results

We start with simple correlations to investigate in how far our input variables and the different income levels move together. Figure 2.1 presents the respective scatter plots for 2005. A first look at the data reveals that geography (distance from equator) correlates with ca. 0.7 with income, and displays higher correlation for the poor vis-à-vis the rich. Trade on the other hand shows a rather weak correlation, which only somewhat increases over the years from roughly 0.2

to 0.3, and without an income-related pattern. Finally, institutional quality is highly correlated to incomes, roughly at 0.8. We also observe a slight but persistent correlation pattern across income levels, with higher coefficients for the poor than for the rich. In summary, income of the poor correlates more strongly with geography and institutions than income of the rich. Trade plays a secondary role among the fundamental factors, and has no distinct correlation dynamics across income groups.

Next, we look at the way the relationships between variables are mirrored in a simple OLS regression of equation (2.2). Results are summarized in table 2.2 across income levels, exemplified again for 2005. First evidence (without taking account of causality issues) generally confirms the literature's findings (in particular Rodrik et al., 2004) with regards to the sign and significance levels of variables. These hold also true for most of the income distribution, with exceptions identified at the top end. Therein, geography tends to lose its significance and the coefficient is on average only one third of its value for the bottom quintile. We also observe a modest decline of the coefficient size for institutions as we go from poor to rich. In general, the coefficient pattern follows a linear trend so that inspection of top and bottom income groups allows to also draw conclusions about median and average. In other words, we find no peculiarities around the middle income groups that require additional interpretation.

Incomes of the poor can be more precisely estimated with the variables at hand than incomes of the rich, as reflected in a decreasing R-square from poor to rich. Results are qualitatively similar also for the other time periods, but point estimates for trade integration are never precise.[17] Altogether, countries more distant from the equator and stronger institutions are likely to have higher incomes, but trade effects are rather inconclusive. Geography generally displays high significance, except for selected time periods when both looking at top income groups and simultaneously controlling for institutions. Institutions in return are always significant at the one percent level and sharply increase the overall regression fit.

[17] Trade displays relatively large standard errors which lead to non-significant coefficients. The coefficient sign is consistently positive only from the year 2000 onwards. An overview of all OLS results across time periods is given in the appendix (see table 2.32).

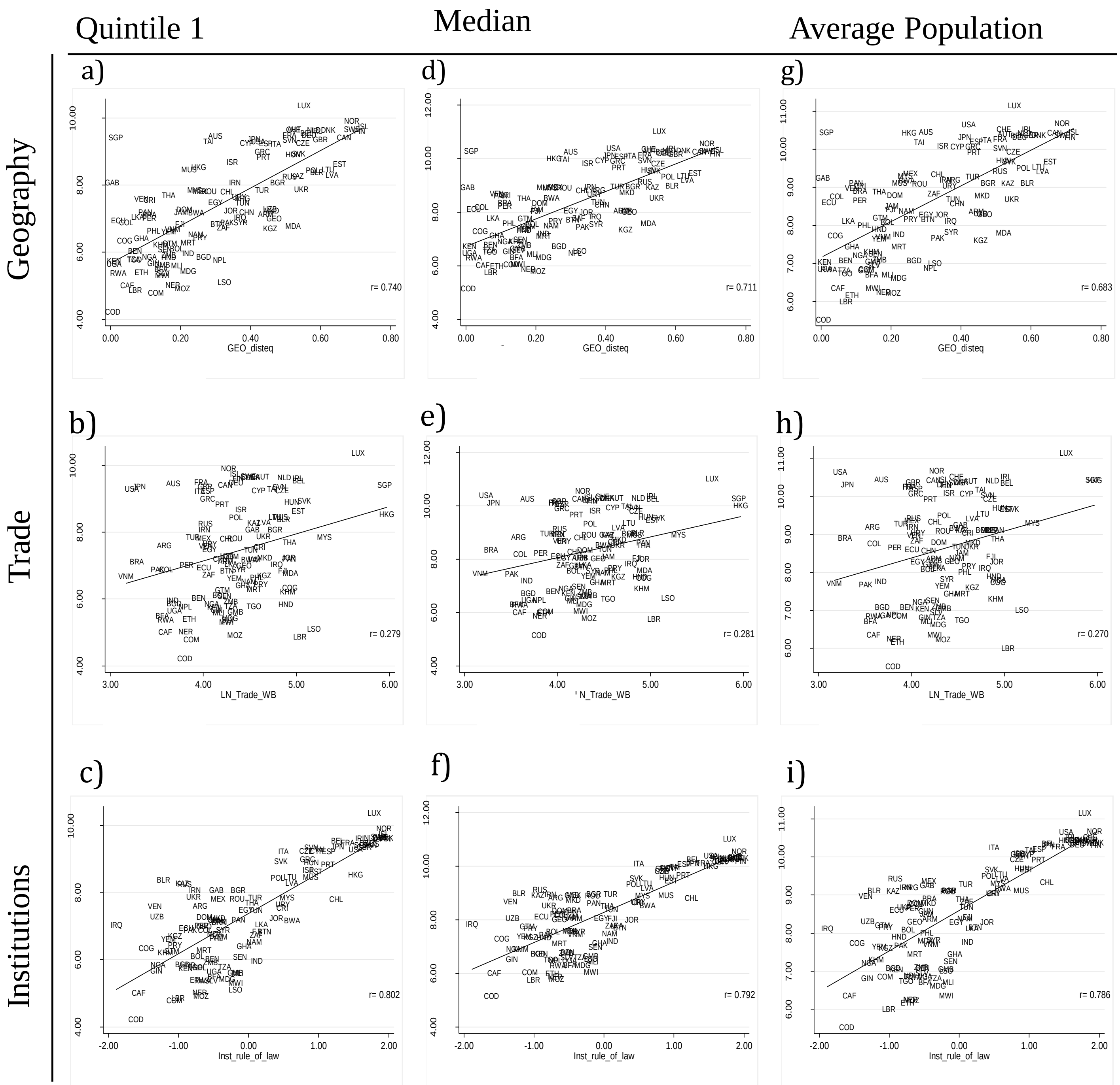


Figure 2.1: Linear correlations between log real GDP 2005 per capita for different income groups, and fundamental development factors (First Quintile for (a)-(c); Median for (d)-(f); Average Population for (g)-(i)). Linear prediction line and correlation coefficient r included.

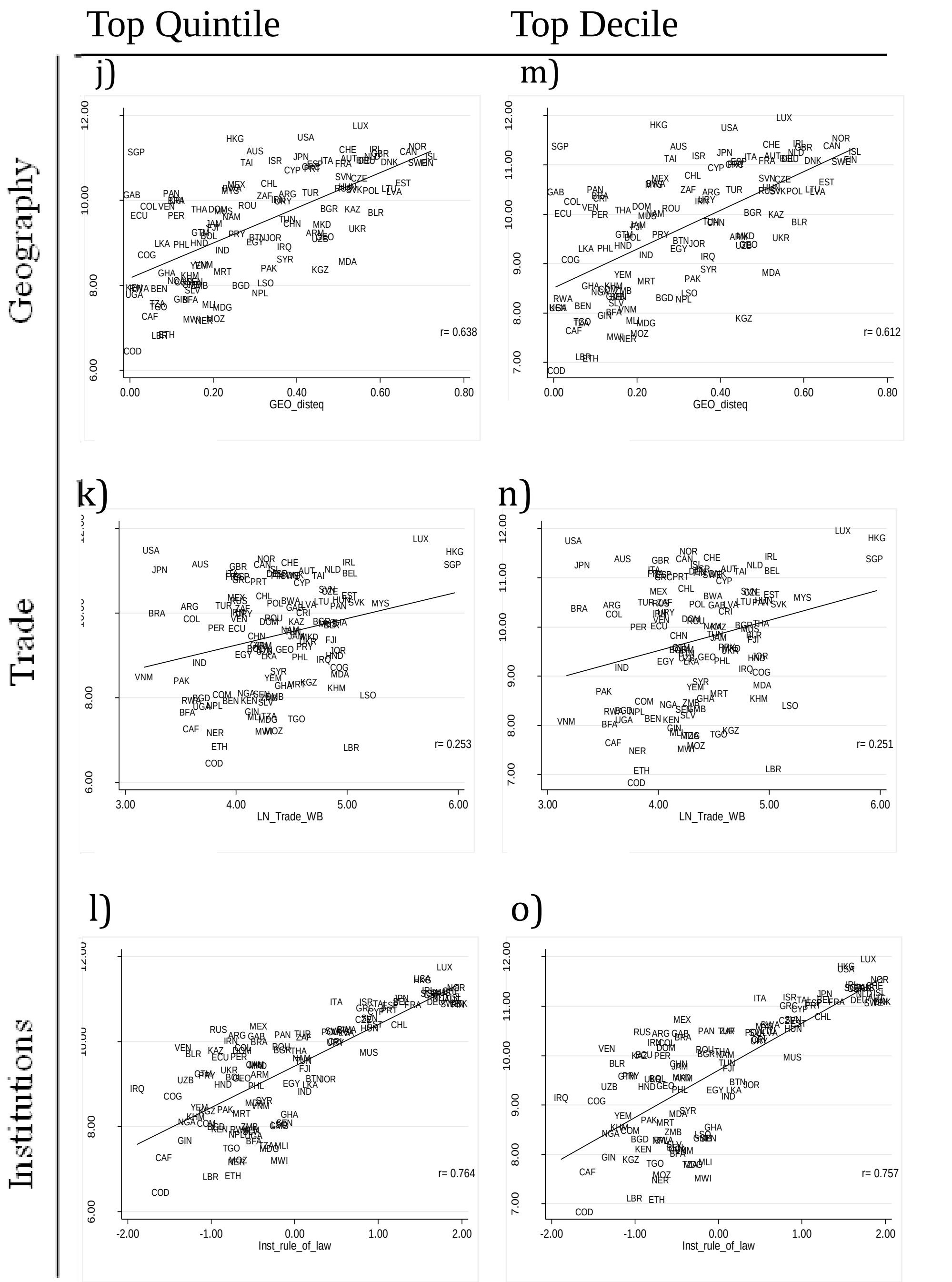


Figure 2.1 continued: Linear correlations between log real GDP 2005 per capita for different income groups, and fundamental development factors (Top Quintile for (j)-(l); Top Decile for (m)-(o)). Linear prediction line and correlation coefficient r included.

Table 2.2: Determinants of income: Base specification, ordinary least squares (OLS) estimates.

2005: Dependent variable = Log GDP per capita of	First Quintile			Median			Average Population			Top Quintile			Top Decile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Sample size	117	117	117	117	117	117	117	117	117	117	117	117	117	117	117
Geography	5.60	5.39	3.02	5.15	4.94	2.63	4.68	4.49	2.24	4.15	3.98	1.80	3.93	3.75	1.56
(GEO)	(0.46)***	(0.51)***	(0.49)***	(0.46)***	(0.50)***	(0.47)***	(0.45)***	(0.49)***	(0.46)***	(0.45)***	(0.49)***	(0.46)***	(0.45)***	(0.49)***	(0.46)***
Trade		0.36	0.16		0.37	0.18		0.33	0.15		0.30	0.12		0.31	0.12
(LN_TRADE_WB)		(0.27)	(0.15)		(0.27)	(0.15)		(0.26)	(0.15)		(0.26)	(0.15)		(0.26)	(0.15)
Institutions			0.82			0.80			0.78			0.75			0.76
(Inst_rule_of_law)			(0.09)***			(0.09)***			(0.08)***			(0.08)***			(0.08)***
RMSE	1.01	1.00	0.75	1.01	1.00	0.77	0.99	0.98	0.76	0.99	0.98	0.78	1.01	1.00	0.80
R-Square	0.54	0.55	0.75	0.50	0.51	0.71	0.46	0.47	0.68	0.40	0.41	0.63	0.37	0.38	0.60

Notes: The dependent variable is per capita GDP in 2005, PPP basis. There are five samples for which the regressions are run: (i) columns (1)-(3) refer to the bottom 20% income group; (ii) columns (4)-(6) regress the median income; (iii) columns (7)-(9) refer to the average per capita GDP; (iv) columns (10)-(12) regress the top 20% income group; and (v) columns (13)-(15) regress the top 10% income group. The regressors are: (i) GEO, the variable for geography, which is measured as the absolute value of latitude of country divided by 90; (ii) trade, the log share of imports and exports to national GDP; and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Instrumental Variables

As outlined before, reverse causality, omitted variable bias and measurement error may influence the OLS method inaccurately, i.e. endogeneity issues would violate OLS consistency. Hence, we proceed to two-stage least squares (IV) regressions with the established instrumental variables described in equations (2.3) and (2.4). After the OLS table showed the step-wise marginal effects of the individual fundamental development factors, results in table 2.3 are now presented directly with all variables included, but split along the six time periods.[18] We add to the summary of findings a note of caution. Due to the analysis across several time periods and extensive data results, our emphasis lies in discussing the broad, robust trends. In other words, we do not address all outliers, since these are in fact the time-specific deviations we try to attenuate through the analysis of more than one period. In addition, the first two time periods suffer from limited sample size and related potential bias. They are useful in extending the overall time period of investigation, and their findings are in line with later, more robust samples so that we see a quite stable pattern over 30 years. Still, the initial periods should be interpreted with caution.

Distance from the equator (geography) displays the most forceful dynamics across income groups in the IV specification. We see a uniform pattern of decreasing coefficients as we go from poor to rich; generally geography turns even negative once we passed the mean income. For 2005, for example, ceteris paribus each latitude degree further away from the equator corresponds to an expected 2.3 percent higher (log) income for the bottom quintile, while the top decile is basically unaffected. However, the variable is generally insignificant, with some exceptions for the first quintile. Hence, for large parts of the population, geography has little importance, which in our specification could be interpreted as institutions being the "deeper" cause and "trumping" geography as argued by Rodrik et al. (2004), and similarly by Acemoglu et al. (2001).

Trade, now instrumented with Frankel and Romer's constructed trade shares, shows two interesting features despite its limited explanatory power for the model. Trade consistently enters the equation with a minus, suggesting a negative effect of trade integration for income levels. This pattern is again reported analogously in Rodrik et al. (2004). Secondly, the coefficient increases in

---

[18] We do not adjust the standard errors in the IV-estimations by using the Delta method as described in Frankel and Romer to account for the generated variable *constructed trade*. Wooldridge (2002, p. 116-117) suggests that such an approach is justified in the case of generated regressors, but not necessarily for generated instruments. See also Frankel and Rose (2002) or Ondrich, Richardson, and Zhang (2006) who apply the same conceptual framework, but do not adjust the standard errors.

size and also in significance as we move towards the rich. This implies that an open economy seems to be more harmful for the rich than for the poor. Roine et al. (2009) report similar findings. Although trade coefficients here are not always significant, the pattern consistency may allow an interpretation. Trade rather leads to an equalization of income levels, as the poor are affected relatively less than the rich. It is our only fundamental variable that displays such an income convergence effect within a country. Potential reasons could be the removal of barriers to entry and consequently higher competition in a more open economy for hitherto monopoly-like or imperfect market structures.

Finally, the employment of Hall and Jones' language data, which constitute our instrument for institutional quality, demonstrates that institutions matter. Institutions display high significance for all incomes, together with an increasing coefficient from bottom to top. The effect on top income groups is on average 20 percent higher than for the poor. Furthermore, while not reported in the output table, there are again strong effects on the other variables once institutions enter the equation, as also seen in the OLS case. Specifically, the size of the geography coefficient drops while its significance vanishes, and trade consistently switches signs to negative when institutions are added. In line with OLS estimates, the R-square decreases substantially for the rich income groups. For the bottom quintile, the model is able to explain around 70 percent of variation. For the rich incomes, this value halves on average, and for one period the R-square is even close to zero for the richest decile. A plausible interpretation is that the variables employed have a much less determining impact for the rich, where omitted, i.e. individual factors play a gradually larger role.

We proceed to a set of tests for probing the validity of the model. The Pagan Hall test suggests that heteroskedasticity is present in selected periods, which leads us to always apply robust standard errors. The test for endogeneity, which is an adjusted version of the Durbin-Wu-Hausman test, yields mixed results regarding the necessity of an instrumental variable from a purely statistical perspective. Generally, endogeneity seems to be an issue, though the first income quintile displays a significant necessity for instrumentation only once, whereas the rich incomes have an opposite pattern. We also report Hansen's J tests of overidentifying restrictions under the null hypothesis that, roughly speaking, the instruments are valid. The results hint at a weakness of the model since in some time periods the p-values are barely insignificant and, hence, suggest potential issues of the instruments employed.

Table 2.3: Determinants of income: Instrumental variable estimates

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	56	56	56	56	56
Geography	0.08	-1.06	-1.54	-2.11	-2.30
(GEO)	(1.28)	(1.30)	(1.33)	(1.45)	(1.48)
Trade	-0.24	-0.28	-0.23	-0.24	-0.24
(LN_TRADE_WB)	(0.26)	(0.23)	(0.23)	(0.24)	(0.24)
Institutions	1.23	1.48	1.42	1.39	1.35
(Inst_rule_of_law)	(0.28)***	(0.29)***	(0.30)***	(0.33)***	(0.34)***
R-Square	0.60	0.60	0.53	0.41	0.34
Pagan Hall test (p-value)	0.51	0.20	0.17	0.16	0.16
Endogeneity test (p-value)	0.50	0.21	0.33	0.43	0.51
Hansen Test (p-value)	0.02	0.03	0.02	0.02	0.01

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	71	71	71	71	71
Geography	1.57	-0.63	-1.47	-2.60	-2.88
(GEO)	(1.81)	(2.19)	(2.22)	(2.43)	(2.48)
Trade	-0.48	-0.47	-0.41	-0.41	-0.40
(LN_TRADE_WB)	(0.34)	(0.36)	(0.35)	(0.38)	(0.38)
Institutions	1.30	1.59	1.59	1.67	1.65
(Inst_rule_of_law)	(0.40)***	(0.48)***	(0.49)***	(0.54)***	(0.55)***
R-Square	0.68	0.55	0.44	0.23	0.15
Pagan Hall test (p-value)	0.70	0.47	0.33	0.28	0.26
Endogeneity test (p-value)	0.18	0.02	0.01	<0.001	<0.001
Hansen Test (p-value)	0.09	0.07	0.07	0.07	0.07

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	107	107	107	107	107
Geography	1.99	0.88	0.09	-0.71	-1.04
(GEO)	(0.73)***	(0.85)	(0.88)	(0.97)	(1.00)
Trade	-0.37	-0.46	-0.47	-0.52	-0.53
(LN_TRADE_WB)	(0.28)	(0.28)*	(0.27)*	(0.28)*	(0.29)*
Institutions	1.29	1.42	1.43	1.47	1.48
(Inst_rule_of_law)	(0.21)***	(0.24)***	(0.25)***	(0.27)***	(0.28)***
R-Square	0.65	0.57	0.50	0.34	0.26
Pagan Hall test (p-value)	0.13	0.77	0.71	0.67	0.69
Endogeneity test (p-value)	0.28	0.09	0.02	<0.001	<0.001
Hansen Test (p-value)	0.03	0.09	0.17	0.27	0.31

Table 2.3 continued: Determinants of income: Instrumental variable estimates

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	84	84	84	84	84
Geography	0.55	-1.59	-2.61	-3.83	-4.27
(GEO)	(1.71)	(2.17)	(2.34)	(2.61)	(2.72)
Trade	-0.35	-0.65	-0.79	-0.96	-1.00
(LN_TRADE_WB)	(0.36)	(0.42)	(0.44)*	(0.48)**	(0.50)**
Institutions	1.61	1.98	2.09	2.24	2.30
(Inst_rule_of_law)	(0.43)***	(0.54)***	(0.58)***	(0.65)***	(0.68)***
R-Square	0.69	0.48	0.34	0.11	0.01
Pagan Hall test (p-value)	0.36	0.44	0.49	0.52	0.50
Endogeneity test (p-value)	0.20	<0.001	<0.001	<0.001	<0.001
Hansen Test (p-value)	0.23	0.56	0.61	0.69	0.68

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	117	117	117	117	117
Geography	2.03	1.39	0.94	0.41	0.12
(GEO)	(0.78)***	(0.86)*	(0.89)	(0.94)	(0.96)
Trade	-0.39	-0.61	-0.70	-0.82	-0.84
(LN_TRADE_WB)	(0.28)	(0.30)**	(0.31)**	(0.33)***	(0.34)***
Institutions	1.25	1.35	1.36	1.38	1.41
(Inst_rule_of_law)	(0.21)***	(0.24)***	(0.25)***	(0.27)***	(0.28)***
R-Square	0.68	0.57	0.50	0.39	0.34
Pagan Hall test (p-value)	0.54	0.42	0.36	0.28	0.20
Endogeneity test (p-value)	0.04	0.01	<0.001	<0.001	<0.001
Hansen Test (p-value)	0.12	0.07	0.07	0.07	0.07

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	91	91	91	91	91
Geography	2.04	1.20	0.66	-0.01	-0.25
(GEO)	(0.64)***	(0.67)*	(0.68)	(0.72)	(0.74)
Trade	-0.15	-0.32	-0.39	-0.48	-0.51
(LN_TRADE_WB)	(0.23)	(0.23)	(0.23)*	(0.24)**	(0.24)**
Institutions	1.00	1.11	1.13	1.16	1.17
(Inst_rule_of_law)	(0.15)***	(0.15)***	(0.16)***	(0.17)***	(0.18)***
R-Square	0.70	0.63	0.58	0.48	0.44
Pagan Hall test (p-value)	0.09	0.04	0.04	0.03	0.02
Endogeneity test (p-value)	0.59	0.11	0.04	0.01	<0.001
Hansen Test (p-value)	0.05	0.03	0.03	0.03	0.03

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) GEO, the variable for geography, which is measured as the absolute value of latitude of country divided by 90; (ii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. The Pagan Hall tests heteroskedasticity for instrumental variables (IV) estimation under the null of homoskedasticity. The endogeneity test is based on the Durbin-Wu-Hausman test, but adjusted here for heteroskedasticity. The Hansen Test follows the standard methodology.

We discuss the underlying reasons for these results now in greater detail as we turn to the first-stage regressions, reported in table 2.4. While estimates for both instrumented variables indicate a reasonable R-square, the first stage model fit for institutions is always better than for trade. This is mainly due to several input variables that significantly affect institutions in addition to the instrumental variables employed. The actual instrument, based on Hall and Jones' language data, is consistently significant for institutional quality.[19] However, constructed trade shares are also significant for institutions, and so is geography. This pattern confirms observations reported by Rodrik et al. (2004). Our first-stage estimate for trade displays a reasonable R-square of roughly 0.5. In the equation, constructed trade shares prevail as key variable, with significance on the one percent level throughout all time intervals. While the other variables have occasionally a significant influence on a nation's trade share as well, only constructed trade shares show a robust pattern over time, thus confirming the validity of the instrument. This leads to a major conclusion: though the instrument for a nation's trade integration proves to be highly significant, trade shares display no positive effect on income levels in the second stage of our model. This makes the outcome even more robust: trade integration does rather not increase a nation's prosperity (it even harms top incomes), but likely leads to income convergence.

Let us now revisit the validity of the IV model. For this, we test for underidentification and for weak instruments. We employ the Lagrange-Multiplier (LM) test using the rank-based rk statistic from Kleibergen and Paap (2006), where significant values indicate valid identification, i.e. the excluded instruments are relevant.[20] Results indicate that the aggregate model is always identified. In addition, the granular identification check for each instrument via the method described by Angrist and Pischke (2009, p. 217-218) yields positive results. Both trade and institutions are identified on significant levels. In the next step, we go one step deeper and analyze via two methods whether strong or weak identification is present. First, we report the first-stage F-statistics and contrast them to the "rule of thumb" values suggested by Staiger and Stock (1997).

---

[19] Depending on the time period, this ranges from the one percent significance level to the ten percent significance level. Thereby, either the component *fraction of the population speaking English as mother tongue*, or the second component *fraction of the population speaking one of the major languages of Western Europe as mother tongue: English, French, German, Portuguese, or Spanish;* or both components are significant

[20] A rejection of the null indicates that the matrix is full column rank, i.e., the model is identified. The rk statistic, also distributed as chi-squared with (L1-K1+1) degrees of freedom, can be seen as a generalization of these tests to the case of non-independently and -identically distributed errors. This approach follows the diagnostics suggested by Bazzi and Clemens (2013).

Table 2.4: First stage results of two-stages least square estimates

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	56	56	71	71	107	107	84	84	117	117	91	91
Geography	0.01	3.49	-0.13	3.73	0.51	2.99	0.41	3.82	0.44	2.79	0.29	2.94
(GEO)	(0.28)	(0.41)***	(0.21)	(0.39)***	(0.21)**	(0.37)***	(0.22)*	(0.32)***	(0.20)**	(0.37)***	(0.25)	(0.44)***
Constructed Trade	0.55	0.18	0.49	0.16	0.42	0.24	0.40	0.23	0.40	0.40	0.46	0.41
(Trade_FR_ROM)	(0.07)***	(0.10)*	(0.06)***	(0.12)	(0.05)***	(0.09)***	(0.06)***	(0.09)**	(0.06)***	(0.10)***	(0.06)***	(0.11)***
Pop. speaking English	0.43	0.90	0.34	0.66	0.25	0.45	0.21	0.15	0.06	0.88	0.11	1.19
(Eng_Lang)	(0.16)***	(0.34)***	(0.15)**	(0.41)*	(0.17)	(0.37)	(0.19)	(0.32)	(0.16)	(0.51)*	(0.14)	(0.37)***
Pop. speaking other European languages	-0.20	0.20	-0.13	0.32	-0.18	0.65	-0.14	0.43	-0.14	0.58	-0.23	0.49
(EUR_Lang)	(0.12)*	(0.21)	(0.11)	(0.18)*	(0.11)*	(0.17)***	(0.10)	(0.16)***	(0.08)*	(0.18)***	(0.10)**	(0.20)**
First-stage F-test	21.8	3.5	26.1	2.43	27.4	9.27	16.8	4.36	20.7	8.8	25.6	12.6
Angrist-Pischke F-statistics (p-value)	<0.001	0.01	<0.001	0.03	<0.001	<0.001	<0.001	0.01	<0.001	<0.001	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	0.04		0.04		<0.001		0.01		<0.001		<0.001	
Kleibergen-Paap Wald rk F statistic	3.49		2.36		8.13		3.43		6.39		12.25	
Stock-Yogo critical values 10%	13.43											
Stock-Yogo critical values 25%	5.45											
R-Square	0.60	0.70	0.54	0.66	0.42	0.50	0.42	0.67	0.39	0.50	0.49	0.57

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) GEO, the variable for geography, which is measured as the absolute value of latitude of country divided by 90; (ii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iii) the proportion of the population of a country that speaks English (Eng_Lang); and (iv) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified.The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

The F-statistics display no issues for trade, but the picture looks different for institutions. Here, only in half of the periods analyzed the F-statistic is close to or above the critical value of ten. This ties back to the problematic values of Hansen's J-test regarding the model validity, which can now be pinpointed to institutions. However, we regard this mostly an issue of finite sample bias as more recent time periods with larger samples yield valid results.

In addition, we probe for the strength of the instruments with the diagnostic developed by Stock and Yogo (2005). For this purpose, we calculate the first-stage Kleibergen-Paap Wald F-statistics which is compared to Stock and Yogo's critical values.[21] Under the null of weak instruments, the test statistic is based on the rejection rate r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5 percent. Weak instruments are hence defined as instruments that will lead to a rejection rate of r when the true rejection rate is 5 percent. Results indicate that we can never reject the null of a rejection rate above 10 percent, but can do so for the less strict threshold of 25 percent.[22] Overall, the equations are always identified, but weak instruments, specifically for institutions, are an issue in some of the time periods. We therefore focus the results discussion on the more robust time periods, even though all time period estimations – independent of weakness or strength of the IV specification – have similar outcomes. Besides, the close alignment with standard literature supports our methodological approach.

Finally, we pick the 2005 period and two countries, Austria and Bangladesh, to understand if the IV estimates are also economically meaningful when examining the bottom quintile and the top decile. In terms of institutional quality, Austria (1.89) ranks considerably higher than Bangladesh (-0.94) for the given time period. In the model, this alone translates into an income difference between the poorest 20 percent in both countries of the factor 34 (log difference increase by 254 percent). The fact that Austria is located 25 degrees of latitude further away from the equator additionally increases the income of its bottom quintile by two thirds (0.5 log difference) compared to the respective Bangladeshi group. Lastly, the trade share of GDP in Bangladesh is at around 80 percent of what Austria reports. Hence, we would expect a 30 percent *reduction* of Austrian relative to Bangladeshi incomes for our group. Multiplying effects of the three

[21] Note that Cragg-Donald or Anderson LM tests cannot be used for neither under- nor weak identification as they are only valid under the assumption of independent and identical distribution.

[22] Critical values have not been tabulated for the Kleibergen-Paap rk statistic which we employ here due to heteroskedasticity concerns. Nonetheless, we follow the literature and apply the critical values for the Cragg-Donald statistic to the Kleibergen-Paap values (see Bazzi & Clemens, 2013; Baum, Schaffer, & Stillman, 2007).

development factors, our model would calculate an income gap of the factor 40 between the two countries (log difference of 3.7). The actual income difference for the bottom quintile amounts to a factor of 45, respectively a log difference of 3.8.

For the richest 10 percent, most of the expected income difference between the two countries is attributed to institutions. This is due to the vanishing role of geography in the specification together with a negative, but in absolute figures small effect of trade. The model predicts a 25-fold difference between the two country's richest groups based on the three variables; the actual income gap amounts to roughly 18 times. A short simulation reveals the overwhelming power of institutions. If geographic and trade variables of Bangladesh were kept constant, but the institutional quality raised to the level of Austria, the estimated Bangladeshi bottom quintile income would increase to $20,000 – higher than the actual income reported for the Austrian bottom quintile. The model suggests that the richest decile would increase its income level even by a factor of 50 to almost $400,000. Actual average income levels reported for Bangladesh during that time period contrast sharply with $1,200.

II. 5. Refinements and Robustness tests

Refinement Analyses

In figure 2.2, we show confidence intervals to illustrate the differences between variable coefficients per income group, exemplified for 2005.[23] Given the criticism of a strict interpretation of significance tests and the question whether the different income groups can be treated as independent samples, we opt for a graphical interpretation. Also, in order to enable better comparison of interval sizes we standardize the regression variables.[24] We see the weaker dynamics between rich and poor incomes for the variables trade and institutions reflected in the graph. Still, there is a distinct positive, and hence income-equalizing trend for trade as we move from top to bottom incomes, whereas effects of institutions are quite stable throughout all income groups. In contrast, geography shows the described strong movement, such that the point estimate for the bottom quintile lies outside of both the 99 percent and 95 percent confidence interval of the top decile. Geographic conditions treat poor and rich strikingly different.

[23] The same illustration for all time periods can be found in figure 2.3 in the appendix.
[24] We perform a regular z-score standardization of variables by taking for each variable the difference between individual values and the mean, and then dividing by the standard deviation.

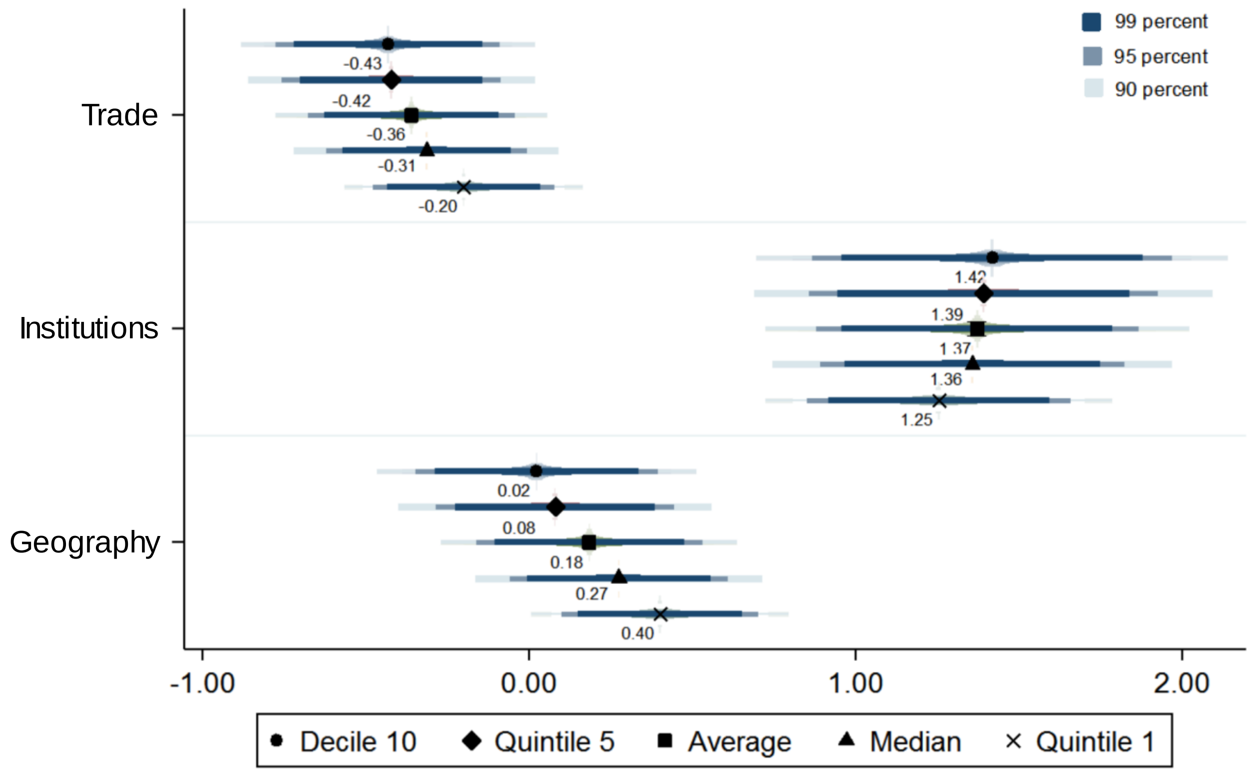


Figure 2.2: Confidence intervals for 2005 per standardized variable of the fundamental development factors (trade, institutions, and geography), each broken down to the respective income group. Coefficients are labeled with their estimated (standardized) coefficients.

Let us then have a closer look at the underlying dynamics of the change in R-square depending on the regressor. We already noted that trade adds nearly no explanatory power to the overall model, whereas geography and institutions drive up the R-square. In order to understand the stand-alone explanatory power of institutions, which is dominating both OLS and IV estimates, we run alternative specifications where this variable enters first, followed by either trade or geography.[25] Results are presented in table 2.5. The weak impact of trade is indeed reflected in this analysis. The marginal effect of geography on the R-squared is positive, but with a clear income group-dependent pattern. The more we move towards the rich, the less of the income variation can be explained via geographical conditions.

[25] Hereby, neither trade nor institutions are instrumented since we are interested more in how much of the data variation can be explained by the fundamental development factors.

Table 2.5: Marginal effect of institutions and other fundamental development factors

(1)	(2)	Fundamental development factor (3)	Institutions (4)	ΔR-squared (5)
Quintile 1	Trade	0.25	1.16	<0.01
		(0.16)	(0.08)***	
	Geography	3.07	0.83	0.10
		(0.51)***	(0.10)***	
Median	Trade	0.25	1.10	0.01
		(0.15)*	(0.07)***	
	Geography	2.68	0.81	0.09
		(0.49)***	(0.09)***	
Average Population	Trade	0.21	1.03	0.01
		(0.14)	(0.07)***	
	Geography	2.28	0.79	0.07
		(0.47)***	(0.09)***	
Top Quintile	Trade	0.17	0.96	<0.01
		(0.14)	(0.07)***	
	Geography	1.84	0.76	0.05
		(0.46)***	(0.08)***	
Top Decile	Trade	0.17	0.93	<0.01
		(0.15)	(0.07)***	
	Geography	1.60	0.77	0.04
		(0.47)***	(0.08)***	

Notes: The dependent variable is per capita GDP in 2005, PPP basis. The table reports estimates of equation (1) when only two regressors are included simultaneously: institutions and either trade or geography, as indicated in the second column. The last column shows the increase in the adjusted R-Squared when a fundamental development factor is included in addition to institutions. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

While not reported in the table, institutions alone also have a larger explanatory power than geography in a single regressor specification. For the first quintile the R-square difference is 0.1; the top decile shows a 0.2 higher R-square when we only regress institutions instead of only geography. The latter outcome again reflects that the richer people are, the less geography matters. Yet, the sign of the estimated effects, the relative size effects between the development factors, and the coefficient dynamics between income groups are consistent.

Additional Controls

We begin our additional robustness tests by probing the reported impact of institutional quality. The need for instrumenting the regressor "institutions" is undisputed. Yet, we are aware that our instrument (based on Hall and Jones, 1999) might impact a country's income not only via institutions, but also via other channels such as specific trade patterns caused by a common language with other countries. To address this concern, we employ settler mortality data as alternative instrument for institutions, which has been proposed by Acemoglu et al. (2001). While methodologically widely recognized, the data availability of settler mortality poses constraints on our sample size. We have to cut down significantly Acemoglu et al.'s (2001) original database of 81 countries, since for many countries there are no data available on income distribution. In return, no settler mortality data exist for most industrialized countries. Specifically European countries, which represent a significant share of our sample, cannot be assigned values or "borrow" from neighboring countries. China would also be missing, which is difficult to justify given its enormous impact on shifts in global trade and income levels over the last 30 years. The settler mortality instrument has furthermore been criticized by Albouy (2012), who lists that 36 countries are assigned mortality rates from other countries, often based on mistaken or conflicting evidence. He concludes that once these cases are controlled for, the instrumental variable estimates become unreliable, and the overall model lacks robustness.

Nevertheless, conducting a robustness check with this alternative instrument does not change the picture (see appendix, tables 2.8 and 2.9). Geography continues to indicate that, relatively speaking, the richer incomes are worse off the further away from the equator they are. Trade also behaves similarly to the core specification, although larger data variation leads to even less precise point estimates. Institutions remain the only highly significant variable and are assigned larger coefficients with this instrument than with Hall and Jones' language data. Overall, poor R-squared values for the second stage estimates in combination with weak first-stage results, and the small sample size (only 2005 has more than 50 countries in the sample) represent validity concerns for this estimation alternative. Thus, while confirming the core specification, we choose to keep language data as preferred instrumental approach.

We also take a closer look at the regressor "geography". While it broadly represents the idea that not all areas of the world have equal natural characteristics, scholars advocate distinct mechanisms how these natural differences affect income levels. The variable in our core specification (distance from the equator) may be regarded as an overarching proxy which could overlook the specific underlying mechanisms. Therefore, we apply two alternative geography variables to test if they drive estimates in a different direction. First, we employ "mean temperature per country" as regressor, which represents different climatic conditions (for an economic growth discussion see for example Diamond, 1997; Montesquieu, [1748] 1989; Sachs, 2001). Results confirm the original geography estimates very closely (see tables 2.10 and 2.11 in the appendix). For rich incomes, having warmer average temperatures is somewhat beneficial (i.e. being close to the equator), whereas for the poor effects are negative. Trade and institutions also behave analogously to the original specification.

The second alternative for geography we apply is the prevalence of malaria, which tests if disease burdens can explain income levels (Gallup & Sachs, 2001; Sachs, 2000). Our robustness test supports this hypothesis (see appendix, tables 2.12 and 2.13). Geography in this specification is now always significant (in contrast to the base specification), but negative effects of malaria are continuously greater as we go from rich to poor. This is similar to previous findings in a sense that rich income groups seem to be unaffected by hot, tropical climate conditions close to the equator. Trade shows very little change vis-à-vis the original specification (trade is positive for the poor, but rather negative for the rich), and institutions remain consistently highly significant, albeit with lower coefficients than before. The overall model has a better explanatory power, so it seems that diseases are the more important underlying variable driving geography than temperature levels.

In a further step, we include two additional controls which have both been identified as important factors for development and income levels: cultural influences and health conditions. For culture, we test on the one hand whether colonial history impacts the results. Adding a European colony dummy does not have a consistently significant effect (see tables 2.14 and 2.15 in the appendix). All other fundamental factors are robust in terms of relative changes across income groups and significance levels. If anything, the colony dummy somehow helps estimate the trade variable more precisely so that its already discussed pattern turns significant for more estimates. However, this does not go at the expense of institutions or geography whose results remain fully robust. Next, we employ the ethno-linguistic fractionalization indicator by Alesina et al. (2003) as alternative cultural variable. This control variable leaves also no sizeable impact (see tables 2.16 and 2.17 in the appendix). Institutions, trade, and geography continue to behave the same way as in the original core specification. Similar to the colonial dummy, we also see no positive changes in the R-square that would suggest keeping either variable.

For measuring health conditions, we take the life expectancy at birth in 1970 as control variable (see tables 2.18 and 2.19 in the appendix). Entering this variable creates some turmoil in the estimates. With few exceptions health is always highly significant, and displays a tendency for a decreasing coefficient as we go from poor to rich. Although the relative patterns for geography, trade, and institutions persist, the latter is no longer persistently significant at the 1 percent level. Still, overall broad robustness remains. The improved R-square for the second stage when including health is opposed to fragile first stage results in such a specification. In particular the very weak first stage F-statistics which lie below our original specification cast doubt on the validity of the overall results. Therefore, we do not adjust our core model despite recognizing the valuable role of health.

We also test the robustness of the model when controlling for the 20-year lags of national per capita incomes, recognizing that including this variable is quite problematic due to endogeneity concerns: historical income levels will affect current trade patterns as well as today's institutional quality, while historical incomes have been influenced in return by past levels of institutional quality and trade integration. As we cannot resolve this methodological issue, we interpret estimates highly cautiously and focus on whether our usual coefficient dynamics across income groups still hold. This is indeed the case for geography and trade, whose dynamics qualitatively persist (see appendix, tables 2.20 and 2.21). Institutional quality, however, is rendered insignificant and displays no longer a clear pattern between poor and rich. We attribute this finding mainly to described endogeneity problems which affect especially this variable. Its first stage results are also very poor and suggest substantial bias, so that we abstain from drawing strong conclusions. Lagged income itself has unsurprisingly a strong effect on today's income at all levels, but we observe no distinct trend between income groups.

Finally, we control for the general income inequality in a country through the GINI index (tables 2.22 and 2.23 in the appendix). We run this check to find out whether our three core variables might affect income groups differently depending on how unequal the overall income distribution is. In other words, this analysis tackles omitted variable bias in yet another dimension to ensure that our core estimates are clear of general inequality patterns in a country. Inspection of results indicates that a control for overall income inequality curbs the coefficient movements between income groups considerably. The GINI Index seems to absorb most of the between-income group variance, as the variable itself displays a consistent sign switch from bottom to top incomes. Consequently, favorable geographic conditions still tend to be relatively more positive for the poor than for the rich, but we no longer observe forceful coefficient dynamics including sign switch. The negative effect for trade integration also persists, but we no longer see a clear pattern between income groups. The same holds true for institutional quality, which remains highly significant, but rather flat. Clearly our control mitigates previously documented effects between income groups.

However, we remain skeptical regarding the econometric validity of this specification. National income inequality, similar to lagged income levels (our previous robustness check) cannot be treated as exogenously given. Rather, it is impacted by our (exogenous) fundamental development factors geography, trade integration, and institutional quality, as well as by our outcome variable itself. Consequently, correlation with the error term and resulting simultaneous equations bias seem highly probable. Often significant Hansen J-test statistics furthermore point at identification issues. We therefore believe that an income inequality control, if not explicitly exogenous, cannot be defended methodologically in our empirical approach. The estimates nonetheless suggest that inclusion leaves core results robust, but curbs the previously documented forceful coefficient dynamics.

The Role of Human Capital

So far we collected evidence for institutions as a likely key variable for income growth, thus following a number of "institutional advocates" (e.g., Acemoglu et al., 2001; Knack and Keefer, 1995; Mauro, 1995). However, a second strand of literature posits that a third variable, namely human capital, plays a key role in this relationship. Findings by Glaeser et al. (2004) indicate that human capital is a more basic source for growth than institutions. Specifically referring to the settler mortality instrument applied by Acemoglu et al. (2001) to measure institutions, the authors argue that the colonists brought human capital in addition to their knowledge of how to build good institutions. Hence, human capital led to enhanced institutions, which subsequently spurred economic growth. Murtin and Wacziarg (2014) also conclude that human capital underlies good institutions. The discussion reflects that the interaction and the way causality runs between human capital, institutional quality, and growth represent important aspects of economic development. Therefore, let us take a closer look at human capital in the context of our research specification.

In general, human capital plays a prominent role in a number of models of endogenous growth (Barro, 1991; Romer, 1990; for an extensive review, see also Savvides & Stengos, 2008). While from a theoretical, and perhaps also from an intuitive point of view one could expect a positive income effect from human capital, empirical findings have been rather mixed. A number of papers attribute a positive impact (Barro, 1991; Glaeser et al., 2004; Mankiw et al., 1992;

Mincer, 1974; Sala-i-Martin et al., 2004). Yet, there is also counterevidence that reports insignificant or even negative effects, often by using an alternative definition of human capital or by applying a different measurement (Benhabib & Spiegel, 1994; Krueger & Lindahl, 2001; Pritchett, 2001; Wolf, 2002). In this context, causality issues have been an ongoing methodological concern (Griliches, 1977). Countries that grow faster have the resources to invest in schools and education so that growth could cause higher human capital.

More recent causality analyses by Hanushek and Woessmann (2011, 2012), however, lend support to human capital causing economic growth, not vice versa. The literature for measuring the effects of human capital on different income groups is inconclusive, too (for a survey, see Psacharopoulos & Woodhall, 1985). Ram (1984, 1989) finds no significant effects of schooling on changes in income distribution. Dollar et al. (2013) conclude similarly when looking at the bottom 20 and 40 percent income groups. De Gregorio and Lee (2002), on the other hand, report that educational factors play some role in altering the overall income distribution.

To assess human capital effects in our empirical framework, we extend now our core specification by this variable. We measure it via primary school enrolment rates, where values for each time period are instrumented with lagged enrolment rates (average of 1970-1979) to address endogeneity issues. School enrolment rates may be criticized as unit of measurement, since they equal human capital with knowledge acquired in school, and also assume that one year of schooling covers the same amount of learning everywhere. However, alternative measures would shrink the sample size significantly, while generally no variable we are aware of is able to perfectly capture all facets of human capital.

Results indicate that human capital plays a highly significant role in explaining income levels (see appendix, tables 2.24 and 2.25). The point estimates are generally very precise, but a pattern along income groups is not discernible. Until 2000, top incomes have lower coefficients assigned than bottom incomes, but this trend reverses for the three following time periods. Human capital does not alter the overall picture of the original specification though. Distance from the equator remains clearly more positive for the poor than for the rich. All geography coefficients increase in direct comparison to the core specification by about 1.5, so that the relative benefit of top incomes from a tropical environment is damped. Trade integration keeps the generally negative income effect which increases in size and significance as we move towards the rich. As an interesting outlier, the first quintile tends to display positive values in this specification. This re-confirms our earlier interpretation that the poor do not suffer from a more open economy. Finally, institutions remain highly significant and show the known pattern of increasing coefficients as we move towards the rich. However, the coefficients in the original specification are on average one third larger than observed here, so that human capital has a sizeable reductive effect.

Altogether we find no clear evidence that human capital is the more basic source of growth than institutions as argued by Glaeser et al. (2004). Two points speak against such a conclusion. First, institutions appear as more robust than human capital in the second stage of our model, as they consistently display highly significant point estimates. Human capital sometimes lacks significance, for example in our last period of 2010 or for some of the top deciles. Secondly, our first stage results show that human capital has a significant effect on institutions only in half of the periods, and even then it is never the strongest predictor.[26] Clearly human capital matters for development throughout all income groups, but we remain skeptical in going as far as calling it more fundamental than institutions.

Robustness of the Outcome Variable

We now turn to a robustness discussion of the dependent variable. For this purpose, we add a control variable, and we also take alternative outcome variables based on the Occupational Wages Worldwide (OWW) database as well as the World Top Incomes Database (WTID). Ciccone and Jarociński (2010) find that international income data play a highly sensitive role for growth regressions. In general, different sources and methods applied per country affect the data quality (for an extensive discussion of this issue, see Atkinson & Brandolini 2001, 2009) and may affect the validity of the results.

The survey data forming the UNU-WIDER database of income distribution that we regularly use are accompanied by a set of cautionary notes from the authors regarding data quality. Industrialized countries, for example, typically measure income distribution with reference to income, not consumption, and so does Latin America. In contrast, Asian and African surveys usually collect consumption data to measure income dispersion (UNU-Wider 2014b). While the database attempts to collect and harmonize both forms of income measurement, we nonetheless add regional dummy control variables to our original specification to check for regional biases

[26] In 1985, 1990, and 1995, where human capital is a significant predictor for institutions on a 5%-level, geography and/or the actual instrument for institutions, namely Hall and Jones' (1999) language data, matter more.

(see appendix, tables 2.26 and 2.27).[27] This also ties back to Balakrishnan et al. (2013) who report major regional differences in their regression model.

Sub-Saharan Africa appears as the only region that influences different income groups significantly. While institutions also remain highly significant, their coefficient pattern across income groups is now fluctuating depending on the period and hence inconclusive. The usual pattern for geography (the richer, the more beneficial to live close to the equator) is also no longer as clearly visible. Finally, the regional dummies seem to paralyze the trade variable which barely moves away from zero. While this does not provide direct evidence that data inconsistencies affect our core results, regional control variables do have a sizeable effect on the results.

We perform a second check on the dependent variable by applying an alternative dataset that is based on individual data. This is to tackle two potential issues of the income distribution data used so far. On the one hand, our usual dataset is newly constructed out of two variables, which individually might not have been designed for such a purpose, thus biasing the results. On the other hand, we only used macro data, as the dependent variable is based on aggregate national level data. Now we employ the most recent Occupational Wages Worldwide (OWW) data to address these two concerns simultaneously.[28] OWW data contain individual wage levels for various occupations per country from poor to rich, and can consequently serve as proxy for the income levels of different groups. Based on the wage data distribution averaging the years 2003-2007 (i.e., our time period 2005), we calculate the respective national deciles and regress them on geography, trade integration, and institutional quality.

Results for 2005 broadly confirm our earlier findings (see tables 2.28 and 2.29 in the appendix). Geography displays the regular switch from positive to negative coefficients as we go from poor to rich. Trade continues to play a quite unimportant role in terms of significance levels and absolute size of coefficients. The latter are, however, not always negative as seen in our regular specification. Institutions also keep their high significance for all incomes, which is in line with earlier findings. Yet, the coefficient pattern across income groups is not as clear as seen before.

---

[27] The definition of world regions is built on the classification of the World Bank, but slightly adjusted to reflect sample specifics. The resulting six regions we use are Sub-Saharan Africa, Middle East and North Africa, Europe and Central Asia, North America, Latin America and the Caribbean, and South-East Asia and Pacific.

[28] Another potential alternative source would be the World Bank PovcalNet database. Although an extensive dataset, it takes again a macro perspective, and is less used for percentile data analysis across the entire distribution of income, i.e. beyond the bottom end. Dykstra, Dykstra, and Sandefur (2014) also warn that estimates of the densities near the bottom and top tails of the distribution could be quite unreliable, while no attempt has been made by the Bank's staff to validate the tool for such purposes.

Overall, the OWW specification reaffirms that the model is better suited for explaining incomes of the poor than of the rich.

Our third alternative specification for the outcome variable resorts to the World Top Incomes Database (WTID) by Alvaredo, Atkinson, Piketty, and Saez (2015). Their database has been a source for a number of significant contributions to the literature on income inequality, in particular for top income groups (Atkinson et al., 2011; Piketty, 2014). Income shares here are calculated exclusively based on tax statistics, which differs from related databases such as the UNU-WIDER. These alternative data may hence serve to replicate previous results and cross-check their robustness. Unfortunately, we can directly compare only one of our five income groups (top decile) as the WTID focuses exclusively on top incomes. However, the data also allow us to newly include the top 1% as additional income group, so that we may extend our usual break-down of income levels.

We estimate effects of the three core explanatory variables on top 10% and top 1% incomes for our usual six time periods individually, and we also estimate a panel. For a digestible summary, in table 2.6 we only document two cross-sections and the panel, but the other four cross-sections are qualitatively identical. The cross-sections may suffer from finite sample bias due to the limited number of countries in the WTID, but nevertheless, the findings strongly reaffirm previous results. This is not surprising given the extreme similarity of the two data sets, which show for the available countries a correlation between 0.95 and 0.98 for the top decile. All regressors show the same pattern for top incomes as before: proximity to the equator is positive, trade integration negative, and institutional quality positive.

Significance levels of the WTID estimates are also similar to the UNU-Wider estimates, but the absolute coefficient size tends to be a bit lower. Still, we observe that within the WTID, effects for the top 1% income group are always larger in absolute size than for the top 10% income group. This reconfirms and extends our previously observed pattern from the UNU-Wider estimates such that, as we move from poor to rich, all regressors follow a distinct linear trend. Specifically, proximity to the equator becomes more positive, trade integration more negative, and institutional quality more positive. The WTID results show that this continuous trend is non-reversing and also holds true for the very high income groups.

Table 2.6: Determinants of income for high income groups: Specifications using World Top Incomes Database (WTID)

	1995			2005			Panel 1985-2010		
Dependent variable = Log GDP per capita of	Top Decile (UNU-Wider)	Top Decile (WTID)	Top 1% (WTID)	Top Decile (UNU-Wider)	Top Decile (WTID)	Top 1% (WTID)	Top Decile (UNU-Wider)	Top Decile (WTID)	Top 1% (WTID)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Observations	107	16	19	117	18	20	192	144	103
Geography	-1.04	-2.29	-10.54	0.12	-0.06	-1.41	-3.83	-1.15	-3.13
(GEO)	(1.00)	(1.15)**	(7.79)	(0.96)	(0.88)	(1.14)	(1.85)**	(0.88)	(1.90)*
Trade	-0.53	-0.17	-0.87	-0.84	-0.23	-0.46	-0.30	-0.21	-0.37
(LN_TRADE_WB)	(0.29)*	(0.17)	(0.66)	(0.34)***	(0.20)	(0.31)	(0.24)	(0.16)	(0.30)
Institutions	1.48	1.42	3.18	1.41	0.92	1.07	1.52	1.12	1.69
(Inst_rule_of_law)	(0.28)***	(0.32)***	(1.93)*	(0.28)***	(0.24)***	(0.39)***	(0.41)***	(0.31)***	(0.61)***
R-Squared	0.26	0.83	n/a	0.34	0.62	0.45	n/a	n/a	n/a

Notes: The dependent variable is per capita GDP on PPP basis. For each specification, there are three alternative dependent variables on which the independent variables are regressed: columns (1), (4), and (7) take the usual data from the UNU-Wider, and can be found analogously in other output tables. Columns (2), (5), and (8) show effects for the top decile of the World Top Incomes Database, and columns (3), (6), and (9) show the top one percent of the World Top Incomes Database.Columns (1) to (6) are cross-section linear 2SLS models, whereas columns (7) to (9) uses the "system GMM", based on Arellano and Bover (1995) and Blundell and Bond (1998). The regressors are: (i) GEO, the variable for geography, which is measured as distance from equator; (ii) trade, the log share of imports and exports to national GDP; and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index . See the Appendix for more detailed variable definitions and sources). "n/a" for R-squared denotes negative values, respectively "not applicable" in the GMM panel estimations. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Robustness of Overall Model

After having discussed the robustness of the left and right hand side of the regression equation, we address here potential weaknesses related to our overall econometric approach. The sample size might be considered insufficient for some of the time periods analyzed. Indeed, only two out of the six periods contain more than 100 countries, while two other periods contain less than 75. This, however, is an inherent limitation to most cross-country regressions that cover several decades. In addition, accompanying tests indicate that finite sample bias is no major issue as outlined earlier.

With the exception of one refinement in table 2.6, the identification strategy taken until here has exclusively resorted to a linear IV model, estimating the effects on our six time periods individually. We now adjust these two methodological features more systematically to probe the effect on our estimates. For this purpose, we employ the Generalized Method of Moments (GMM) estimator and transform the six individual time periods into one panel data set. While this has the downside of substantially reducing the sample size, it can exploit the time series feature of the dataset. The overall model specification then follows Arellano and Bover (1995) and Blundell and Bond (1998) and is also referred to as "system GMM".[29] It is designed for panels that may contain fixed as well as idiosyncratic errors which are heteroskedastic and correlated within but not across individual cases (Roodman, 2009). We run the model in two variations, where the first option employs our usual instrumental variables, while the second uses as instruments all available lags of the regressors themselves in levels. Each option is applied on three alternative panel sets with increasing sample size, namely 1985-2010 (32 countries), 1990-2010 (43 countries), and 1995-2010 (55 countries).

Results for the first option are given in table 2.7. The GMM estimator closely confirms the findings of our linear model. Geography displays its usual pattern as richer income groups increasingly benefit from equator proximity, and we see the sign switch in two of the three panels. Trade integration has a negative effect which increases in absolute size and significance for higher income groups. The core findings for institutions, namely strongly positive effects that increase as we go from bottom to top, are also reflected in the results. However, the Arellano-Bond test for

---

[29] Note that a closely related model based on Arellano-Bond (1991), also known as "difference GMM", is inappropriate in this context since differencing strictly eliminates the (fixed) geography variable.

autocorrelation of order two points at potential endogeneity problems, especially for higher income groups.

When using lags of the regressors as instruments in our GMM option two, we can only confirm the relative coefficient movements per fundamental development variable across income groups. However, the overall model appears less robust (see appendix, table 2.30). Significance levels vary, and the absolute coefficient values appear less meaningful. In addition, the autocorrelation test of order one suggests that the lags are not as strong in explaining the contemporaneous variables as the first GMM option shown here. Hence, we prefer our model with explicit instrumental variables, although both GMM specifications broadly confirm the relative coefficient dynamics from our previous cross-section estimates.

We then address the concern that pooling together income groups from countries that otherwise show large differences might affect our estimates. For instance, our sample for 2005 contains the bottom income quintile of Norway and Ethiopia. In absolute terms though, Norway's poorest 20% earn 100 times more than the respective Ethiopian income group, and even 20 times more than the Ethiopian top quintile. We hypothesize that the coefficient dynamics of the three core fundamental factors impact all economies similarly. Nevertheless, we split the countries in a low-income and high-income sub-sample to create a more homogenous income level per sample.[30] Due to the reduced sample size, we limit our analysis to the cross-sections of 1995, 2005, as well as a panel analysis of 1990 to 2010 (see table 2.31 in the appendix). Results indicate that our previous findings also hold in these sub-samples. Compared to low-income countries the model appears to be relatively better suited for high-income countries, and the panel estimates seem most reliable as they contain most data points. The sample containing the upper income half behaves very closely to the full sample, and displays similarly high explanatory power and significance levels. Interestingly, trade is now not always negative but, similar to geography, switches signs as we move from poor to rich.

A further discussion point pertains to conflicting priorities regarding the number of explanatory variables in cross-country growth regressions. While too many variables result in fragile results due to the naturally limited sample size, too few variables attract criticism for being

[30] We also experimented with a split in three or more sub-samples. However, due to the resulting highly limited sample size, estimates become biased and can no longer be meaningfully interpreted.

Table 2.7: Determinants of income: Dynamic panel-data estimation, one step system GMM

1985-2010 (6 periods): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	32	32	32	32	32
Geography	-0.29	-1.86	-2.58	-3.58	-3.83
(GEO)	(1.57)	(1.67)	(1.70)	(1.83)**	(1.85)**
Trade	0.12	-0.08	-0.15	-0.26	-0.30
(LN_TRADE_WB)	(0.29)	(0.21)	(0.22)	(0.24)	(0.24)
Institutions	1.22	1.43	1.46	1.53	1.52
(Inst_rule_of_law)	(0.32)***	(0.36)***	(0.37)***	(0.40)***	(0.41)***
Arellano-Bond test for AR (1) (p-value)	0.03	0.01	0.01	0.01	0.01
Arellano-Bond test for AR (2) (p-value)	0.31	0.02	0.02	0.02	0.02
Hansen Test (p-value)	0.47	0.46	0.34	0.30	0.28

1990-2010 (5 periods): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	43	43	43	43	43
Geography	2.14	0.49	-0.24	-1.18	-1.44
(GEO)	(1.36)	(1.64)	(1.77)	(1.98)	(2.02)
Trade	-0.14	-0.34	-0.45	-0.59	-0.64
(LN_TRADE_WB)	(0.26)	(0.28)	(0.31)	(0.34)*	(0.35)*
Institutions	0.88	1.10	1.14	1.21	1.21
(Inst_rule_of_law)	(0.31)***	(0.38)***	(0.41)***	(0.47)***	(0.48)***
Arellano-Bond test for AR (1) (p-value)	0.94	0.01	<0.001	0.01	0.01
Arellano-Bond test for AR (2) (p-value)	0.40	0.03	0.01	0.02	0.01
Hansen Test (p-value)	0.25	0.25	0.21	0.21	0.21

1995-2010 (4 periods): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	55	55	55	55	55
Geography	1.23	-0.37	-1.08	-2.01	-2.36
(GEO)	(1.53)	(1.83)	(1.89)	(2.07)	(2.14)
Trade	-0.26	-0.48	-0.56	-0.70	-0.75
(LN_TRADE_WB)	(0.31)	(0.34)	(0.35)*	(0.37)*	(0.39)**
Institutions	1.24	1.47	1.49	1.57	1.60
(Inst_rule_of_law)	(0.37)***	(0.44)***	(0.45)***	(0.49)***	(0.51)***
Arellano-Bond test for AR (1) (p-value)	0.01	0.01	0.01	0.01	0.01
Arellano-Bond test for AR (2) (p-value)	0.33	0.01	0.01	0.28	0.86
Hansen Test (p-value)	0.06	0.10	0.12	0.16	0.17

Notes: The dependent variable is per capita GDP on PPP basis. For each specification, there are five samples for which the two-step dynamic panel-data estimations are run: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. Three panel specifications are analyzed: 1985-2010 (six time periods), 1990-2010 (five time periods), and 1995-2010 (four time periods). The model used, known as "system GMM", is based on Arellano and Bover (1995) and Blundell and Bond (1998). The regressors are: (i) GEO, the variable for geography, which is measured as distance from equator; (ii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. The Arellano-Bond tests for autocorrelation and is applied to the differenced residuals. The Hansen Test for over-identifying restrictions follows the standard methodology.

“randomly” selected, and for yielding findings that are non-robust to variable selection (Levine & Renelt, 1992). Ciccone and Jarociński (2010) report that a Bayesian model averaging also offers little help. While it allows for a larger number of regressors, the results are very sensitive to minor measurement errors. This dilemma hence remains so far unsolved and a valid point of criticism which also applies to this work.

We decide to follow standard literature by employing a parsimonious model which includes the variables considered most fundamental for development. The complex causalities and interdependencies of “only” three variables already require careful analyses and interpretations, in particular when examining if their impact differs depending on income groups. Since economic development will never be fully explainable with econometric modeling, a limited set of variables at least allows to get the fundamentals right. The choice of instruments included in the model follows the most widely accepted variables, recognizing that a final consensus on the “right” instruments for trade or institutions is yet to be reached.

Moreover, cross-country growth regressions per se are subject to substantial criticism, which both refers to methodological issues and comparability of data (Atkinson & Brandolini, 2001; Solow, 1986; Mankiw, Phelps, & Romer, 1995). Proponents of randomized experiments (Banerjee & Duflo, 2009) and country-specific “growth diagnostics” (Hausmann, Rodrik, & Velasco, 2006) even tend to regard cross-country growth regressions as generally uninformative. This ongoing controversy has been fueled by additional recent criticism. Eberhardt and Teal (2011) list as common pitfalls cross-section correlation or dependence, which standard empirical models do not take into account, as well as non-stationarity of at least some of the data. Acemoglu (2010) condemns the widespread use of instruments without theory, and Deaton (2010, p.425) argues similarly that there is “a good deal of misunderstanding in the literature about the use of instrumental variables”, and is skeptical whether instruments actually contribute to more credibility in applied econometrics. In particular, Deaton points at the key difference of an instrument being exogenous or merely external, and at the fact that the commonly observed heterogeneity is not a technical problem but a serious symptom of some deeper economic reason. However, his call for an even stronger link of empirics to theoretical mechanisms is difficult to implement here. There is little theoretical guidance on expected income group-specific economic effects, but by using instruments that have passed the most rigorous reviews and reflect a clear underlying theoretical model, we are confident to minimize instrument “misunderstandings”.

Bazzi and Clemens (2013) criticize the simultaneous use of instruments for several endogenous variables. We recognize this point, and indeed singular elements of the constructed trade instrument, such as population, are also “recycled” for other instruments. This would violate the exclusion restriction in case the other studies can argue convincingly that in their context the instrumental variable is more valid. We believe the sound theoretical background of the gravity model together with the high observed correlation and explanatory power of constructed trade shares for actual trade nevertheless justifies the continued use of the instrument.[31] Also, Bazzi and Clemens acknowledge that “new users of [an] instrument bear the burden of showing that other important findings using that instrument do not invalidate its use in the new case” (2013, p. 181): an advantage for the established instruments used in this research. Finally, the authors recommend the extensive use of tests for probing validity of the respective specification. We incorporated this advice through a broad set of tests accompanying the empirical results. With this battery of statistical evidence, we hope to report valid and robust results. After all, the main contribution of this work lies less in its methodological innovativeness; but rather in a careful application of established empirical standards for exploring whether fundamental development factors impact income groups differently.

II. 6. CONCLUDING REMARKS

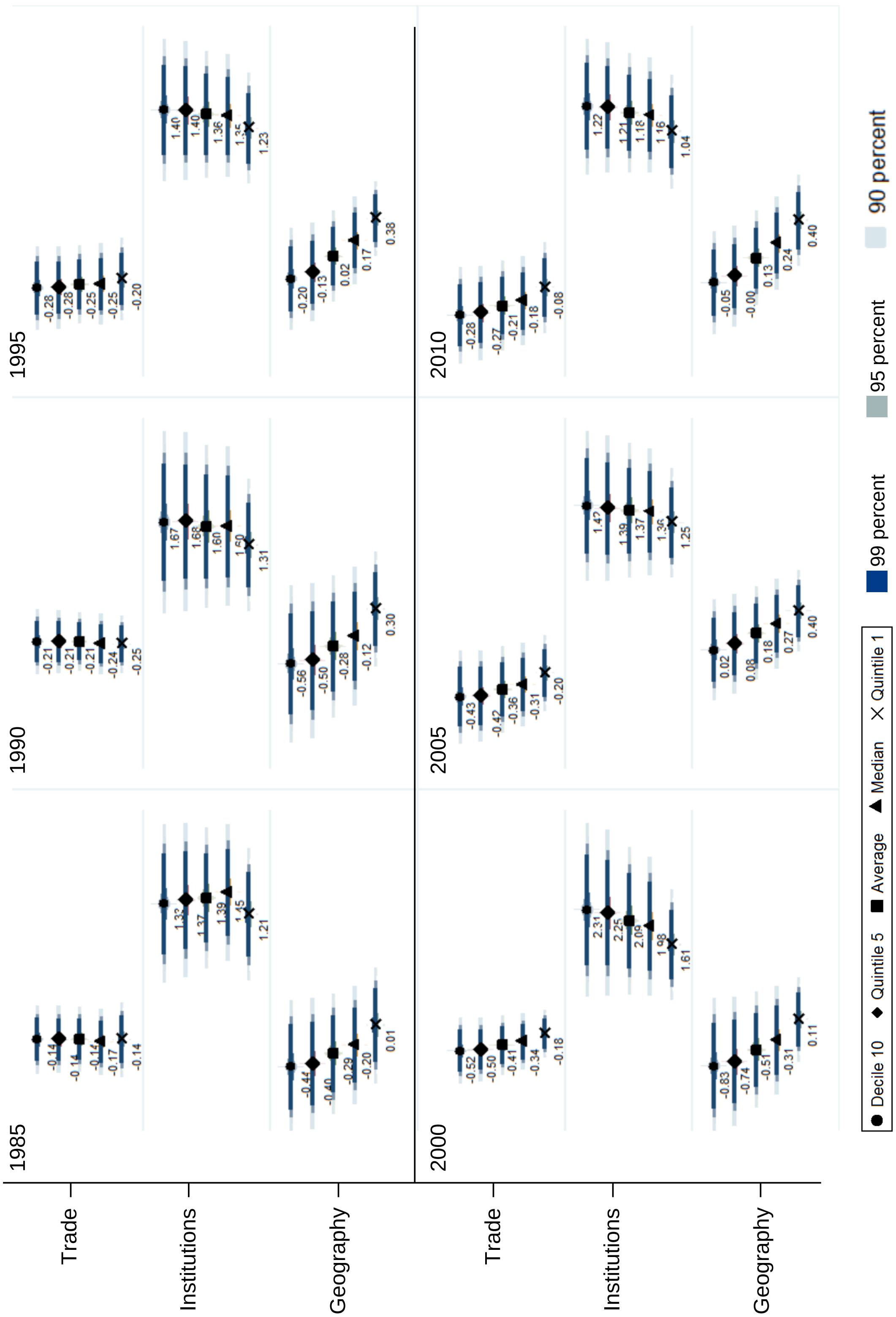
While there exists a large literature which aims to identify causal factors for economic development and which analyzes the link between growth and inequality, this literature is relatively mute on how development factors affect different income groups within developing and developed countries. Thus, there remains substantial ambiguity pertaining to whether different development factors, that change average income levels, actually reach all strata of society equally. It could well be that only specific income groups benefit, respectively suffer, from certain geographic conditions, from changes in trade integration or from institutional improvements. This chapter shed new light on how geography, trade and institutions causally affect different income deciles. Thereby, we offered an answer to the question whether the established fundamental development factors in the literature affect lower income groups differently than higher income

[31] These statistical findings are reported in the original paper by Frankel and Romer (1999), but can be also confirmed in this research as demonstrated before.

groups. Based on the recognized econometric methodology, we analyzed a newly constructed dataset of 138 countries.

The systematic analysis of five income groups over six time periods, covering in total 30 years, yields a number of interesting results: Favorable geographic conditions show an important difference between poor and rich. We observe a consistent pattern of decreasing coefficients, and geography turns even negative for high income groups. The point estimates for the top incomes lie outside the 95 percent confidence interval of the poorest income in most cases. However, the results for geography are usually not statistically significant. In an alternative specification where geography is measured by the prevalence of malaria, the same pattern holds, but the point estimates turn statistically significant. Similar to Rodrik et al. (2004), we observe a negative effect of trade integration for all income levels, but rich incomes display higher absolute coefficient values and significance levels than poor incomes. We interpret this as an equalizing effect of trade different income levels within a country. Institutional quality, on the other hand, affects all income groups positively and at high significance levels, but the coefficient for high incomes is approximately 20 percent higher than the coefficient for low incomes. Coefficient trends move evenly across income groups so that results for median and average income groups are close to a linear interpolation of top and bottom incomes. Our results are consistent over time but the explanatory power of the empirical analysis increases for lower incomes.

We corroborated the findings through a large number of robustness tests. These indicate that world regions, lagged income values and overall inequality in a country have a sizeable effect, though the latter two suffer from likely endogeneity issues. The control variables health and human capital, the latter instrumented with lags, also enter significantly, but none of the controls alter the described relative effects of the fundamental development factors. Specifically, results do not suggest that human capital is a more basic source for growth than institutions. We also documented the model's overall validity through a set of additional econometric tests. The transformation of our econometric findings into what might be called "pro-poor" or "inclusive development policies" is a formidable challenge to which this study might only serve as a reference. Nevertheless, the evidence for the adverse role of geographic conditions for the poor in the form of a disproportional disease burden, the equalizing effect of trade, and the relatively higher influence of institutional quality on high income groups may serve as fresh input for development policy discussions and further research.



Appendix Figure 2.3: Confidence intervals for all time periods per standardized variable of the fundamental development factors (trade, institutions, and geography), each broken down to the respective income group. (Standardized) regression coefficients are labeled with their estimated values.

Appendix Table 2.8: Determinants of income: Specifications using settler mortality instrument

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	30	30	30	30	30
Geography	-2.64	-1.78	-1.71	-1.77	-1.68
(GEO)	(2.12)	(1.77)	(1.60)	(1.61)	(1.55)
Trade	-0.21	-0.09	0.03	0.11	0.13
(LN_TRADE_WB)	(0.45)	(0.42)	(0.41)	(0.43)	(0.43)
Institutions	1.63	1.46	1.30	1.20	1.10
(Inst_rule_of_law)	(0.49)***	(0.36)***	(0.31)***	(0.30)***	(0.30)***
R-Square	0.19	0.32	0.29	0.24	0.23
Pagan Hall test (p-value)	0.74	0.82	0.44	0.28	0.22
Endogeneity test (p-value)	0.08	0.05	0.05	0.08	0.11

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	43	43	43	43	43
Geography	-4.77	-5.69	-5.68	-6.07	-6.17
(GEO)	(3.86)	(3.89)	(3.70)	(3.77)	(3.78)*
Trade	-0.18	0.07	0.20	0.31	0.35
(LN_TRADE_WB)	(0.62)	(0.63)	(0.60)	(0.62)	(0.62)
Institutions	2.74	2.76	2.57	2.51	2.47
(Inst_rule_of_law)	(1.02)***	(1.01)***	(0.96)***	(0.98)***	(0.98)***
R-Square	<0.001	<0.001	<0.001	<0.001	<0.001
Pagan Hall test (p-value)	0.96	0.93	0.88	0.83	0.82
Endogeneity test (p-value)	0.02	0.01	0.01	0.01	0.01

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	49	49	49	49	49
Geography	-9.28	-8.01	-7.21	-7.25	-7.07
(GEO)	(6.19)	(4.89)*	(4.06)*	(3.83)*	(3.65)**
Trade	-0.20	-0.10	-0.04	0.01	0.05
(LN_TRADE_WB)	(0.98)	(0.81)	(0.69)	(0.66)	(0.63)
Institutions	3.86	3.42	3.08	2.98	2.87
(Inst_rule_of_law)	(1.61)**	(1.26)***	(1.03)***	(0.97)***	(0.92)***
R-Square	<0.001	<0.001	<0.001	<0.001	<0.001
Pagan Hall test (p-value)	0.57	0.76	0.90	0.93	0.93
Endogeneity test (p-value)	0.08	0.03	0.02	0.01	0.01

Appendix Table 2.8 continued: Determinants of income: Specifications using settler mortality instrument

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	48	48	48	48	48
Geography	-3.27	-4.76	-5.65	-6.52	-7.25
(GEO)	(4.52)	(4.99)	(5.18)	(5.44)	(5.72)
Trade	-0.23	-0.23	-0.26	-0.26	-0.22
(LN_TRADE_WB)	(0.54)	(0.61)	(0.65)	(0.70)	(0.74)
Institutions	2.62	2.86	2.99	3.14	3.30
(Inst_rule_of_law)	(1.28)**	(1.40)**	(1.46)**	(1.54)**	(1.63)**
R-Square	<0.001	<0.001	<0.001	<0.001	<0.001
Pagan Hall test (p-value)	0.98	0.97	0.92	0.86	0.87
Endogeneity test (p-value)	0.05	0.02	0.02	0.02	0.02

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	55	55	55	55	55
Geography	-2.68	-3.47	-3.82	-4.08	-4.17
(GEO)	(2.30)	(2.71)	(2.91)	(3.10)	(3.18)
Trade	-0.48	-0.61	-0.68	-0.74	-0.74
(LN_TRADE_WB)	(0.48)	(0.54)	(0.58)	(0.62)	(0.63)
Institutions	2.12	2.35	2.44	2.48	2.50
(Inst_rule_of_law)	(0.60)***	(0.70)***	(0.76)***	(0.81)***	(0.83)***
R-Square	0.17	<0.001	<0.001	<0.001	<0.001
Pagan Hall test (p-value)	0.71	0.70	0.75	0.76	0.74
Endogeneity test (p-value)	0.04	0.01	0.01	0.01	0.01

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	41	41	41	41	41
Geography	-3.18	-4.43	-4.95	-5.43	-5.57
(GEO)	(3.10)	(3.41)	(3.58)	(3.79)	(3.87)
Trade	-0.63	-0.71	-0.74	-0.74	-0.74
(LN_TRADE_WB)	(0.57)	(0.66)	(0.70)	(0.74)	(0.77)
Institutions	1.91	2.22	2.31	2.37	2.39
(Inst_rule_of_law)	(0.65)***	(0.76)***	(0.82)***	(0.90)***	(0.93)***
R-Square	0.09	<0.001	<0.001	<0.001	<0.001
Pagan Hall test (p-value)	0.62	0.59	0.63	0.65	0.67
Endogeneity test (p-value)	0.14	0.04	0.04	0.04	0.04

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) GEO, the variable for geography, which is measured as the absolute value of latitude of country divided by 90; (ii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented with settler mortality rates following Acemoglu et al. (2001). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.The Pagan Hall tests of heteroskedasticity for instrumental variables (IV) estimation under the null of homoskedasticity. The endogeneity test is based on the Durbin-Wu-Hausman test, but adjusted here for heteroskedasticity.

Appendix Table 2.9: First stage estimates of two-stages least square estimates using settler mortality as instrument

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	30	30	43	43	49	49	48	48	55	55	41	41
Geography	0.62	2.25	0.37	2.78	0.39	3.16	0.53	2.94	0.33	2.62	0.26	2.93
(GEO)	(0.52)	(0.69)***	(0.46)	(0.61)***	(0.53)	(0.59)***	(0.52)	(0.63)***	(0.39)	(0.78)***	(0.49)	(1.02)***
Constructed Trade	0.58	-0.11	0.58	-0.07	0.48	0.02	0.46	0.05	0.63	0.24	0.64	0.27
(Trade_FR_ROM)	(0.12)***	(0.15)	(0.11)***	(0.14)	(0.09)***	(0.14)	(0.11)***	(0.13)	(0.09)***	(0.14)*	(0.12)***	(0.21)
Settler Mortality	-0.06	-0.33	-0.03	-0.20	-0.04	-0.17	-0.03	-0.14	-0.15	-0.28	-0.13	-0.32
(Inst_sett_mort)	(0.08)	(0.10)***	(0.06)	(0.09)**	(0.06)	(0.07)**	(0.08)	(0.10)	(0.06)**	(0.10)***	(0.07)*	(0.14)**
First-stage F-test	11.6	6.4	27.0	2.6	14.4	2.8	13.5	1.4	27.7	4.1	13.8	3.1
Angrist-Pischke F-statistics (p-value)	<0.001	<0.001	<0.001	0.03	<0.001	0.02	<0.001	0.11	<0.001	0.02	<0.001	0.06
Kleibergen-Paap LM test (p-value)	0.04		0.06		0.03		0.12		0.01		0.04	
Kleibergen-Paap Wald rk F statistic	4.41		2.34		2.85		1.25		3.34		2.36	
Stock-Yogo critical values 10%							7.03					
Stock-Yogo critical values 25%							3.63					
R-Square	0.43	0.57	0.47	0.52	0.37	0.49	0.34	0.39	0.54	0.43	0.53	0.40

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) GEO, the variable for geography, which is measured as the absolute value of latitude of country divided by 90; (ii) constructed trade, the instrument for trade obtained from Frankel and Romer; and (iii) settler mortality in the country following Acemoglu et al. (2001). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are repoted in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.10: Determinants of income: Specifications, using alternative geography variable Mean Temperature

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	54	54	54	54	54
Geography	-0.05	-0.02	-0.01	-0.01	-0.01
(Mean Temperature)	(0.03)	(0.02)	(0.02)	(0.02)	0.02
Trade	-0.24	-0.27	-0.23	-0.24	-0.25
(LN_TRADE_WB)	(0.25)	(0.22)	(0.21)	(0.23)	(0.23)
Institutions	0.80	1.02	0.93	0.86	0.81
(Inst_rule_of_law)	(0.27)***	(0.17)***	(0.16)	(0.16)***	(0.16)***
R-Square	0.68	0.71	0.67	0.56	0.50
Hansen Test (p-value)	0.04	0.02	0.01	0.01	0.01

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	67	67	67	67	67
Geography	-0.04	-0.01	0.01	0.02	0.03
(Mean Temperature)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Trade	-0.38	-0.32	-0.26	-0.25	-0.24
(LN_TRADE_WB)	(0.31)	(0.29)	(0.28)	(0.30)	(0.30)
Institutions	1.16	1.32	1.27	1.28	1.25
(Inst_rule_of_law)	(0.32)***	(0.33)***	(0.32)***	(0.34)***	(0.34)***
R-Square	0.72	0.68	0.61	0.47	0.41
Hansen Test (p-value)	0.08	0.01	0.01	<0.001	<0.001

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	86	86	86	86	86
Geography	-0.04	-0.01	0.01	0.03	0.03
(Mean Temperature)	(0.02)**	(0.03)	(0.03)	(0.03)	(0.03)
Trade	-0.47	-0.42	-0.38	-0.37	-0.36
(LN_TRADE_WB)	(0.28)*	(0.28)	(0.28)	(0.31)	(0.33)
Institutions	1.23	1.45	1.48	1.55	1.56
(Inst_rule_of_law)	(0.26)***	(0.33)***	(0.37)***	(0.42)***	(0.44)***
R-Square	0.69	0.65	0.57	0.38	0.29
Hansen Test (p-value)	0.10	0.06	0.07	0.07	0.08

Appendix Table 2.10 continued: Determinants of income: Specifications, using geography variable Mean Temperature

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	76	76	76	76	76
Geography	0.01	0.05	0.08	0.10	0.11
(Mean Temperature)	(0.04)	(0.05)	(0.06)	(0.06)	(0.07)
Trade	-0.29	-0.67	-0.88	-1.11	-1.17
(LN_TRADE_WB)	(0.41)	(0.51)	(0.55)*	(0.60)*	(0.63)*
Institutions	1.74	2.12	2.19	2.32	2.33
(Inst_rule_of_law)	(0.41)***	(0.58)***	(0.65)***	(0.74)***	(0.78)***
R-Square	0.69	0.46	0.30	0.01	<0.001
Hansen Test (p-value)	0.19	0.26	0.27	0.29	0.27

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	91	91	91	91	91
Geography	-0.01	0.01	0.02	0.03	0.03
(Mean Temperature)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Trade	-0.26	-0.48	-0.60	-0.75	-0.78
(LN_TRADE_WB)	(0.27)	(0.31)	(0.33)*	(0.37)**	(0.38)**
Institutions	1.44	1.54	1.54	1.56	1.55
(Inst_rule_of_law)	(0.28)	(0.34)***	(0.38)***	(0.43)***	(0.44)**
R-Square	0.69	0.59	0.53	0.41	0.36
Hansen Test (p-value)	0.13	0.05	0.03	0.02	0.02

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	69	69	69	69	69
Geography	-0.05	-0.03	-0.02	-0.01	0.01
(Mean Temperature)	(0.02)***	(0.02)	(0.02)	(0.02)	(0.03)
Trade	-0.15	-0.31	-0.37	-0.46	-0.49
(LN_TRADE_WB)	(0.22)	(0.21)	(0.21)*	(0.22)**	(0.23)**
Institutions	0.89	1.04	1.04	1.07	1.07
(Inst_rule_of_law)	(0.19)***	(0.17)***	(0.18)***	(0.21)***	(0.23)***
R-Square	0.74	0.68	0.64	0.55	0.51
Hansen Test (p-value)	0.09	0.03	0.01	0.01	0.01

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) GEO, the variable for geography, which is measured via the Mean Temperature (CID Harvard University, 2002) ; (ii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.11: First stage estimates of two-stages least square estimates using alternative geography variable Mean Temperature

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	54	54	67	67	86	86	76	76	91	91	69	69
Geography	-0.01	-0.08	0.01	-0.09	0.01	-0.06	-0.01	-0.09	<0.001	-0.09	<0.001	-0.08
(Mean Temperature)	(0.01)	(0.01)***	(0.01)	(0.01)***	(0.01)	(0.02)***	(0.01)	(0.01)***	(0.01)	(0.01)***	(0.01)	(0.01)***
Constructed Trade	0.57	0.20	0.54	0.11	0.47	0.08	0.42	0.25	0.48	0.30	0.55	0.36
(Trade_FR_ROM)	(0.07)***	(0.12)*	(0.07)***	(0.12)	(0.07)***	(0.14)	(0.08)***	(0.11)**	(0.07)***	(0.11)***	(0.09)***	(0.14)***
Pop. speaking English	0.42	1.15	0.39	0.89	0.41	0.76	0.27	0.54	0.17	0.94	0.33	1.39
(Eng_Lang)	(0.15)***	(0.34)***	(0.14)***	(0.36)**	(0.16)***	(0.35)**	(0.21)	(0.29)**	(0.14)	(0.43)**	(0.18)*	(0.43)***
Pop. speaking other European languages	-0.18	-0.07	-0.10	0.03	-0.09	0.30	-0.14	0.26	-0.07	0.19	-0.17	0.09
(EUR_Lang)	(0.11)	(0.20)	(0.11)	(0.20)	(0.10)	(0.22)	(0.10)	(0.19)	(0.09)	(0.18)	(0.10)	(0.21)
First-stage F-test	21.5	4.2	32.3	2.4	17.9	3.2	9.8	3.8	16.6	3.9	16.5	5.0
Angrist-Pischke F-statistics (p-value)	<0.001	0.01	<0.001	0.04	<0.001	0.01	<0.001	0.03	<0.001	0.03	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	0.14		0.14		0.06		0.08		0.08		0.10	
Kleibergen-Paap Wald rk F statistic	3.79		2.25		3.00		2.02		2.27		3.99	
Stock-Yogo critical values 10%	13.43											
Stock-Yogo critical values 25%	5.45											
R-Square	0.58	0.69	0.54	0.62	0.42	0.40	0.35	0.60	0.43	0.55	0.49	0.61

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) GEO, the variable for geography, which is measured via the Mean Temperature (CID Harvard University, 2002) (ii) constructed trade, theinstrument for trade obtained from Frankel and Romer; (iii) the proportion of the population of a country that speaks English (Eng_Lang); and (iv) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is underidentified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993)to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.12: Determinants of income: Specifications, using alternative geography variable Malaria

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	56	56	56	56	56
Geography	-0.75	-0.74	-0.72	-0.67	-0.63
(Malaria)	(0.41)*	(0.38)**	(0.34)**	(0.33)**	(0.33)*
Trade	-0.10	-0.21	-0.19	-0.23	-0.23
(LN_TRADE_WB)	(0.24)	(0.19)	(0.18)	(0.18)	(0.18)
Institutions	0.98	1.07	0.93	0.83	0.77
(Inst_rule_of_law)	(0.16)***	(0.12)***	(0.11)***	(0.11)***	(0.11)***
R-Square	0.62	0.71	0.69	0.61	0.56
Hansen Test (p-value)	0.08	0.04	0.02	0.01	0.01

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	71	71	71	71	71
Geography	-1.84	-1.46	-1.13	-0.90	-0.82
(Malaria)	(0.40)***	(0.32)***	(0.31)***	(0.34)***	(0.34)**
Trade	-0.08	-0.16	-0.15	-0.20	-0.20
(LN_TRADE_WB)	(0.27)	(0.23)	(0.22)	(0.22)	(0.23)
Institutions	0.93	0.93	0.85	0.78	0.74
(Inst_rule_of_law)	(0.17)***	(0.15)***	(0.14)***	(0.15)***	(0.15)***
R-Square	0.79	0.80	0.75	0.66	0.61
Hansen Test (p-value)	0.20	0.83	0.30	0.07	0.04

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	107	107	107	107	107
Geography	-1.79	-1.48	-1.14	-0.89	-0.75
(Malaria)	(0.42)***	(0.31)***	(0.29)***	(0.34)***	(0.38)**
Trade	0.08	-0.12	-0.23	-0.38	-0.43
(LN_TRADE_WB)	(0.27)	(0.20)	(0.19)	(0.22)*	(0.25)*
Institutions	0.66	0.84	0.94	1.04	1.08
(Inst_rule_of_law)	(0.30)**	(0.17)***	(0.16)***	(0.23)***	(0.28)***
R-Square	0.67	0.75	0.74	0.63	0.55
Hansen Test (p-value)	0.60	0.89	0.48	0.29	0.24

Appendix Table 2.12 continued: Determinants of income: Specifications, using geography variable Malaria

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	84	84	84	84	84
Geography	-1.35	-1.13	-1.03	-0.92	-0.96
(Malaria)	(0.59)**	(0.37)***	(0.40)***	(0.54)*	(0.56)*
Trade	0.25	-0.14	-0.29	-0.49	-0.50
(LN_TRADE_WB)	(0.40)	(0.24)	(0.26)	(0.36)	(0.37)
Institutions	0.93	1.12	1.11	1.14	1.10
(Inst_rule_of_law)	(0.44)**	(0.20)***	(0.25)***	(0.40)***	(0.41)***
R-Square	0.78	0.80	0.77	0.68	0.65
Hansen Test (p-value)	0.43	0.98	0.55	0.29	0.22

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	117	117	117	117	117
Geography	-1.38	-1.48	-1.46	-1.41	-1.37
(Malaria)	(0.26)***	(0.25)***	(0.25)***	(0.25)***	(0.26)***
Trade	0.02	-0.18	-0.28	-0.41	-0.45
(LN_TRADE_WB)	(0.20)	(0.19)	(0.18)	(0.19)**	(0.20)**
Institutions	0.94	0.94	0.92	0.91	0.93
(Inst_rule_of_law)	(0.14)***	(0.13)***	(0.13)***	(0.14)***	(0.15)***
R-Square	0.75	0.76	0.76	0.72	0.69
Hansen Test (p-value)	0.54	0.88	0.56	0.31	0.21

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	91	91	91	91	91
Geography	-1.82	-1.76	-1.65	-1.50	-1.42
(Malaria)	(0.29)***	(0.28)***	(0.28)***	(0.30)***	(0.32)***
Trade	0.27	0.02	-0.10	-0.26	-0.32
(LN_TRADE_WB)	(0.17)	(0.15)	(0.14)	(0.16)*	(0.17)*
Institutions	0.66	0.74	0.75	0.78	0.79
(Inst_rule_of_law)	(0.13)***	(0.09)***	(0.08)***	(0.10)***	(0.12)***
R-Square	0.74	0.78	0.78	0.73	0.69
Hansen Test (p-value)	0.62	0.98	0.65	0.29	0.22

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) GEO, the variable for geography, which is measured via the Malaria Index 1994 by Gallup and Sachs (1994); (ii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.13: First stage estimates of two-stages least square estimates using alternative geography variable Malaria

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	56	56	71	71	107	107	84	84	117	117	91	91
Geography	-0.14	-1.00	-0.04	-1.24	-0.40	-1.06	-0.29	-1.33	-0.34	-0.96	-0.37	-1.19
(Malaria)	(0.25)	(0.28)***	(0.12)	(0.24)***	(0.11)***	(0.18)***	(0.14)**	(0.20)***	(0.10)***	(0.17)***	(0.15)***	(0.21)***
Constructed Trade	0.53	0.41	0.48	0.31	0.41	0.25	0.40	0.31	0.40	0.43	0.44	0.46
(Trade_FR_ROM)	(0.07)***	(0.11)***	(0.06)***	(0.13)***	(0.05)***	(0.11)**	(0.05)***	(0.11)***	(0.05)***	(0.09)***	(0.05)***	(0.11)***
Pop. speaking English	0.42	1.87	0.32	1.40	0.33	1.03	0.29	0.87	0.12	1.37	0.14	1.61
(Eng_Lang)	(0.15)***	(0.33)***	(0.15)**	(0.58)**	(0.16)**	(0.54)*	(0.20)	(0.63)	(0.14)	(0.61)**	(0.14)	(0.49)***
Pop. speaking other European languages	-0.24	-0.19	-0.15	-0.24	-0.34	0.10	-0.26	-0.18	-0.27	0.15	-0.35	-0.04
(EUR_Lang)	(0.13)*	(0.36)	(0.12)	(0.32)	(0.10)***	(0.27)	(0.10)***	(0.29)	(0.08)***	(0.28)	(0.09)***	(0.30)
First-stage F-test	30.1	17.0	31.5	3.5	32.4	3.0	21.5	3.5	24.5	7.0	34.6	8.2
Angrist-Pischke F-statistics (p-value)	<0.001	<0.001	<0.001	0.11	<0.001	0.06	<0.001	0.53	<0.001	0.02	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	0.04		0.24		0.12		0.60		0.05		0.03	
Kleibergen-Paap Wald rk F statistic	12.60		1.39		1.76		0.40		2.57		6.07	
Stock-Yogo critical values 10%	13.43											
Stock-Yogo critical values 25%	5.45											
R-Square	0.60	0.42	0.54	0.35	0.46	0.32	0.44	0.39	0.42	0.34	0.52	0.41

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) GEO, the variable for geography, which is measured via the Malaria Index 1994 by Gallup and Sachs (1994); (ii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iii) the proportion of the population of a country that speaks English (Eng_Lang); and (iv) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified.The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.14: Determinants of income: Control variable European Colony

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	56	56	56	56	56
European Colony	-0.32	-0.10	-0.04	0.03	0.07
(EUR_colony)	(0.34)	(0.31)	(0.28)	(0.28)	(0.28)
Geography	-0.82	-1.31	-1.61	-1.97	-2.05
(GEO)	(1.54)	(1.40)	1.35	(1.37)	(1.38)
Trade	-0.35	-0.31	-0.24	-0.22	-0.20
(LN_TRADE_WB)	(0.31)	(0.26)	(0.24)	(0.25)	(0.25)
Institutions	1.34	1.51	1.42	1.37	1.32
(Inst_rule_of_law)	(0.30)***	(0.28)***	(0.28)***	(0.30)***	(0.30)***
R-Square	0.60	0.60	0.53	0.42	0.36
Hansen Test (p-value)	0.02	0.03	0.02	0.01	0.01

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	71	71	71	71	71
European Colony	-0.58	-0.43	-0.31	-0.20	-0.15
(EUR_colony)	(0.29)**	(0.33)	(0.31)	(0.33)	(0.33)
Geography	-0.29	-1.97	-2.39	-3.15	-3.27
(GEO)	(1.96)	(2.12)	(2.04)	(2.14)	(2.14)
Trade	-0.69	-0.62	-0.51	-0.47	-0.45
(LN_TRADE_WB)	(0.41)*	(0.43)	(0.42)	(0.44)	(0.44)
Institutions	1.52	1.75	1.69	1.72	1.69
(Inst_rule_of_law)	(0.44)***	(0.49)***	(0.48)***	(0.50)***	(0.51)***
R-Square	0.64	0.48	0.38	0.19	0.12
Hansen Test (p-value)	0.21	0.15	0.12	0.10	0.09

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	107	107	107	107	107
European Colony	-0.54	-0.38	-0.23	-0.10	-0.07
(EUR_colony)	(0.28)**	(0.28)	(0.26)	(0.26)	(0.26)
Geography	0.54	-0.14	-0.52	-0.97	-1.21
(GEO)	(1.07)	(1.11)	(1.10)	(1.14)	(1.16)
Trade	-0.48	-0.53	-0.51	-0.54	-0.54
(LN_TRADE_WB)	(0.31)	(0.31)*	(0.29)*	(0.30)*	(0.30)*
Institutions	1.44	1.52	1.49	1.49	1.49
(Inst_rule_of_law)	(0.24)***	(0.25)***	(0.25)***	(0.27)***	(0.28)***
R-Square	0.62	0.53	0.47	0.32	0.25
Hansen Test (p-value)	0.12	0.19	0.24	0.30	0.33

Appendix Table 2.14 continued: Determinants of income: Control variable European Colony

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	84	84	84	84	84
European Colony	-0.62	-0.45	-0.28	-0.10	-0.06
(EUR_colony)	(0.30)**	(0.36)	(0.37)	(0.39)	(0.41)
Geography	-1.28	-2.93	-3.43	-4.12	-4.40
(GEO)	(1.86)	(2.24)	(2.31)	(2.49)*	(2.59)*
Trade	-0.55	-0.80	-0.88	-0.99	-1.02
(LN_TRADE_WB)	(0.49)	(0.46)*	(0.47)*	(0.51)**	(0.52)**
Institutions	1.81	2.13	2.18	2.27	2.32
(Inst_rule_of_law)	(0.46)***	(0.56)***	(0.58)***	(0.63)***	(0.65)***
R-Square	0.64	0.41	0.29	0.08	0.01
Hansen Test (p-value)	0.61	0.78	0.73	0.73	0.70

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	117	117	117	117	117
European Colony	-0.93	-0.94	-0.70	-0.58	-0.54
(EUR_colony)	(0.29)***	(0.32)***	(0.32)**	(0.34)*	(0.34)
Geography	-0.27	-0.70	-0.80	-1.04	-1.23
(GEO)	(1.22)	(1.34)	(1.35)	(1.41)	(1.42)
Trade	-0.76	-0.94	-0.98	-1.05	-1.05
(LN_TRADE_WB)	(0.33)**	(0.37)***	(0.38)***	(0.40)***	(0.41)***
Institutions	1.49	1.57	1.55	1.54	1.56
(Inst_rule_of_law)	(0.27)***	(0.31)***	(0.31)***	(0.33)***	(0.33)***
R-Square	0.60	0.46	0.39	0.27	0.23
Hansen Test (p-value)	0.29	0.18	0.15	0.14	0.13

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	91	91	91	91	91
European Colony	-0.55	-0.52	-0.45	-0.40	-0.36
(EUR_colony)	(0.36)	(0.33)	(0.31)	(0.31)	(0.31)
Geography	0.76	-0.01	-0.39	-0.93	-1.08
(GEO)	(1.10)	(1.07)	(1.02)	(1.04)	(1.04)
Trade	-0.36	-0.52	-0.56	-0.63	-0.65
(LN_TRADE_WB)	(0.26)	(0.27)*	(0.28)**	(0.29)**	(0.30)**
Institutions	1.12	1.22	1.23	1.25	1.25
(Inst_rule_of_law)	(0.15)***	(0.17)***	(0.18)***	(0.20)***	(0.21)***
R-Square	0.68	0.59	0.54	0.43	0.39
Hansen Test (p-value)	0.07	0.06	0.05	0.05	0.05

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) European Colony, which is a dummy variable whether a country has been colonized by a European country (ii) GEO, the variable for geography, which is measured as distance from equator; (iii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iv) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.15: First stage estimates of two-stages least square estimates using control variable European Colony

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institutions	Trade	Institutions	Trade	Institutions	Trade	Institutions	Trade	Institutions	Trade	Institutions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	56	56	71	71	107	107	84	84	117	117	91	91
European Colony	0.35	-0.25	0.13	-0.16	0.05	-0.10	-0.09	-0.15	0.03	0.18	0.03	0.07
(EUR_colony)	(0.17)**	(0.27)	(0.16)	(0.28)	(0.14)	(0.20)	(0.15)	(0.22)	(0.13)	(0.20)	(0.16)	(0.22)
Geography	0.60	3.06	0.09	3.45	0.60	2.79	0.23	3.53	0.50	3.15	0.34	3.06
(GEO)	(0.38)	(0.58)***	(0.36)	(0.58)***	(0.36)*	(0.55)***	(0.36)	(0.50)***	(0.31)	(0.52)***	(0.35)	(0.54)***
Constructed Trade	0.60	0.14	0.50	0.14	0.42	0.24	0.39	0.22	0.40	0.41	0.46	0.42
(Trade_FR_ROM)	(0.08)***	(0.11)	(0.07)***	(0.13)	(0.05)***	(0.09)***	(0.06)***	(0.10)**	(0.06)***	(0.10)***	(0.06)***	(0.12)***
Pop. speaking English	0.36	0.95	0.31	0.71	0.23	0.49	0.24	0.19	0.05	0.83	0.11	1.17
(Eng_Lang)	(0.16)**	(0.38)**	(0.16)*	(0.43)*	(0.19)	(0.38)	(0.20)	(0.34)	(0.16)	(0.51)*	(0.15)	(0.38)***
Pop. speaking other European languages	-0.25	0.23	-0.15	0.33	-0.19	0.67	-0.12	0.45	-0.15	0.55	-0.23	0.48
(EUR_Lang)	(0.13)*	(0.20)	(0.11)	(0.17)*	(0.11)*	(0.16)***	(0.10)	(0.16)***	(0.09)*	(0.18)***	(0.10)**	(0.20)**
First-stage F-test	20.4	3.4	24.1	2.8	27.2	10.1	16.3	4.8	16.9	8.6	17.8	16.8
Angrist-Pischke F-statistics (p-value)	<0.001	0.01	<0.001	0.02	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	0.02		0.02		<0.001		0.01		0.01		<0.001	
Kleibergen-Paap Wald rk F statistic	3.37		2.64		8.71		3.93		5.71		11.9	
Stock-Yogo critical values 10%							13.43					
Stock-Yogo critical values 25%							5.45					
R-Square	0.63	0.71	0.55	0.66	0.42	0.50	0.42	0.67	0.39	0.51	0.49	0.57

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) European Colony, which is a dummy that takes 1 if the country had been colonized by a European country in the past; (ii) GEO, the variable for geography, which is measured as distance from the equator; (iii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iv) the proportion of the population of a country that speaks English (Eng_Lang); and (v) the proportion of the popultion of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance atthe 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap,2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.16: Determinants of income: Control variable ethnolinguistic fractionalization

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	56	56	56	56	56
Ethnolinguistic fractionalization	-0.12	0.13	0.18	0.23	0.22
(Ethnoling_frac)	(0.42)	(0.39)	(0.36)	(0.38)	(0.38)
Geography	0.26	-0.70	-1.14	-1.67	-1.85
(GEO)	(1.14)	(1.14)	(1.18)	(1.29)	(1.32)
Trade	-0.23	-0.27	-0.23	-0.23	-0.23
(LN_TRADE_WB)	(0.26)	(0.23)	(0.22)	(0.23)	(0.23)
Institutions	1.17	1.42	1.35	1.32	1.28
(Inst_rule_of_law)	(0.27)***	(0.26)***	(0.27)***	(0.29)***	(0.30)***
R-Square	0.61	0.62	0.56	0.45	0.38
Hansen Test (p-value)	0.02	0.02	0.01	0.01	0.01

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	71	71	71	71	71
Ethnolinguistic fractionalization	-0.80	-0.56	-0.33	-0.19	-0.13
(Ethnoling_frac)	(0.39)**	(0.40)	(0.40)	(0.44)	(0.45)
Geography	1.16	-0.92	-1.63	-2.69	-2.94
(GEO)	(1.71)	(2.08)	(2.15)	(2.38)	(2.43)
Trade	-0.42	-0.43	-0.38	-0.39	-0.39
(LN_TRADE_WB)	(0.33)	(0.34)	(0.34)	(0.37)	(0.38)
Institutions	1.24	1.55	1.56	1.65	1.64
(Inst_rule_of_law)	(0.37)***	(0.46)***	(0.48)***	(0.54)***	(0.55)***
R-Square	0.71	0.58	0.46	0.24	0.16
Hansen Test (p-value)	0.16	0.11	0.09	0.08	0.08

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	107	107	107	107	107
Ethnolinguistic fractionalization	-0.55	-0.38	-0.29	-0.18	-0.17
(Ethnoling_frac)	(0.35)	(0.37)	(0.40)	(0.45)	(0.48)
Geography	1.89	0.85	0.08	-0.68	-1.02
(GEO)	(0.64)***	(0.76)	(0.80)	(0.89)	(0.92)
Trade	-0.33	-0.42	-0.43	-0.50	-0.51
(LN_TRADE_WB)	(0.27)	(0.27)	(0.26)*	(0.28)*	(0.28)*
Institutions	1.18	1.33	1.36	1.41	1.42
(Inst_rule_of_law)	(0.21)***	(0.24)***	(0.25)***	(0.28)***	(0.29)***
R-Square	0.67	0.61	0.54	0.38	0.30
Hansen Test (p-value)	0.06	0.11	0.19	0.28	0.32

Appendix Table 2.16 continued: Determinants of income: Control variable ethnolinguistic fractionalization

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	84	84	84	84	84
Ethnolinguistic fractionalization	-0.77	-0.93	-0.87	-0.90	-0.93
(Ethnoling_frac)	(0.42)*	(0.47)**	(0.50)*	(0.56)	(0.58)
Geography	0.30	-1.87	-2.87	-4.09	-4.54
(GEO)	(1.70)	(2.07)	(2.22)	(2.48)*	(2.59)*
Trade	-0.30	-0.58	-0.72	-0.89	-0.93
(LN_TRADE_WB)	(0.33)	(0.38)	(0.40)*	(0.43)	(0.45)**
Institutions	1.50	1.84	1.96	2.12	2.17
(Inst_rule_of_law)	(0.39)***	(0.49)***	(0.54)***	(0.60)***	(0.63)***
R-Square	0.73	0.56	0.43	0.22	0.12
Hansen Test (p-value)	0.43	0.85	0.85	0.92	0.90

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	117	117	117	117	117
Ethnolinguistic fractionalization	-0.30	-0.42	-0.48	-0.52	-0.61
(Ethnoling_frac)	(0.37)	(0.39)	(0.39)	(0.42)	(0.42)
Geography	1.97	1.30	0.83	0.29	-0.02
(GEO)	(0.73)***	(0.80)*	(0.81)	(0.85)	(0.86)
Trade	-0.36	-0.57	-0.66	-0.77	-0.79
(LN_TRADE_WB)	(0.27)	(0.29)**	(0.30)**	(0.32)***	(0.32)**
Institutions	1.19	1.27	1.28	1.29	1.31
(Inst_rule_of_law)	(0.21)***	(0.24)***	(0.25)***	(0.27)***	(0.28)***
R-Square	0.69	0.60	0.55	0.44	0.41
Hansen Test (p-value)	0.15	0.08	0.08	0.08	0.09

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	91	91	91	91	91
Ethnolinguistic fractionalization	-0.80	-0.72	-0.60	-0.49	-0.43
(Ethnoling_frac)	(0.38)**	(0.36)**	(0.37)*	(0.40)	(0.41)
Geography	1.90	1.09	0.58	-0.06	-0.28
(GEO)	(0.60)***	(0.61)*	(0.61)	(0.65)	(0.67)
Trade	-0.11	-0.28	-0.35	-0.45	-0.48
(LN_TRADE_WB)	(0.21)	(0.21)	(0.21)*	(0.23)**	(0.23)**
Institutions	0.88	1.00	1.03	1.08	1.09
(Inst_rule_of_law)	(0.15)***	(0.15)***	(0.16)***	(0.18)***	(0.19)***
R-Square	0.73	0.67	0.62	0.52	0.48
Hansen Test (p-value)	0.09	0.06	0.05	0.04	0.04

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) Ethnolinguistic fractionalization following Alesina et al. (2003) (ii) GEO, the variable for geography, which is measured as distance from equator; (iii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iv) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.17: First stage estimates of two-stages least square estimates using control variable ethnolinguistic fractionalization

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	56	56	71	71	107	107	84	84	117	117	91	91
Ethnolinguistic fractionalization	0.01	-0.43	0.14	0.01	0.07	-0.76	-0.31	0.06	-0.05	-0.66	-0.08	-0.80
(Ethnoling_frac)	(0.30)	(0.40)	(0.22)	(0.38)	(0.20)	(0.33)**	(0.26)	(0.33)	(0.23)	(0.33)**	(0.23)	(0.35)**
Geography	0.01	3.18	-0.01	3.74	0.56	2.39	0.14	3.86	0.41	2.34	0.24	2.45
(GEO)	(0.37)	(0.54)***	(0.27)	(0.54)***	(0.27)**	(0.48)***	(0.35)	(0.44)***	(0.29)	(0.48)***	(0.30)	(0.53)***
Constructed Trade	0.55	0.17	0.49	0.16	0.42	0.23	0.40	0.23	0.40	0.38	0.45	0.38
(Trade_FR_ROM)	(0.07)***	(0.11)*	(0.06)***	(0.12)	(0.05)***	(0.09)***	(0.06)***	(0.09)**	(0.05)***	(0.09)***	(0.06)***	(0.11)***
Pop. speaking English	0.43	0.95	0.33	0.66	0.23	0.52	0.28	0.14	0.06	0.92	0.11	1.18
(Eng_Lang)	(0.16)***	(0.34)***	(0.15)**	(0.42)	(0.18)	(0.40)	(0.21)	(0.33)	(0.15)	(0.51)*	(0.14)	(0.39)***
Pop. speaking other European languages	-0.20	0.15	-0.12	0.32	-0.17	0.52	-0.17	0.44	-0.15	0.48	-0.24	0.38
(EUR_Lang)	(0.13)	(0.21)	(0.12)	(0.19)*	(0.12)	(0.18)***	(0.11)	(0.17)***	(0.09)	(0.20)**	(0.10)**	(0.21)*
First-stage F-test	21.3	3.7	25.4	2.5	26.9	7.0	17.2	4.3	20.9	7.3	26.6	9.5
Angrist-Pischke F-statistics (p-value)	<0.001	0.01	<0.001	0.03	<0.001	<0.001	<0.001	0.01	<0.001	<0.001	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	0.04		0.04		0.01		0.01		0.01		<0.001	
Kleibergen-Paap Wald rk F statistic	3.67		2.35		5.89		3.47		4.81		8.97	
Stock-Yogo critical values 10%	13.43											
Stock-Yogo critical values 25%	5.45											
R-Square	0.60	0.71	0.54	0.66	0.42	0.53	0.44	0.67	0.39	0.52	0.49	0.60

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) Ethnolinguistic fractionalization following Alesina et al. (2003); (ii) GEO, the variable for geography, which is measured as distance from the equator; (iii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iv) the proportion of the population of a country that speaks English (Eng_Lang); and (v) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.18: Determinants of income: Control variable health

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	56	56	56	56	56
Health	4.29	3.57	3.43	3.35	3.34
(LN_Health_lifeexp)	(1.35)***	(0.99)***	(0.89)***	(0.90)***	(0.92)***
Geography	1.40	0.57	0.29	-0.08	-0.19
(GEO)	(1.09)	(0.77)	(0.77)	(0.81)	(0.82)
Trade	-0.07	-0.14	-0.10	-0.11	-0.11
(LN_TRADE_WB)	(0.21)	(0.16)	(0.16)	(0.16)	(0.16)
Institutions	0.25	0.53	0.44	0.38	0.32
(Inst_rule_of_law)	(0.39)	(0.26)**	(0.24)*	(0.24)*	(0.24)
R-Square	0.76	0.84	0.82	0.76	0.72
Hansen Test (p-value)	0.22	0.02	0.01	0.01	0.01

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	71	71	71	71	71
Health	3.71	3.07	2.57	2.23	2.12
(LN_Health_lifeexp)	(0.76)***	(0.66)***	(0.76)***	(0.92)**	(0.97)**
Geography	3.17	1.00	0.08	-1.05	-1.32
(GEO)	(1.17)***	(0.92)	(1.02)	(1.30)	(1.37)
Trade	-0.17	-0.22	-0.20	-0.23	-0.23
(LN_TRADE_WB)	(0.20)	(0.17)	(0.18)	(0.20)	(0.21)
Institutions	0.24	0.64	0.75	0.88	0.89
(Inst_rule_of_law)	(0.37)	(0.28)**	(0.32)**	(0.40)**	(0.42)**
R-Square	0.86	0.86	0.81	0.68	0.63
Hansen Test (p-value)	0.85	0.23	0.13	0.09	0.08

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	107	107	107	107	107
Health	0.05	0.08	0.08	0.10	0.11
(LN_Health_lifeexp)	(0.35)	(0.28)	(0.23)	(0.19)	(0.18)
Geography	1.98	0.87	0.08	-0.71	-1.05
(GEO)	(0.73)***	(0.83)	(0.87)	(0.94)	(0.97)
Trade	-0.37	-0.45	-0.45	-0.50	-0.51
(LN_TRADE_WB)	(0.28)	(0.28)	(0.27)*	(0.28)*	(0.28)*
Institutions	1.29	1.40	1.41	1.44	1.45
(Inst_rule_of_law)	(0.21)***	(0.23)***	(0.24)***	(0.26)***	(0.27)***
R-Square	0.65	0.58	0.52	0.36	0.29
Hansen Test (p-value)	0.03	0.09	0.16	0.27	0.31

Appendix Table 2.18 continued: Determinants of income: Control variable health

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	83	83	83	83	83
Health	2.22	2.03	1.87	1.79	1.63
(LN_Health_lifeexp)	(0.83)***	(1.11)*	(1.24)	(1.46)	(1.63)
Geography	1.73	-0.48	-1.57	-2.81	-3.32
(GEO)	(1.38)	(1.75)	(2.01)	(2.41)	(2.67)
Trade	-0.01	-0.34	-0.49	-0.67	-0.74
(LN_TRADE_WB)	(0.28)	(0.33)	(0.36)	(0.41)*	(0.44)*
Institutions	0.87	1.30	1.45	1.63	1.75
(Inst_rule_of_law)	(0.47)*	(0.61)**	(0.70)**	(0.84)**	(0.94)*
R-Square	0.84	0.76	0.67	0.50	0.38
Hansen Test (p-value)	0.13	0.33	0.45	0.60	0.60

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	116	116	116	116	116
Health	2.97	3.30	3.26	3.23	3.10
(LN_Health_lifeexp)	(0.39)***	(0.40)***	(0.41)***	(0.44)***	(0.48)***
Geography	1.21	0.50	0.07	-0.45	-0.69
(GEO)	(0.48)***	(0.47)	(0.47)	(0.50)	(0.53)
Trade	0.03	-0.14	-0.24	-0.37	-0.40
(LN_TRADE_WB)	(0.16)	(0.16)	(0.16)	(0.17)**	(0.19)**
Institutions	0.73	0.77	0.79	0.81	0.86
(Inst_rule_of_law)	(0.13)***	(0.13)***	(0.14)***	(0.16)***	(0.18)***
R-Square	0.85	0.84	0.82	0.76	0.72
Hansen Test (p-value)	0.52	0.16	0.11	0.11	0.10

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	90	90	90	90	90
Health	3.37	3.52	3.37	3.26	3.23
(LN_Health_lifeexp)	(0.55)***	(0.53)***	(0.52)***	(0.54)***	(0.55)***
Geography	1.13	0.26	-0.23	-0.86	-1.09
(GEO)	(0.46)***	(0.42)	(0.41)	(0.43)**	(0.45)**
Trade	0.13	-0.03	-0.11	-0.21	-0.24
(LN_TRADE_WB)	(0.16)	(0.13)	(0.13)	(0.14)	(0.15)
Institutions	0.53	0.62	0.66	0.71	0.72
(Inst_rule_of_law)	(0.12)***	(0.10)***	(0.10)***	(0.12)***	(0.13)***
R-Square	0.85	0.85	0.83	0.78	0.76
Hansen Test (p-value)	0.58	0.40	0.27	0.22	0.23

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) Health measured as life expectancy at birth in 1970 (number of years) and taken from the World Bank Development Indicators; (ii) GEO, the variable for geography, which is measured as distance from equator; (iii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iv) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.19: First stage estimates of two-stages least square estimates using control variable health

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	56	56	71	71	107	107	83	83	116	116	90	90
Health	0.75	2.55	0.72	1.40	-0.06	-0.10	0.42	1.29	0.88	1.14	0.66	1.14
(LN_Health_lifeexp)	(0.44)*	(0.58)***	(0.28)***	(0.51)***	(0.08)	(0.16)	(0.40)	(0.49)***	(0.32)***	(0.48)**	(0.36)*	(0.56)**
Geography	-0.40	2.10	-0.62	2.78	0.55	3.07	0.12	2.95	-0.19	1.98	-0.15	2.19
(GEO)	(0.34)	(0.46)***	(0.31)**	(0.48)***	(0.23)**	(0.40)***	(0.33)	(0.43)***	(0.32)	(0.54)***	(0.39)	(0.64)***
Constructed Trade	0.53	0.11	0.48	0.14	0.42	0.24	0.39	0.21	0.39	0.38	0.44	0.38
(Trade_FR_ROM)	(0.07)***	(0.09)	(0.06)***	(0.12)	(0.05)***	(0.09)***	(0.06)***	(0.10)**	(0.05)***	(0.09)***	(0.05)***	(0.10)***
Pop. speaking English	0.41	0.83	0.34	0.65	0.24	0.45	0.22	0.16	0.05	0.88	0.11	1.18
(Eng_Lang)	(0.17)**	(0.28)***	(0.15)**	(0.42)	(0.17)	(0.36)	(0.20)	(0.36)	(0.15)	(0.52)*	(0.15)	(0.35)***
Pop. speaking other European languages	-0.32	-0.19	-0.27	0.06	-0.17	0.68	-0.20	0.21	-0.30	0.38	-0.33	0.32
(EUR_Lang)	(0.12)***	(0.22)	(0.12)**	(0.21)	(0.11)	(0.18)***	(0.12)	(0.19)	(0.11)***	(0.21)*	(0.13)***	(0.24)
First-stage F-test	20.8	3.0	24.5	1.0	27.2	8.9	17.3	1.8	22.7	6.9	27.0	10.2
Angrist-Pischke F-statistics (p-value)	<0.001	0.02	<0.001	0.26	<0.001	<0.001	<0.001	0.25	<0.001	0.01	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	0.11		0.28		<0.001		0.25		0.01		0.01	
Kleibergen-Paap Wald rk F statistic	2.88		0.88		7.77		0.91		4.01		9.01	
Stock-Yogo critical values 10%	13.43											
Stock-Yogo critical values 25%	5.45											
R-Square	0.62	0.79	0.58	0.70	0.42	0.50	0.44	0.71	0.45	0.54	0.52	0.61

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) Health measured as life expectancy at birth in 1970 (number of years) and taken from the World Bank Development Indicators; (ii) GEO, the variable for geography, which is measured as distance from the equator; (iii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iv) the proportion of the population of a country that speaks English (Eng_Lang); and (v) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reportedin parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.20: Determinants of income: Control variable lagged income

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	54	54	54	54	54
20-year lagged income	1.03	0.87	0.81	0.76	0.75
(LN_Income_Lag)	(0.24)***	(0.15)***	(0.14)	(0.15)***	(0.16)***
Geography	2.44	1.51	1.23	0.91	0.82
(GEO)	(1.37)*	(0.75)**	(0.71)*	(0.73)	(0.74)
Trade	0.05	-0.01	0.04	0.04	0.04
(LN_TRADE_WB)	(0.17)	(0.11)	(0.11)	(0.13)	(0.14)
Institutions	-0.27	0.08	0.02	-0.02	-0.07
(Inst_rule_of_law)	(0.44)	(0.24)	(0.22)	(0.23)	(0.23)
R-Square	0.76	0.89	0.87	0.80	0.76
Hansen Test (p-value)	0.42	0.97	0.35	0.13	0.10

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	68	68	68	68	68
20-year lagged income	0.98	1.16	1.18	1.26	1.28
(LN_Income_Lag)	(0.39)***	(0.33)***	(0.30)***	(0.32)***	(0.33)***
Geography	5.09	3.87	3.02	2.39	2.20
(GEO)	(2.52)**	(2.05)*	(1.76)*	(1.72)	(1.73)
Trade	-0.27	-0.19	-0.10	-0.07	-0.06
(LN_TRADE_WB)	(0.22)	(0.15)	(0.13)	(0.14)	(0.14)
Institutions	-0.49	-0.60	-0.62	-0.74	-0.80
(Inst_rule_of_law)	(0.87)	(0.73)	(0.64)	(0.64)	(0.65)
R-Square	0.72	0.75	0.75	0.68	0.64
Hansen Test (p-value)	0.99	0.71	0.72	0.56	0.57

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	91	91	91	91	91
20-year lagged income	1.70	1.07	0.70	0.34	0.16
(LN_Income_Lag)	(1.23)	(0.58)*	(0.43)	(0.61)	(0.76)
Geography	7.91	3.13	0.42	-2.42	-3.69
(GEO)	(7.50)	(3.37)	(2.55)	(3.90)	(4.88)
Trade	-0.05	-0.16	-0.19	-0.26	-0.28
(LN_TRADE_WB)	(0.40)	(0.20)	(0.17)	(0.24)	(0.29)
Institutions	-1.78	-0.29	0.49	1.30	1.68
(Inst_rule_of_law)	(2.84)	(1.30)	(0.99)	(1.45)	(1.80)
R-Square	0.32	0.77	0.83	0.64	0.44
Hansen Test (p-value)	0.62	0.11	0.05	0.21	0.28

Appendix Table 2.20 continued: Determinants of income: Control variable lagged income

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	81	81	81	81	81
20-year lagged income	1.23	0.83	0.75	0.60	0.57
(LN_Income_Lag)	(0.82)	(0.30)***	(0.36)**	(0.61)	(0.66)
Geography	5.91	1.17	-0.30	-2.47	-2.96
(GEO)	(5.76)	(2.00)	(2.58)	(4.46)	(4.83)
Trade	0.50	-0.02	-0.18	-0.41	-0.46
(LN_TRADE_WB)	(0.66)	(0.25)	(0.31)	(0.51)	(0.55)
Institutions	-1.13	0.31	0.59	1.14	1.24
(Inst_rule_of_law)	(2.32)	(0.82)	(1.03)	(1.77)	(1.91)
R-Square	0.70	0.93	0.91	0.77	0.70
Hansen Test (p-value)	0.96	0.86	0.76	0.79	0.73

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	100	100	100	100	100
20-year lagged income	0.87	0.98	0.99	0.99	1.00
(LN_Income_Lag)	(0.12)***	(0.11)***	(0.11)***	(0.13)***	(0.14)***
Geography	1.68	1.26	0.93	0.44	0.37
(GEO)	(0.74)**	(0.72)*	(0.94)	(0.83)	(0.87)
Trade	0.05	-0.08	-0.15	-0.25	-0.26
(LN_TRADE_WB)	(0.15)	(0.14)	(0.14)	(0.15)	(0.16)*
Institutions	0.23	0.12	0.08	0.07	0.04
(Inst_rule_of_law)	(0.27)	(0.26)	(0.27)	(0.31)	(0.33)
R-Square	0.90	0.91	0.90	0.87	0.85
Hansen Test (p-value)	0.28	0.80	0.89	0.56	0.49

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	51	51	51	51	51
20-year lagged income	0.96	0.99	1.01	1.03	1.04
(LN_Income_Lag)	(0.12)***	(0.08)***	(0.07)***	(0.08)***	(0.08)***
Geography	2.13	0.85	0.36	-0.36	-0.52
(GEO)	(0.97)**	(0.59)	(0.50)	(0.51)	(0.55)
Trade	0.24	0.07	0.01	-0.08	-0.12
(LN_TRADE_WB)	(0.16)	(0.10)	(0.09)	(0.10)	(0.11)
Institutions	-0.22	-0.05	-0.07	-0.05	-0.06
(Inst_rule_of_law)	(0.29)	(0.17)	(0.14)	(0.13)	(0.15)
R-Square	0.89	0.94	0.95	0.93	0.92
Hansen Test (p-value)	0.63	0.94	0.38	0.10	0.09

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) a 20-year lag of the national income per capita in PPPs, taken from the Penn World Tables; (ii) GEO, the variable for geography, which is measured as distance from equator; (iii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iv) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.21: First stage estimates of two-stages least square estimates with control variable lagged income

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	54	54	68	68	91	91	81	81	100	100	51	51
20-year lagged income	-0.14	0.28	0.08	0.47	0.01	0.40	0.14	0.37	0.13	0.39	0.13	0.25
(LN_Income_Lag)	(0.09)	(0.11)**	(0.07)	(0.10)***	(0.06)	(0.07)***	(0.08)*	(0.08)***	(0.07)*	(0.08)***	(0.06)**	(0.12)**
Geography	0.42	2.76	-0.53	2.39	0.01	2.71	-0.21	2.45	-0.42	2.17	-0.61	2.78
(GEO)	(0.44)	(0.52)***	(0.32)*	(0.49)***	(0.31)	(0.39)***	(0.40)	(0.42)***	(0.36)	(0.49)***	(0.38)	(0.63)***
Constructed Trade	0.57	0.12	0.50	-0.01	0.45	0.04	0.39	0.11	0.40	0.15	0.43	0.23
(Trade_FR_ROM)	(0.07)***	(0.10)	(0.06)	(0.09)	(0.06)***	(0.07)	(0.06)***	(0.07)	(0.06)***	(0.07)**	(0.05)***	(0.13)*
Pop. speaking English	0.49	0.73	0.34	0.34	0.34	0.16	0.25	0.19	0.11	0.58	0.09	1.01
(Eng_Lang)	(0.17)***	(0.29)**	(0.14)**	(0.31)	(0.16)**	(0.24)	(0.22)	(0.23)	(0.17)	(0.33)*	(0.15)	(0.43)**
Pop. speaking other European languages	-0.07	-0.11	-0.19	-0.20	-0.10	0.05	-0.26	-0.03	-0.20	-0.09	-0.38	0.16
(EUR_Lang)	(0.16)	(0.23)	(0.13)	(0.16)	(0.12)	(0.15)	(0.13)*	(0.17)	(0.12)*	(0.19)	(0.12)***	(0.23)
First-stage F-test	23.93	2.10	25.35	0.79	25.82	0.36	16.42	1.07	17.97	1.98	38.11	2.14
Angrist-Pischke F-statistics (p-value)	<0.001	0.06	<0.001	0.31	<0.001	0.65	<0.001	0.70	<0.001	0.22	<0.001	0.05
Kleibergen-Paap LM test (p-value)	0.13		0.32		0.62		0.71		0.26		0.10	
R-Square	0.61	0.73	0.58	0.76	0.45	0.75	0.46	0.77	0.44	0.69	0.70	0.75

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) a 20-year lag of the national income per capita in PPPs, taken from the Penn World Tables (ii) GEO, the variable for geography, which is measured as distance from the equator; (iii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iv) the proportion of the population of a country that speaks English (Eng_Lang); and (v) the proportion of the population of a country that speaksany Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006).The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.22: Determinants of income: Control variable GINI index

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	36	36	36	36	36
GINI Index	0.01	0.02	0.02	0.03	0.03
(Gini)	(0.01)	(0.01)*	(0.01)**	(0.01)**	(0.01)**
Geography	0.25	-0.48	-0.66	-0.90	-0.99
(GEO)	(1.13)	(1.16)	(1.12)	(1.13)	(1.13)
Trade	-0.48	-0.36	-0.26	-0.20	-0.18
(LN_TRADE_WB)	(0.34)	(0.32)	(0.31)	(0.31)	(0.31)
Institutions	1.16	1.46	1.38	1.35	1.30
(Inst_rule_of_law)	(0.21)***	(0.25)***	(0.25)***	(0.26)***	(0.27)***
R-Square	0.41	0.37	0.31	0.26	0.23
Hansen Test (p-value)	0.07	0.06	0.04	0.03	0.03

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	47	47	47	47	47
GINI Index	-0.02	0.01	0.02	0.03	0.03
(Gini)	(0.02)	(0.01)	(0.01)	(0.01)**	(0.01)***
Geography	2.03	1.43	1.42	1.21	1.20
(GEO)	(1.47)	(1.62)	(1.53)	(1.57)	(1.57)
Trade	-0.58	-0.55	-0.51	-0.52	-0.52
(LN_TRADE_WB)	(0.36)*	(0.40)	(0.39)	(0.40)	(0.40)
Institutions	1.20	1.31	1.19	1.14	1.08
(Inst_rule_of_law)	(0.36)***	(0.41)***	(0.39)***	(0.41)***	(0.41)***
R-Square	0.61	0.48	0.42	0.33	0.31
Hansen Test (p-value)	0.04	0.05	0.10	0.17	0.22

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	75	75	75	75	75
GINI Index	-0.02	0.01	0.01	0.03	0.03
(Gini)	(0.01)*	(0.01)	(0.01)	(0.01)**	(0.01)**
Geography	1.02	0.88	0.68	0.51	0.30
(GEO)	(0.95)	(1.02)	(1.02)	(1.06)	(1.10)
Trade	-0.38	-0.39	-0.38	-0.39	-0.40
(LN_TRADE_WB)	(0.34)	(0.34)	(0.33)	(0.33)	(0.34)
Institutions	1.71	1.78	1.73	1.70	1.70
(Inst_rule_of_law)	(0.39)***	(0.39)***	(0.40)***	(0.41)***	(0.43)***
R-Square	0.41	0.27	0.20	0.12	0.05
Hansen Test (p-value)	0.03	0.09	0.16	0.26	0.28

Appendix Table 2.22 continued: Determinants of income: Control variable GINI index

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	58	58	58	58	58
GINI Index	-0.01	0.02	0.04	0.05	0.06
(Gini)	(0.02)	(0.02)	(0.02)**	(0.02)***	(0.02)***
Geography	1.06	0.93	1.19	1.24	1.46
(GEO)	(1.80)	(1.94)	(1.77)	(1.75)	(1.65)
Trade	-0.71	-0.69	-0.67	-0.66	-0.59
(LN_TRADE_WB)	(0.49)	(0.51)	(0.46)	(0.44)	(0.42)
Institutions	1.69	1.75	1.58	1.50	1.43
(Inst_rule_of_law)	(0.62)***	(0.64)***	(0.58)***	(0.56)***	(0.52)***
R-Square	0.27	0.16	0.26	0.33	0.37
Hansen Test (p-value)	0.16	0.32	0.43	0.56	0.55

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	107	107	107	107	107
GINI Index	-0.01	0.01	0.03	0.04	0.05
(Gini)	(0.01)	(0.01)	(0.01)***	(0.01)***	(0.01)***
Geography	1.92	2.04	2.16	2.18	2.11
(GEO)	(0.82)**	(0.87)	(0.81)***	(0.80)***	(0.79)***
Trade	-0.50	-0.63	-0.61	-0.64	-0.62
(LN_TRADE_WB)	(0.33)	(0.35)*	(0.33)*	(0.32)**	(0.31)**
Institutions	1.27	1.29	1.24	1.20	1.20
(Inst_rule_of_law)	(0.20)***	(0.21)***	(0.20)***	(0.19)***	(0.19)***
R-Square	0.68	0.60	0.59	0.55	0.55
Hansen Test (p-value)	0.09	0.07	0.09	0.11	0.12

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	86	86	86	86	86
GINI Index	-0.01	0.01	0.03	0.04	0.04
(Gini)	(0.01)	(0.01)	(0.01)**	(0.01)***	(0.01)
Geography	2.59	2.31	2.27	2.07	2.04
(GEO)	(0.94)***	(0.94)***	(0.87)***	(0.87)**	(0.88)**
Trade	-0.19	-0.32	-0.30	-0.32	-0.33
(LN_TRADE_WB)	(0.24)	(0.23)	(0.22)	(0.21)	(0.21)
Institutions	0.94	1.00	0.96	0.95	0.93
(Inst_rule_of_law)	(0.16)***	(0.16)***	(0.15)***	(0.15)***	(0.15)***
R-Square	0.74	0.69	0.68	0.64	0.62
Hansen Test (p-value)	0.02	0.02	0.02	0.03	0.03

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) the GINI index, taken from the World Bank; (ii) GEO, the variable for geography, which is measured as distance from equator; (iii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iv) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.23: First stage estimates of two-stages least square estimates with GINI Index control

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	36	36	47	47	75	75	58	58	107	107	86	86
GINI Index	0.01	-0.01	0.01	-0.01	0.01	-0.01	0.01	-0.01	0.01	-0.02	0.01	-0.01
(Gini)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)*	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)*	(0.01)*	(0.01)
Geography	0.35	1.90	0.21	2.15	1.08	1.71	0.80	2.33	0.68	2.59	0.79	3.20
(GEO)	(0.57)	(0.61)***	80.46)	(0.71)***	(0.29)***	(0.47)***	(0.43)*	(0.61)***	(0.22)***	(0.40)***	(0.24)***	(0.41)***
Constructed Trade	0.61	0.18	0.44	0.08	0.45	0.09	0.34	0.09	0.34	0.34	0.41	0.32
(Trade_FR_ROM)	(0.10)***	(0.16)	(0.08)***	(0.17)	(0.06)***	(0.10)	(0.07)***	(0.11)	(0.05)***	(0.08)***	(0.05)***	(0.10)***
Pop. speaking English	0.60	1.77	0.42	1.06	0.23	0.67	0.21	0.66	-0.03	0.71	0.10	1.09
(Eng_Lang)	(0.34)*	(0.39)***	(0.26)	(0.64)*	(0.31)	(0.60)	(0.29)	(0.58)	(0.17)	(0.47)	(0.17)	(0.40)***
Pop. speaking other European languages	-0.26	0.12	-0.27	0.21	-0.23	0.59	-0.20	0.31	-0.11	0.77	-0.24	0.67
(EUR_Lang)	(0.22)	(0.32)	(0.16)*	(0.26)	(0.16)	(0.22)***	(0.19)	(0.28)	(0.09)	(0.17)***	(0.10)**	(0.17)***
First-stage F-test	15.92	26.93	12.02	1.68	17.77	4.57	9.23	1.73	18.55	12.74	32.45	11.54
Angrist-Pischke F-statistics (p-value)	<0.001	<0.001	<0.001	0.14	<0.001	0.002	<0.001	0.10	<0.001	<0.001	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	<0.001		0.10		0.01		0.15		<0.001		<0.001	
R-Square	0.55	0.55	0.43	0.43	0.50	0.33	0.28	0.37	0.41	0.61	0.53	0.64

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) the GINI Index as taken from the World Bank; (ii) GEO, the variable for geography, which is measured as distance from the equator; (iii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iv) the proportion of the population of a country that speaks English (Eng_Lang); and (v) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006).The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.24: Determinants of income: Control variable human capital

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	55	55	55	55	55
School enrolment rate 1985	2.90	1.80	1.32	1.11	1.04
(HC_schoolenr_85)	(0.90)***	(0.51)***	(0.52)***	(0.62)*	(0.66)
Geography	1.03	0.17	-0.18	-0.63	-0.78
(GEO)	(1.23)	(0.91)	(0.85)	(0.89)	(0.90)
Trade	0.16	-0.03	-0.04	-0.06	-0.07
(LN_TRADE_WB)	(0.25)	(0.20)	(0.19)	(0.19)	(0.20)
Institutions	0.84	1.08	1.01	0.97	0.92
(Inst_rule_of_law)	(0.31)***	(0.23)***	(0.21)***	(0.21)***	(0.21)***
R-Square	0.65	0.71	0.67	0.59	0.54
Hansen Test (p-value)	0.59	0.09	0.02	0.01	0.01

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	69	69	69	69	69
School enrolment rate 1990	1.49	1.28	1.14	1.04	1.01
(HC_schoolenr_90)	(0.49)***	(0.40)***	(0.37)***	(0.39)***	(0.40)***
Geography	3.05	0.85	-0.04	-1.17	-1.45
(GEO)	(1.21)***	(1.41)	(1.47)	(1.71)	(1.76)
Trade	-0.11	-0.16	-0.13	-0.16	-0.16
(LN_TRADE_WB)	(0.24)	(0.22)	(0.22)	(0.23)	(0.24)
Institutions	0.76	1.08	1.11	1.20	1.19
(Inst_rule_of_law)	(0.31)***	(0.33)***	(0.35)***	(0.40)***	(0.41)***
R-Square	0.78	0.73	0.66	0.51	0.45
Hansen Test (p-value)	0.58	0.21	0.15	0.11	0.10

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	89	89	89	89	89
School enrolment rate 1995	1.98	1.54	1.23	1.06	0.97
(HC_schoolenr_95)	(0.51)***	(0.43)***	(0.37)***	(0.39)***	(0.40)**
Geography	2.43	0.53	-0.59	-1.72	-2.19
(GEO)	(1.10)**	(1.14)	(1.28)	(1.50)	(1.60)
Trade	0.13	-0.07	-0.17	-0.26	-0.29
(LN_TRADE_WB)	(0.25)	(0.24)	(0.24)	(0.27)	(0.29)
Institutions	0.76	1.15	1.30	1.45	1.51
(Inst_rule_of_law)	(0.31)***	(0.31)***	(0.34)***	(0.40)***	(0.43)***
R-Square	0.74	0.71	0.62	0.45	0.35
Hansen Test (p-value)	0.28	0.25	0.36	0.51	0.54

Appendix Table 2.24 continued: Determinants of income: Control variable human capital

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	79	79	79	79	79
School enrolment rate 2000	2.40	2.84	2.80	2.98	3.00
(HC_schoolenr_00)	(0.68)***	(0.64)***	(0.73)***	(0.87)***	(0.95)***
Geography	2.44	0.64	-0.34	-1.37	-1.74
(GEO)	(1.55)	(1.74)	(1.87)	(2.08)	(2.20)
Trade	0.04	-0.22	-0.35	-0.48	-0.52
(LN_TRADE_WB)	(0.31)	(0.31)	(0.31)	(0.34)	(0.35)
Institutions	0.98	1.24	1.34	1.43	1.47
(Inst_rule_of_law)	(0.41)**	(0.46)***	(0.49)***	(0.55)***	(0.59)***
R-Square	0.80	0.72	0.65	0.53	0.46
Hansen Test (p-value)	0.70	0.88	0.91	0.83	0.89

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	98	98	98	98	98
School enrolment rate 2005	3.85	4.73	4.94	5.20	5.19
(HC_schoolenr_05)	(1.01)***	(1.17)***	(1.20)***	(1.27)***	(1.29)***
Geography	2.45	1.68	1.22	0.62	0.40
(GEO)	(1.26)**	(1.38)	(1.42)	(1.51)	(1.53)
Trade	-0.33	-0.53	-0.61	-0.71	-0.73
(LN_TRADE_WB)	(0.29)	(0.31)*	(0.31)**	(0.32)**	(0.33)**
Institutions	1.12	1.22	1.23	1.25	1.27
(Inst_rule_of_law)	(0.31)***	(0.34)***	(0.36)***	(0.38)***	(0.39)***
R-Square	0.65	0.53	0.47	0.36	0.32
Hansen Test (p-value)	0.70	0.53	0.48	0.46	0.43

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	72	72	72	72	72
School enrolment rate 2010	9.08	10.61	10.99	11.50	11.63
(HC_schoolenr_10)	(5.02)*	(5.64)*	(5.73)*	(5.88)**	(5.88)**
Geography	3.10	2.09	1.50	0.71	0.46
(GEO)	(1.21)***	(1.34)	(1.39)	(1.47)	(1.50)
Trade	0.06	-0.06	-0.09	-0.15	-0.17
(LN_TRADE_WB)	(0.31)	(0.34)	(0.35)	(0.37)	(0.38)
Institutions	0.87	1.02	1.04	1.10	1.11
(Inst_rule_of_law)	(0.27)***	(0.30)***	(0.30)***	(0.32)***	(0.33)***
R-Square	0.23	0.01	0.01	0.01	0.01
Hansen Test (p-value)	0.62	0.59	0.56	0.54	0.55

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) Human Capital measured as 5-year average around the given time period, and instrumented by average primary school enrolment rates 1970-79. Data are in logs and taken from the UNESCO Institute for Statistics; (ii) GEO, the variable for geography, which is measured as distance from equator; (iii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iv) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.25: First stage estimates of two-stages least square estimates with control variable human capital

	1985			1990			1995			2000			2005			2010		
Dependent variable =	Trade	Institu-tions	Human Capital	Trade	Institu-tions	Human Capital	Trade	Institu-tions	Human Capital	Trade	Institu-tions	Human Capital	Trade	Institu-tions	Human Capital	Trade	Institu-tions	Human Capital
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Sample size	55	55	55	69	69	69	89	89	89	79	79	79	98	98	98	72	72	72
School enrolment rate 1970s	0.06	0.62	0.57	-0.33	0.31	0.66	0.24	0.21	0.54	0.31	0.17	0.29	0.29	0.13	0.21	0.23	0.01	0.10
(HC_schoolenr_70)	(0.30)	(0.28)**	(0.06)***	(0.24)	(0.15)**	(0.08)***	(0.10)***	(0.11)**	(0.07)***	(0.11)***	(0.12)	(0.05)***	(0.09)***	(0.11)	(0.05)***	(0.11)**	(0.14)	(0.06)*
Geography	-0.02	3.09	-0.10	-0.33	3.39	-0.10	-0.01	3.36	-0.02	1.37	3.62	-0.06	0.06	3.22	-0.20	0.02	3.62	-0.14
(GEO)	(0.30)	(0.42)***	(0.08)	(0.24)	(0.43)***	(0.07)	(0.24)	(0.44)***	(0.07)	(0.23)	(0.37)***	(0.06)	(0.21)	(0.41)***	(0.06)***	(0.27)	(0.43)***	(0.06)**
Constructed Trade	0.55	0.19	-0.03	0.50	0.18	-0.03	0.43	0.17	-0.02	0.39	0.24	0.01	0.39	0.28	0.01	0.43	0.30	-0.01
(Trade_FR_ROM)	(0.08)***	(0.10)**	(0.01)	(0.06)***	(0.12)	(0.01)*	(0.05)***	(0.09)*	(0.02)	(0.06)***	(0.10)**	(0.01)	(0.05)***	(0.09)***	(0.01)	(0.05)***	(0.09)***	(0.01)
Pop. speaking English	0.43	1.06	-0.05	0.40	0.77	-0.02	0.41	0.43	-0.02	0.25	0.24	-0.07	0.09	0.76	-0.05	0.12	1.09	-0.02
(Eng_Lang)	(0.17)**	(0.35)***	(0.03)	(0.15)***	(0.44)*	(0.03)	(0.18)**	(0.40)	(0.04)	(0.20)	(0.36)	(0.04)*	(0.15)	(0.48)	(0.04)	(0.15)	(0.34)***	(0.03)
Pop. speaking other European languages	-0.22	-0.03	0.03	-0.24	0.14	0.03	-0.23	0.41	0.04	-0.23	0.30	0.06	-0.18	0.37	0.01	-0.23	0.41	-0.01
(EUR_Lang)	(0.16)	(0.25)	(0.03)	(0.12)*	(0.22)	(0.03)	(0.12)*	(0.19)**	(0.03)	(0.12)**	(0.20)	(0.03)**	(0.10)*	(0.20)*	(0.02)	(0.13)*	(0.22)*	(0.03)
First-stage F-test	16.3	4.42	29.8	19.2	2.9	28.8	21.5	5.4	21.6	13.8	3.6	11.1	17.8	4.0	6.9	19.2	7.9	1.1
Angrist-Pischke F-statistics (p-value)	<0.001	0.01	<0.001	<0.001	0.11	<0.001	<0.001	0.01	<0.001	<0.001	0.07	<0.001	<0.001	0.01	<0.001	<0.001	<0.001	0.17
Kleibergen-Paap LM test (p-value)		0.11			0.15			0.02			0.09			0.02			0.32	
R-Square	0.60	0.72	0.77	0.57	0.66	0.86	0.47	0.59	0.74	0.45	0.68	0.56	0.43	0.57	0.41	0.52	0.66	0.16

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) Human Capital lags measured as logs of average primary school enrolment rates 1970-79 and taken from the UNESCO Institute for Statistics; (ii) GEO, the variable for geography, which is measured as distance from the equator; (iii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iv) the proportion of the population of a country that speaks English (Eng_Lang); and (v) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.26: Determinants of income: Regional dummy control variables

1985 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	56	56	56	56	56
Sub-Saharan Africa	-2.10	-1.87	-1.42	-1.11	-0.96
(Subsah_AFR)	(0.65)***	(0.51)***	(0.47)***	(0.46)**	(0.45)**
Middle East and North Africa	-0.35	-0.56	-0.44	-0.33	-0.27
(MEast_NAfr)	(0.42)	(0.37)	(0.36)	(0.37)	(0.37)
Europe and Central Asia	0.31	-0.18	-0.16	-0.21	-0.19
(Eur_Asia)	(0.27)	(0.20)	(0.19)	(0.20)	(0.21)
North America					
(NorthAm)					
Latin America	-0.53	-0.56	-0.31	-0.12	-0.04
(LatAM)	(0.50)	(0.45)	(0.42)	(0.42)	(0.42)
South-East Asia and Pacific	-0.97	-1.34	-1.24	-1.22	-1.19
(SE_Asia)	(0.39)***	(0.34)***	(0.33)***	(0.33)***	(0.34)***
Geography	-2.20	-1.94	-1.97	-2.04	-2.12
(GEO)	(1.19)*	(1.02)*	(0.99)**	(0.99)**	(0.98)**
Trade	-0.09	-0.10	-0.10	-0.13	-0.15
(LN_TRADE_WB)	(0.22)	(0.19)	(0.18)	(0.19)	(0.18)
Institutions	1.06	1.17	1.13	1.11	1.08
(Inst_rule_of_law)	(0.27)***	(0.21)***	(0.20)***	(0.20)***	(0.21)***
R-Square	0.77	0.81	0.79	0.73	0.69
Hansen Test (p-value)	0.71	0.30	0.32	0.30	0.32

1990 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	71	71	71	71	71
Sub-Saharan Africa	-2.25	-1.88	-1.46	-1.13	-0.97
(Subsah_AFR)	(0.37)***	(0.33)***	(0.32)***	(0.34)***	(0.36)***
Middle East and North Africa	-0.86	-0.89	-0.81	-0.69	-0.63
(MEast_NAfr)	(0.30)***	(0.31)***	(0.30)***	(0.31)**	(0.31)**
Europe and Central Asia	0.08	-0.09	-0.12	-0.17	-0.16
(Eur_Asia)	(0.18)	(0.20)	(0.20)	(0.21)	(0.21)
North America					
(NorthAm)					
Latin America	-0.80	-0.52	-0.29	-0.03	0.08
(LatAM)	(0.29)***	(0.31)*	(0.30)	(0.30)	(0.31)
South-East Asia and Pacific	-0.81	-0.97	-0.90	-0.84	-0.78
(SE_Asia)	(0.26)***	(0.27)***	(0.26)***	(0.26)***	(0.26)***
Geography	0.86	0.10	-0.16	-0.58	-0.65
(GEO)	(0.93)	(0.91)	(0.92)	(0.98)	(1.01)
Trade	-0.10	-0.11	-0.09	-0.11	-0.11
(LN_TRADE_WB)	(0.21)	(0.21)	(0.21)	(0.23)	(0.24)
Institutions	0.68	0.83	0.85	0.90	0.90
(Inst_rule_of_law)	(0.20)***	(0.21)***	(0.22)***	(0.23)***	(0.24)***
R-Square	0.87	0.85	0.80	0.71	0.67
Hansen Test (p-value)	0.31	0.19	0.23	0.27	0.32

1995 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	107	107	107	107	107
Sub-Saharan Africa	-1.36	-1.05	-0.80	-0.65	-0.59
(Subsah_AFR)	(0.31)***	(0.30)***	(0.29)***	(0.31)**	(0.31)*
Middle East and North Africa					
(MEast_NAfr)					
Europe and Central Asia	0.46	0.40	0.30	0.19	0.19
(Eur_Asia)	(0.22)**	(0.24)*	(0.24)	(0.25)	(0.26)
North America	0.43	0.57	0.51	0.46	0.48
(NorthAm)	(0.25)*	(0.28)**	(0.29)*	(0.32)	(0.33)
Latin America	-0.22	0.12	0.31	0.47	0.55
(LatAM)	(0.24)	(0.25)	(0.26)	(0.27)*	(0.27)**
South-East Asia and Pacific	-0.09	-0.17	-0.16	-0.18	-0.14
(SE_Asia)	(0.24)	(0.27)	(0.27)	(0.28)	(0.28)
Geography	0.17	0.25	0.24	0.18	0.06
(GEO)	(0.77)	(0.68)	(0.65)	(0.67)	(0.67)
Trade	0.01	-0.01	0.01	-0.02	-0.01
(LN_TRADE_WB)	(0.18)	(0.17)	(0.17)	(0.17)	(0.18)
Institutions	0.92	0.86	0.82	0.78	0.76
(Inst_rule_of_law)	(0.11)***	(0.09)***	(0.09)***	(0.09)***	(0.10)***
R-Square	0.79	0.79	0.77	0.71	0.68
Hansen Test (p-value)	0.11	0.35	0.80	0.62	0.50

Appendix Table 2.26 continued: Determinants of income: Regional dummy control variables

2000 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	84	84	84	84	84
Sub-Saharan Africa	-1.50	-1.32	-1.12	-0.96	-0.85
(Subsah_AFR)	(0.27)***	(0.28)***	(0.30)***	(0.37)***	(0.40)**
Middle East and North Africa					
(MEast_NAfr)					
Europe and Central Asia	0.26	0.31	0.33	0.31	0.28
(Eur_Asia)	(0.30)	(0.32)	(0.32)	(0.36)	(0.38)
North America	-0.10	0.28	0.35	0.42	0.44
(NorthAm)	(0.46)	(0.44)	(0.41)	(0.43)	(0.44)
Latin America	-0.58	-0.11	0.16	0.42	0.57
(LatAM)	(0.25)**	(0.26)	(0.28)	(0.34)	(0.37)
South-East Asia and Pacific	-0.49	-0.48	-0.40	-0.32	-0.28
(SE_Asia)	(0.24)**	(0.28)*	(0.30)	(0.36)	(0.39)
Geography	-0.15	0.11	0.11	0.03	0.17
(GEO)	(0.90)	(0.78)	(0.78)	(0.84)	(0.83)
Trade	-0.11	-0.07	-0.10	-0.13	-0.11
(LN_TRADE_WB)	(0.25)	(0.22)	(0.21)	(0.21)	(0.20)
Institutions	1.10	0.92	0.86	0.81	0.79
(Inst_rule_of_law)	(0.27)***	(0.23)***	(0.21)***	(0.21)***	(0.20)***
R-Square	0.85	0.86	0.85	0.80	0.78
Hansen Test (p-value)	0.43	0.83	0.97	0.74	0.85

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	117	117	117	117	117
Sub-Saharan Africa	-1.53	-1.45	-1.29	-1.18	-1.12
(Subsah_AFR)	(0.27)***	(0.28)***	(0.29)***	(0.30)***	(0.30)***
Middle East and North Africa					
(MEast_NAfr)					
Europe and Central Asia	0.43	0.47	0.36	0.26	0.17
(Eur_Asia)	(0.29)	(0.29)*	(0.29)	(0.28)	(0.29)
North America	0.01	0.36	0.32	0.36	0.33
(NorthAm)	(0.37)	(0.35)	(0.35)	(0.35)	(0.36)
Latin America	-0.44	-0.04	0.17	0.37	0.46
(LatAM)	(0.29)	(0.29)	(0.28)	(0.28)	(0.27)*
South-East Asia and Pacific	-0.59	-0.45	-0.36	-0.27	-0.26
(SE_Asia)	(0.30)**	(0.31)	(0.30)	(0.29)	(0.29)
Geography	-0.40	-0.41	-0.22	-0.16	-0.07
(GEO)	(0.70)	(0.68)	(0.68)	(0.69)	(0.69)
Trade	-0.14	-0.08	-0.09	-0.09	-0.07
(LN_TRADE_WB)	(0.19)	(0.18)	(0.17)	(0.17)	(0.17)
Institutions	0.96	0.88	0.85	0.81	0.81
(Inst_rule_of_law)	(0.12)***	(0.10)***	(0.09)***	(0.08)***	(0.09)***
R-Square	0.83	0.83	0.81	0.78	0.76
Hansen Test (p-value)	0.50	0.42	0.37	0.29	0.30

2010 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	91	91	91	91	91
Sub-Saharan Africa	-1.29	-1.86	-1.73	-1.76	-1.79
(Subsah_AFR)	(0.35)***	(0.34)***	(0.38)***	(0.42)***	(0.44)***
Middle East and North Africa	0.01	-0.81	-0.89	-1.09	-1.20
(MEast_NAfr)	(0.27)	(0.23)***	(0.21)***	(0.20)***	(0.19)***
Europe and Central Asia	0.79	0.24	0.12	-0.08	-0.16
(Eur_Asia)	(0.22)***	(0.19)	(0.16)	(0.14)	(0.13)
North America					
(NorthAm)					
Latin America	0.06	-0.30	-0.18	-0.17	-0.19
(LatAM)	(0.33)	(0.31)	(0.30)	(0.29)	(0.29)
South-East Asia and Pacific	-0.02	-0.59	-0.60	-0.73	-0.78
(SE_Asia)	(0.31)	(0.29)**	(0.27)**	(0.26)***	(0.26)***
Geography	0.18	-0.25	-0.31	-0.56	-0.65
(GEO)	(0.68)	(0.66)	(0.64)	(0.64)	(0.64)
Trade	-0.15	-0.10	-0.10	-0.10	-0.10
(LN_TRADE_WB)	(0.18)	(0.18)	(0.17)	(0.16)	(0.16)
Institutions	0.77	0.74	0.73	0.71	0.70
(Inst_rule_of_law)	(0.09)***	(0.09)***	(0.08)***	(0.08)***	(0.07)***
R-Square	0.83	0.82	0.81	0.77	0.76
Hansen Test (p-value)	0.48	0.32	0.25	0.18	0.16

Notes: The dependent variable is per capita GDP on PPP basis. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: Regional dummies for (i) Sub-saharan Africa, (ii) for Middle East and North Africa, (iii) for Europe and Central Asia, (iv) for North America, (v) for Latin America and the Carribean, (vi) for South-East Asia and the Pacific incl. Oceania, (vii) GEO, the variable for geography, which is measured as distance from equator; (viii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (ix) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Missing value indicates that variable was dropped due to collinearity. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.27: First stage estimates of two-stages least square estimates using Regional Dummy control variables

	1985		1990		1995		2000		2005		2010	
Dependent variable =	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions	Trade	Institu-tions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample size	56	56	71	71	107	107	84	84	117	117	91	91
Sub-Saharan Africa	0.33	-0.41	-0.19	-0.39	-0.01	-0.16	0.07	0.16	0.01	0.61	-0.36	-0.09
(Subsah_AFR)	(0.39)	(0.43)	(0.36)	(0.43)	(0.16)	(0.27)	(0.17)	(0.30)	(0.14)	(0.29)**	(0.30)	(0.52)
Middle East and North Africa	-0.17	-0.48	-0.23	-0.52							-0.49	-0.40
(MEast_NAfr)	(0.34)	(0.46)	(0.33)	(0.44)							(0.27)*	(0.52)
Europe and Central Asia	-0.24	-0.09	-0.16	-0.26	0.26	-0.34	0.19	0.45	0.20	0.15	-0.14	-0.57
(Eur_Asia)	(0.27)	(0.37)	(0.31)	(0.39)	(0.14)*	(0.28)	(0.17)	(0.33)	(0.14)	(0.32)	(0.23)	(0.40)
North America					0.37	0.38	0.28	1.15	0.27	0.94		
(NorthAm)					(0.36)	(0.43)	(0.42)	(0.49)**	(0.31)	(0.51)*		
Latin America	0.22	-0.80	0.04	-0.95	0.50	-0.73	0.22	-0.11	0.35	-0.29	-0.03	-0.84
(LatAM)	(0.36)	(0.38)**	(0.34)	(0.39)**	(0.17)***	(0.31)**	(0.17)	(0.35)	(0.14)***	(0.33)	(0.29)	(0.50)*
South-East Asia and Pacific	0.13	0.07	-0.03	0.22	0.20	0.36	0.33	0.45	0.31	1.16	-0.08	0.21
(SE_Asia)	(0.29)	(0.37)	(0.31)	(0.38)	(0.15)	(0.29)	(0.19)*	(0.34)	(0.20)	(0.34)***	(0.27)	(0.46)
Geography	0.91	2.78	-0.02	3.13	0.28	3.15	0.31	3.00	0.29	3.46	0.17	3.61
(GEO)	(0.60)	(0.70)***	(0.49)	(0.61)***	(0.47)	(0.69)***	(0.43)	(0.66)***	(0.46)	(0.67)***	(0.49)	(0.86)***
Constructed Trade	0.58	0.24	0.52	0.27	0.44	0.34	0.43	0.29	0.44	0.49	0.47	0.49
(Trade_FR_ROM)	(0.09)***	(0.11)**	(0.08)***	(0.12)**	(0.05)***	(0.09)***	(0.06)***	(0.11)***	(0.06)***	(0.09)***	(0.06)***	(0.11)***
Pop. speaking English	0.39	0.36	0.38	0.10	0.42	-0.16	0.21	-0.21	0.14	0.05	0.15	0.57
(Eng_Lang)	(0.20)**	(0.35)	(0.16)**	(0.31)	(0.15)***	(0.23)	(0.19)	(0.20)	(0.14)	(0.32)	(0.18)	(0.26)**
Pop. speaking other European languages	-0.23	0.67	-0.27	0.90	-0.53	1.27	-0.18	0.67	-0.32	1.33	-0.37	1.12
(EUR_Lang)	(0.18)	(0.17)***	(0.14)*	(0.20)***	(0.10)***	(0.18)***	(0.11)*	(0.22)***	(0.09)***	(0.16)***	(0.10)***	(0.16)***
First-stage F-test	17.5	7.7	19.5	8.5	31.0	24.8	18.7	5.0	21.9	38.7	24.2	31.0
Angrist-Pischke F-statistics (p-value)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.01	<0.001	<0.001	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	0.01		0.01		0.01		0.03		<0.001		<0.001	
Kleibergen-Paap Wald rk F statistic	6.94		7.92		23.03		2.89		24.3		23.2	
Stock-Yogo critical values 10%	13.43											
Stock-Yogo critical values 25%	5.45											
R-Square	0.64	0.76	0.56	0.74	0.47	0.58	0.45	0.71	0.44	0.63	0.52	0.64

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: Regional dummies for (i) Sub-saharan Africa, (ii) for Middle East and North Africa, (iii) for Europe and Central Asia, (iv) for North America, (v) for Latin America and the Carribean, (vi) for South-East Asia and the Pacific incl. Oceania, ; (vii) GEO, the variable for geography, which is measured as distance from the equator; (viii) constructed trade, the instrument for trade obtained from Frankel and Romer; (ix) the proportion of the population of a country that speaks English (Eng_Lang); and (x) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Missing value indicates that variable was dropped due to collinearity. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified.
The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.28: Determinants of income: Estimates using OWW data

2005 (IV): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	63	63	63	63	63
Geography	1.22	0.94	0.51	-0.06	-0.35
(GEO)	(0.64)*	(0.69)	(0.75)	(0.19)	(1.01)
Trade	-0.18	-0.07	-0.01	0.15	0.27
(LN_TRADE_WB)	(0.18)	(0.20)	(0.23)	(0.29)	(0.34)
Institutions	1.28	1.27	1.22	1.11	1.03
(Inst_rule_of_law)	(0.16)***	(0.17)***	(0.20)***	(0.25)***	(0.29)***
R-Square	0.59	0.51	0.36	0.06	0.01
Pagan Hall test (p-value)	0.36	0.38	0.37	0.23	0.19
Endogeneity test (p-value)	0.07	0.03	0.04	0.10	0.15
Hansen Test (p-value)	0.15	0.17	0.11	0.06	0.07

Notes: The dependent variable are hourly wages from the OWW database with country-specific calibration and imputation, converted into USD using official average exchanges rates 2003-2007. There are five samples for which the IV regressions are run per time period: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. The regressors are: (i) GEO, the variable for geography, which is measured as the absolute value of latitude of country divided by 90; (ii) trade, the log share of imports and exports to national GDP which is instrumented following Frankel and Romer (1999); and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index, which is instrumented following Hall and Jones (1999). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. The Pagan Hall tests of heteroskedasticity for instrumental variables (IV) estimation under the null of homoskedasticity. The endogeneity test is based on the Durbin-Wu-Hausman test, but adjusted here for heteroskedasticity. The Hansen Test follows the standard methodology.

Appendix Table 2.29: First stage estimates of two-stages least square estimates using OWW data

Dependent variable =	Trade	Institutions
	(1)	(2)
Sample size	63	63
Geography	0.57	2.09
(GEO)	(0.47)	(0.65)***
Constructed Trade	0.42	0.34
(Trade_FR_ROM)	(0.09)***	(0.13)***
Pop. speaking English	-0.35	1.49
(Eng_Lang)	(0.24)	(0.29)***
Pop. speaking other European languages	0.04	0.37
(EUR_Lang)	(0.16)	(0.27)
First-stage F-test	11.2	10.8
Angrist-Pischke F-statistics (p-value)	<0.001	<0.001
Kleibergen-Paap LM test (p-value)	0.01	
Kleibergen-Paap Wald rk F statistic	6.40	
Stock-Yogo critical values 10%	13.43	
Stock-Yogo critical values 25%	5.45	
R-Square	0.31	0.37

Notes: The dependent variable is the Rule of Law Index (Inst_rule_of_law) for even columns, and trade (LN_TRADE_WB) as share of imports and exports over nominal GDP for uneven columns. The regressors are: (i) GEO, the variable for geography, which is measured as the absolute value of latitude of country divided by 90; (ii) constructed trade, the instrument for trade obtained from Frankel and Romer; (iii) the proportion of the population of a country that speaks English (Eng_Lang); and (iv) the proportion of the population of a country that speaks any Western European Language (EUR_Lang). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. Angrist and Pischke (2009) propose a conditional first-stage F-statistic for the case of multiple endogenous variables under the null that the equation is under-identified. The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (Kleibergen and Paap, 2006). The first-stage Kleibergen-Paap Wald F-statistics is the generalization from Cragg and Donald (1993) to non-independently and -identically distributed errors. Below, we report the critical values from Stock and Yogo (2005) under the null of weak instruments, i.e. the rejection rate of r (here given as 10 percent and 25 percent) that may be tolerated if the true rejection rate should be 5%. Although critical values do not exist for the Kleibergen-Paap statistic, we follow the literature suggested in Baum, Schaffer and Stillman (2007) and applied in Bazzi and Clemens (2013), and use the Stock and Yogo critical values as point of comparison.

Appendix Table 2.30: Determinants of income: Core specifications, dynamic panel-data estimation, one step system GMM

1985-2010 (6 periods): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	32	32	32	32	32
Geography	2.32	1.71	0.78	-0.24	-0.63
(GEO)	(1.87)	(1.83)	(1.79)	(1.85)	(1.88)
Trade	0.67	0.61	0.58	0.55	0.55
(LN_TRADE_WB)	(0.27)**	(0.25)**	(0.23)***	(0.23)**	(0.23)**
Institutions	0.58	0.54	0.60	0.65	0.66
(Inst_rule_of_law)	(0.37)	(0.35)	(0.35)*	(0.36)*	(0.37)*
Arellano-Bond test for AR (1) (p-value)	0.27	0.02	0.01	0.02	0.05
Arellano-Bond test for AR (2) (p-value)	0.71	0.43	0.21	0.28	0.31
Hansen Test (p-value)	0.77	0.72	0.73	0.71	0.71

1990-2010 (5 periods): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	43	43	43	43	43
Geography	5.29	2.88	1.69	0.27	-0.27
(GEO)	(3.27)*	(3.00)	(2.89)	(2.95)	(2.99)
Trade	0.21	0.21	0.15	0.12	0.11
(LN_TRADE_WB)	(0.33)	(0.32)	(0.33)	(0.36)	(0.37)
Institutions	0.31	0.48	0.57	0.68	0.72
(Inst_rule_of_law)	(0.46)	(0.42)	(0.40)	(0.42)	(0.43)*
Arellano-Bond test for AR (1) (p-value)	0.64	0.02	<0.001	0.01	0.11
Arellano-Bond test for AR (2) (p-value)	0.90	0.29	0.19	0.20	0.19
Hansen Test (p-value)	0.12	0.09	0.10	0.13	0.13

1995-2010 (4 periods): Dependent variable = Log GDP per capita of	First Quintile	Median	Average Popula-tion	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)
Sample size	55	55	55	55	55
Geography	13.71	11.76	10.00	8.43	7.77
(GEO)	(5.37)***	(5.05)**	(4.65)**	(4.41)**	(4.32)*
Trade	1.04	0.94	0.80	0.69	0.67
(LN_TRADE_WB)	(0.74)	(0.73)	(0.67)	(0.65)	(0.64)
Institutions	-1.06	-0.91	-0.72	-0.57	-0.51
(Inst_rule_of_law)	(0.81)	(0.73)	(0.66)	(0.62)	(0.60)
Arellano-Bond test for AR (1) (p-value)	0.20	0.15	0.15	0.24	0.27
Arellano-Bond test for AR (2) (p-value)	0.39	0.29	0.26	0.28	0.17
Hansen Test (p-value)	0.38	0.11	0.10	0.23	0.30

Notes: The dependent variable is per capita GDP on PPP basis. For each specification, there are five samples for which the two-step dynamic panel-data estimations are run: (1) refer to the bottom 20% income group; (2) regress the median income; (3) refer to the average per capita GDP; (4) regress the top 20% income group; and (5) regress the top 10% income group. Three panel specifications are analyzed: 1985-2010 (six time periods), 1990-2010 (five time periods), and 1995-2010 (four time periods). The model used, known as "system GMM", is based on Arellano and Bover (1995) and Blundell and Bond (1998). Variables are used as bases for "GMM-style" instrument sets described in Holtz-Eakin, Newey, and Rosen (1988) and Arellano and Bond (1991). The regressors are: (i) GEO, the variable for geography, which is measured as distance from equator; (ii) trade, the log share of imports and exports to national GDP; and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index . See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively. The Arellano-Bond tests for autocorrelation and is applied to the differenced residuals. The Hansen Test for over-identifying restrictions follows the standard methodology.

Appendix Table 2.31: Determinants of income: Core specifications, sample split into low-income and high-income incountries

Low-income countries	1995 (linear IV)					2005 (linear IV)					Panel 1990-2010 (GMM)				
Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile	First Quintile	Median	Average Population	Top Quintile	Top Decile	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Observations	54	54	54	54	54	59	59	59	59	59	108	108	108	108	108
Geography	2.07	0.76	-0.36	-1.24	-1.61	3.80	4.25	4.09	3.95	3.69	0.38	-2.03	-3.18	-4.41	-4.75
(GEO)	(1.18)*	(1.70)	(1.97)	(2.29)	(2.37)	(2.03)*	(4.36)	(5.00)	(5.72)	(6.09)	(1.99)	(2.78)	(3.15)	(3.58)	(3.70)
Trade	-0.32	-0.05	0.12	0.21	0.25	0.89	2.14	2.54	2.96	3.20	-0.49	-0.50	-0.63	-0.74	-0.79
(LN_TRADE_WB)	(0.70)	(0.97)	(1.13)	(1.32)	(1.35)	(2.24)	(4.38)	(4.99)	(5.66)	(5.96)	(0.47)	(0.69)	(0.78)	(0.87)	(0.90)
Institutions	1.73	2.94	3.61	4.25	4.35	-2.70	-6.52	-7.48	-8.56	-9.00	2.10	3.29	3.81	4.34	4.47
(Inst_rule_of_law)	(1.62)	(2.13)	(2.39)	(2.75)	(2.85)	(5.42)	(9.30)	(10.50)	(11.81)	(12.41)	(1.49)	(2.21)	(2.51)	(2.83)	(2.91)
R-Squared	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

High-income countries	1995 (linear IV)					2005 (linear IV)					Panel 1990-2010 (GMM)				
Dependent variable = Log GDP per capita of	First Quintile	Median	Average Population	Top Quintile	Top Decile	First Quintile	Median	Average Population	Top Quintile	Top Decile	First Quintile	Median	Average Population	Top Quintile	Top Decile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Observations	53	53	53	53	53	58	58	58	58	58	107	107	107	107	107
Geography	2.45	1.36	0.60	-0.25	-0.58	1.36	0.72	0.28	-0.26	-0.47	1.07	0.52	0.09	-0.51	-0.71
(GEO)	(0.60)***	(0.49)***	(0.43)	(0.43)	(0.44)	(0.48)***	(0.38)	(0.33)	(0.31)	(0.30)	(0.69)	(0.72)	(0.76)	(0.86)	(0.90)
Trade	0.07	-0.02	-0.06	-0.11	-0.13	0.21	0.10	0.02	-0.08	-0.11	0.36	0.15	0.08	-0.05	-0.08
(LN_TRADE_WB)	(0.18)	(0.16)	(0.15)	(0.15)	(0.15)	(0.15)	(0.13)	(0.12)	(0.11)	(0.12)	(0.17)**	(0.18)	(0.19)	(0.21)	(0.21)
Institutions	0.39	0.43	0.45	0.47	0.48	0.69	0.66	0.62	0.59	0.59	0.46	0.44	0.43	0.43	0.42
(Inst_rule_of_law)	(0.21)**	(0.16)***	(0.14)***	(0.13)***	(0.14)***	(0.13)***	(0.11)***	(0.09)	(0.09)***	(0.09)***	(0.14)***	(0.14)***	(0.15)***	(0.18)**	(0.19)**
R-Squared	0.72	0.73	0.70	0.61	0.55	0.69	0.69	0.72	0.69	0.68	n/a	n/a	n/a	n/a	n/a

Notes: The dependent variable is per capita GDP on PPP basis, estimated for different income groups. The sample for low-income countries contains all countries up until and including the median GDP per capita income level of of all countries. The sample for high-income countries contains all other countries above that. Columns (1) to (10) report results from a cross-section linear 2SLS model that takes average values from a 5-year time interval. Columns (11) to (15) uses the "system GMM", based on Arellano and Bover (1995) and Blundell and Bond (1998). The regressors are: (i) GEO, the variable for geography, which is measured as distance from equator; (ii) trade, the log share of imports and exports to national GDP; and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index . See the Appendix for more detailed variable definitions and sources). "n/a" for R-squared denotes negative values, respectively "not applicable" in the GMM panel estimations. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.32: Income determinants. Base specification, ordinary least squares (OLS) estimates for all time periods

1985: Dependent variable = Log GDP per capita of	First Quintile			Median			Average Population			Top Quintile			Top Decile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Sample size	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56
Geography	4.92	5.11	1.87	4.75	4.82	1.63	4.05	4.01	1.15	3.38	3.28	0.66	3.03	2.91	0.44
(GEO)	(0.56)***	(0.68)***	(0.68)***	(0.45)***	(0.53)***	(0.62)**	(0.41)***	(0.46)***	(0.61)*	(0.40)***	(0.43)***	(0.66)	(0.39)***	(0.42)***	(0.68)
Trade		-0.20	-0.38		-0.07	-0.25		0.04	-0.13		0.10	-0.05		0.12	-0.02
(LN_TRADE_WB)		(0.32)	(0.30)		(0.27)	(0.24)		(0.24)	(0.21)		(0.24)	(0.21)		(0.23)	(0.21)
Institutions			0.83			0.82			0.74			0.67			0.64
(Inst_rule_of_law)			(0.16)***			(0.15)***			(0.14)***			(0.15)***			(0.15)***
RMSE	0.94	0.94	0.81	0.82	0.83	0.68	0.76	0.76	0.63	0.75	0.75	0.65	0.74	0.75	0.65
Adj. R-Square	0.49	0.49	0.62	0.54	0.53	0.69	0.50	0.49	0.65	0.42	0.41	0.57	0.37	0.36	0.51

1990: Dependent variable = Log GDP per capita of	First Quintile			Median			Average Population			Top Quintile			Top Decile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Sample size	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71
Geography	6.63	6.64	3.93	5.61	5.58	2.86	4.79	4.73	2.18	3.97	3.89	1.45	3.64	3.55	1.20
(GEO)	(0.51)***	(0.53)***	(0.79)***	(0.46)***	(0.47)***	(0.79)***	(0.42)***	(0.43)***	(0.76)***	(0.42)***	(0.42)***	(0.79)*	(0.42)***	(0.41)***	(0.81)
Trade		-0.03	-0.13		0.09	-0.02		0.17	0.07		0.22	0.12		0.24	0.15
(LN_TRADE_WB)		(0.19)	(0.16)		(0.20)	(0.17)		(0.20)	(0.17)		(0.21)	(0.18)		(0.21)	(0.19)
Institutions			0.68			0.68			0.64			0.61			0.59
(Inst_rule_of_law)			(0.14)***			(0.15)***			(0.15)***			(0.16)***			(0.16)***
RMSE	0.89	0.90	0.79	0.82	0.83	0.71	0.78	0.78	0.67	0.79	0.79	0.69	0.80	0.79	0.70
R-Square	0.67	0.67	0.74	0.63	0.63	0.73	0.58	0.58	0.69	0.48	0.48	0.60	0.44	0.44	0.56

1995: Dependent variable = Log GDP per capita of	First Quintile			Median			Average Population			Top Quintile			Top Decile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Sample size	107	107	107	107	107	107	107	107	107	107	107	107	107	107	107
Geography	5.68	5.67	2.94	4.90	4.90	2.35	4.15	4.16	1.71	3.42	3.43	1.12	3.10	3.11	0.87
(GEO)	(0.50)***	(0.51)***	(0.49)***	(0.46)***	(0.48)***	(0.46)***	(0.44)***	(0.45)***	(0.44)***	(0.43)***	(0.45)***	(0.46)**	(0.43)***	(0.45)***	(0.47)*
Trade		0.03	-0.03		-0.01	-0.05		-0.01	-0.06		-0.02	-0.06		-0.02	-0.06
(LN_TRADE_WB)		(0.23)	(0.16)		(0.20)	(0.13)		(0.19)	(0.13)		(0.19)	(0.14)		(0.19)	(0.15)
Institutions			0.90			0.85			0.81			0.77			0.74
(Inst_rule_of_law)			(0.09)***			(0.09)***			(0.08)***			(0.08)***			(0.08)***
RMSE	1.09	1.09	0.85	0.98	0.99	0.75	0.93	0.93	0.70	0.92	0.92	0.72	0.91	0.92	0.73
R-Square	0.49	0.49	0.69	0.47	0.46	0.69	0.41	0.41	0.66	0.33	0.32	0.59	0.29	0.28	0.55

Appendix Table 2.32 continued: Income determinants. Base specification, ordinary least squares (OLS) estimates for all time periods

2000: Dependent variable = Log GDP per capita of	First Quintile			Median			Average Population			Top Quintile			Top Decile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Sample size	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97
Geography	6.01	5.77	3.35	5.21	4.97	2.66	4.56	4.33	2.03	3.90	3.68	1.41	3.69	3.44	1.20
(GEO)	(0.43)***	(0.46)***	(0.42)***	(0.42)***	(0.44)***	(0.41)***	(0.42)***	(0.45)***	(0.43)***	(0.43)***	(0.46)***	(0.45)***	(0.44)***	(0.47)***	(0.45)***
Trade		0.29	0.13		0.29	0.14		0.28	0.13		0.26	0.11		0.29	0.15
(LN_TRADE_WB)		(0.22)	(0.13)		(0.22)	(0.14)		(0.22)	(0.15)		(0.23)	(0.16)		(0.24)	(0.17)
Institutions			0.82			0.78			0.78			0.77			0.76
(Inst_rule_of_law)			(0.08)***			(0.07)***			(0.07)***			(0.08)***			(0.08)***
RMSE	0.94	0.94	0.71	0.90	0.90	0.68	0.89	0.89	0.67	0.92	0.91	0.70	0.93	0.92	0.73
R-Square	0.61	0.62	0.78	0.57	0.57	0.76	0.50	0.51	0.72	0.41	0.42	0.65	0.38	0.39	0.62

2005: Dependent variable = Log GDP per capita of	First Quintile			Median			Average Population			Top Quintile			Top Decile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Sample size	117	117	117	117	117	117	117	117	117	117	117	117	117	117	117
Geography	5.60	5.39	3.02	5.15	4.94	2.63	4.68	4.49	2.24	4.15	3.98	1.80	3.93	3.75	1.56
(GEO)	(0.46)***	(0.51)***	(0.49)***	(0.46)***	(0.50)***	(0.47)***	(0.45)***	(0.49)***	(0.46)***	(0.45)***	(0.49)***	(0.46)***	(0.45)***	(0.49)***	(0.46)***
Trade		0.36	0.16		0.37	0.18		0.33	0.15		0.30	0.12		0.31	0.12
(LN_TRADE_WB)		(0.27)	(0.15)		(0.27)	(0.15)		(0.26)	(0.15)		(0.26)	(0.15)		(0.26)	(0.15)
Institutions			0.82			0.80			0.78			0.75			0.76
(Inst_rule_of_law)			(0.09)***			(0.09)***			(0.08)***			(0.08)***			(0.08)***
RMSE	1.01	1.00	0.75	1.01	1.00	0.77	0.99	0.98	0.76	0.99	0.98	0.78	1.01	1.00	0.80
R-Square	0.54	0.55	0.75	0.50	0.51	0.71	0.46	0.47	0.68	0.40	0.41	0.63	0.37	0.38	0.60

2010: Dependent variable = Log GDP per capita of	First Quintile			Median			Average Population			Top Quintile			Top Decile		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Sample size	91	91	91	91	91	91	91	91	91	91	91	91	91	91	91
Geography	5.19	4.90	2.57	4.59	4.29	1.94	4.07	3.77	1.48	3.45	3.16	0.92	3.21	2.92	0.71
(GEO)	(0.57)***	(0.60)***	(0.57)***	(0.56)***	(0.59)***	(0.55)***	(0.54)***	(0.56)***	(0.52)***	(0.52)***	(0.54)***	(0.51)*	(0.52)***	(0.54)***	(0.51)
Trade		0.40	0.15		0.43	0.18		0.42	0.17		0.41	0.17		0.40	0.17
(LN_TRADE_WB)		(0.25)	(0.13)		(0.26)	(0.13)		(0.25)*	(0.12)		(0.24)*	(0.12)		(0.24)*	(0.12)
Institutions			0.77			0.77			0.76			0.74			0.73
(Inst_rule_of_law)			(0.10)***			(0.09)***			(0.09)***			(0.09)***			(0.09)***
RMSE	1.00	0.98	0.76	1.00	0.98	0.75	0.97	0.95	0.73	0.97	0.95	0.74	0.97	0.95	0.74
R-Square	0.51	0.53	0.72	0.45	0.47	0.69	0.40	0.43	0.66	0.32	0.35	0.61	0.29	0.32	0.58

Notes: The dependent variable is per capita GDP in 2005, PPP basis. There are five samples for which the core regressions are run: (i) columns (1)-(3) refer to the bottom 20% income group; (ii) columns (4)-(6) regress the median income; (iii) columns (7)-(9) refer to the average per capita GDP; (iv) columns (10)-(12) regress the top 20% income group; and (v) columns (13)-(15) regress the top 10% income group. The regressors are: (i) GEO, the variable for geography, which is measured as the absolute value of latitude of country divided by 90; (ii) trade, the log share of imports and exports to national GDP; and (iii) Institutions (Inst_rule_of_law), taken from the Rule of Law Index. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 2.33: Data and Sources

Variable Name	Description	Source	Available Years	Remarks
Country	Name of country	Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" available for download at www.ggdc.net/pwt	1983-2012	Czechoslovakia was continued as Czech Republic after 1989. Germany classified as "West Germany" until 1989.
First Quintile	Log of GDP per capita of first income quintile per time period (Output-side real GDP at current PPPs)	Own construction, based on: Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" available for download at www.ggdc.net/pwt, combined with World Income Inequality Database (WIID3.0B) available at http://www.wider.unu.edu/research/WIID-3b/en_GB/database/. The alternative robustness specification uses the Occupational Wages around the World (OWW) database, which are derived from the ILO October Inquiry database, and are available for download at http://www.nber.org/oww/. Income shares are then constructed by splitting the overall available wage distribution per country in the respective percentiles examined.	1983-2012 for core specification using Penn World Tables and WIID / 2003-2007 for OWW data	
Median	Median of all quintile logs of GDP per capita per time period (Output-side real GDP at current PPPs)			
Average Population	Simple average of log of GDP per capita per time period (Output-side real GDP at current PPPs)			
Top Quintile	Log of GDP per capita of fifth income quintile per time period (Output-side real GDP at current PPPs)			
Top Decile	Log of GDP per capita of tenth income decile per time period (Output-side real GDP at current PPPs)			
Inst_rule_of_law	Rule of Law index (from World Governance Indicators).	World Bank. http://info.worldbank.org/governance/wgi/index.aspx#home	1983-2012	1996 approximates for all 1983-1996. 1997, 1999, 2001 taken as average of the 1996 and 1998, 1998 and 2000, and 2000 and 2002, respectively.
Eng_Lang	Fraction of the population speaking English as mother tongue	Hall, R., & Jones, C.I. (1999). Why Do Some Countries Produce So Much More per Worker than Others? *Quarterly Journal of Economics, 114* , 83-116.	1983-2012	
EUR_Lang	Fraction of the population speaking one of the major languages of Western Europe as mother tongue: English, French, German, Portuguese, or Spanish	Hall, R., & Jones, C.I. (1999). Why Do Some Countries Produce So Much More per Worker than Others? *Quarterly Journal of Economics, 114* , 83-116.	1983-2012	
Inst_sett_mort	Log of settler mortalities in European colonies	Acemoglu, D., Johnson, S., & Robinson, J.A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *The American Economic Review, 91* (5), 1369-1401.	1983-2012	
LN_Trade_WB	Log of nominal trade per country (sum of exports and imports of goods and services measured as a share of gross domestic product).	World Bank national accounts data, and OECD National Accounts data files. http://data.worldbank.org/indicator/NE.TRD.GNFS.ZS?cid=GPD_31	1983-2012	
Trade_FR_ROM	Log of predicted trade shares computed following Frankel and Romer	Frankel, J.A., & Romer, D. (1999). Does Trade Cause Growth? The American Economic Review, 89(3), 379-399. Own construction for missing countries	1985 (fixed)	Instrumental variable for trade
HC_schoolenr_70	Log of primary gross school enrollment rates, averaging 1970-1979	World Bank, UNESCO Institute for Statistics	1970-1979	
HC_schoolenr_85	Log of primary gross school enrollment rates, averaging 1983-1987		1983-1987	
HC_schoolenr_90	Log of primary gross school enrollment rates, averaging 1988-1992		1988-1992	
HC_schoolenr_95	Log of primary gross school enrollment rates, averaging 1993-1997		1993-1997	
HC_schoolenr_00	Log of primary gross school enrollment rates, averaging 1998-2002		1998-2002	
HC_schoolenr_05	Log of primary gross school enrollment rates, averaging 2003-2007		2003-2007	
HC_schoolenr_10	Log of primary gross school enrollment rates, averaging 2008-2012		2008-2012	
GEO_disteq	Mean distance to equator, measured as abs(latitude of country centroid)/90	Own construction, based on John L. Gallup, Andrew D. Mellinger, and Jeffrey D. Sachs' Geography Datasets; http://www.cid.harvard.edu/ciddata/geographydata.htm	1983-2012	
Malaria	Malaria Index 1994	Gallup and Sachs (1998)	1983-2012	
Meantemp	Average temperature in given country (Celsius)	CID Harvard University (2002)	1983-2012	
Health_lifeexp	Life expectancy at birth in 1970 (number of years)	World Bank World Development Indicators	1970	
Ethnoling_frac	Ethnic fractionalization using Ethnicity data points between 1979-2001	Alesina, A., Devleeschauwer, A., Easterly, W. Kurlat, S., & Wacziarg, R. (2003). Fractionalization. Journal of Economic Growth, 8, 155-194.	1979-2001	Available values per country used as proxy for all time periods
EUR_colony	Dummy variable taking value 1 if country had a European colonizer, 0 otherwise	Own construction based on World Bank definition of world regions	1983-2012	
Subsah_AFR	Dummy variable taking value 1 if country is located in Sub-Saharan Africa, 0 otherwise			
MEast_Nafr	Dummy variable taking value 1 if country is located in Middle East or North Africa, 0 otherwise			
Eur_Asia	Dummy variable taking value 1 if country is located in Europe or Central Asia, 0 otherwise			
North_Am	Dummy variable taking value 1 if country is located in North America, 0 otherwise			
LatAm	Dummy variable taking value 1 if country is located in Latin America, 0 otherwise			
SE_Asia	Dummy variable taking value 1 if country is located in South or South-East Asia, 0 otherwise			
GINI	Measure of statistical dispersion to represent the income distribution within a country; commonly used measure of inequality.	World Bank World Development Indicators		
WTID	World Top Incomes Database	Alvaredo et al. (2015), retrieved from http://topincomes.parisschoolofeconomics.eu/#Database:	1983-2012	

Appendix Table 2.34: Overview of countries per time period in regular specification using WIID data

1983-1987	1988-1992	1993-1997	1998-2002	2003-2007	2008-2012
56	*71*	*107*	*84*	*117*	*91*
Argentina	Argentina	Argentina	Angola	Argentina	Angola
Australia	Australia	Armenia	Argentina	Armenia	Argentina
Austria	Bangladesh	Austria	Austria	Australia	Armenia
Bangladesh	Belgium	Bangladesh	Bangladesh	Austria	Australia
Belgium	Bolivia	Belarus	Belgium	Bangladesh	Austria
Bolivia	Brazil	Belgium	Belize	Belarus	Bangladesh
Botswana	Bulgaria	Belize	Bolivia	Belgium	Barbados
Brazil	Canada	Bolivia	Brazil	Benin	Belarus
Canada	Chile	Brazil	Bulgaria	Bhutan	Belgium
Chile	Colombia	Bulgaria	Burkina Faso	Bolivia	Bhutan
Costa Rica	Costa Rica	Burkina Faso	Burundi	Botswana	Bolivia
Cote d`Ivoire	Cote d`Ivoire	Cameroon	Cambodia	Brazil	Brazil
Denmark	Czech Republic	Cambodia	Cameroon	Bulgaria	Bulgaria
Dominican Republic	Denmark	Canada	Canada	Burkina Faso	Burkina Faso
Ecuador	Dominican Republic	Chile	Cape Verde	Cambodia	Cambodia
Finland	Ecuador	China	Chile	Canada	Canada
France	Egypt	Colombia	China	Central African Republic	Central African Republic
Germany	El Salvador	Costa Rica	Colombia	Chile	Chile
Ghana	Finland	Cote d`Ivoire	Costa Rica	China	China
Guatemala	France	Czech Republic	Cote d`Ivoire	Colombia	Colombia
Honduras	Gambia, The	Denmark	Czech Republic	Comoros	Costa Rica
Hungary	Germany	Djibouti	Denmark	Congo	Cote d`Ivoire
India	Ghana	Dominican Republic	Dominican Republic	Costa Rica	Croatia
Indonesia	Guatemala	Ecuador	Ecuador	Cyprus	Cyprus
Iran	Guinea	Egypt	El Salvador	Czech Republic	Czech Republic
Ireland	Honduras	El Salvador	Ethiopia	Dem. Rep. Congo	Denmark
Israel	Hungary	Estonia	Finland	Denmark	Dominican Republic
Italy	India	Ethiopia	France	Dominican Republic	Ecuador
Jordan	Indonesia	Finland	Gambia, The	Ecuador	Egypt
Korea, Republic of	Iran	France	Germany	Egypt	El Salvador
Lesotho	Israel	Gambia, The	Ghana	El Salvador	Estonia
Luxembourg	Italy	Georgia	Greece	Estonia	Ethiopia
Malawi	Jamaica	Germany	Guatemala	Ethiopia	Fiji
Malaysia	Jordan	Ghana	Guinea-Bissau	Fiji	Finland
Mauritania	Kenya	Guinea-Bissau	Honduras	Finland	France
Mexico	Korea, Republic of	Greece	Hong Kong	France	Germany
Morocco	Luxembourg	Guinea	Hungary	Gabon	Greece
Nepal	Malaysia	Honduras	India	Gambia, The	Honduras
Netherlands	Mali	Hong Kong	Iran	Georgia	Hong Kong
New Zealand	Mauritania	Hungary	Ireland	Germany	Hungary
Nigeria	Mexico	Indonesia	Israel	Ghana	Iceland
Norway	Morocco	Iran	Italy	Greece	India
Pakistan	Netherlands	Ireland	Jamaica	Guatemala	Ireland
Paraguay	New Zealand	Israel	Kenya	Guinea	Italy
Peru	Niger	Italy	Korea, Republic of	Honduras	Japan
Philippines	Nigeria	Jamaica	Laos	Hong Kong	Jordan
Poland	Norway	Japan	Luxembourg	Hungary	Kazakhstan
Sri Lanka	Pakistan	Jordan	Madagascar	Iceland	Kyrgyzstan
Sweden	Panama	Kazakhstan	Malawi	India	Laos
Thailand	Paraguay	Kenya	Mali	Iran	Latvia
Tunisia	Peru	Korea, Republic of	Mauritania	Iraq	Lithuania
Turkey	Philippines	Kyrgyzstan	Mexico	Ireland	Luxembourg
United Kingdom	Poland	Laos	Mongolia	Israel	Macedonia
United States	Portugal	Latvia	Morocco	Italy	Madagascar
Uruguay	Romania	Lesotho		Jamaica	Malawi
Venezuela	Russia	Lithuania		Japan	Malaysia
	Senegal	Luxembourg		Jordan	Mali
	Spain	Macedonia		Kazakhstan	Mexico
	Sri Lanka	Madagascar		Kenya	Moldova
	Sweden	Malaysia		Kyrgyzstan	Namibia
	Switzerland	Mali		Latvia	Nepal

Appendix Table 2.34 continued: Overview of countries per time period in regular specification using WIID data

	1988-1992	1993-1997	1998-2002	2003-2007	2008-2012
	Tanzania	Mauritania	Netherlands	Lesotho	Netherlands
	Thailand	Mexico	Norway	Liberia	Niger
	Tunisia	Moldova	Pakistan	Lithuania	Norway
	Uganda	Mongolia	Panama	Luxembourg	Pakistan
	United Kingdom	Morocco	Paraguay	Macedonia	Panama
	United States	Mozambique	Peru	Madagascar	Paraguay
	Uruguay	Namibia	Philippines	Malawi	Peru
	Venezuela	Nepal	Poland	Malaysia	Philippines
	Yemen	Netherlands	Portugal	Mali	Poland
	Zambia	New Zealand	Romania	Mauritania	Portugal
		Niger	Russia	Mauritius	Romania
		Nigeria	Senegal	Mexico	Russia
		Norway	South Africa	Moldova	Singapore
		Pakistan	Spain	Mozambique	Slovak Republic
		Panama	Sri Lanka	Namibia	Slovenia
		Paraguay	Suriname	Nepal	South Africa
		Peru	Sweden	Netherlands	Spain
		Philippines	Switzerland	Niger	Sudan
		Poland	Taiwan	Nigeria	Sweden
		Portugal	Tanzania	Norway	Switzerland
		Romania	Thailand	Pakistan	Taiwan
		Russia	Tunisia	Panama	Thailand
		Senegal	Uganda	Paraguay	Turkey
		Slovak Republic	United Kingdom	Peru	Uganda
		Slovenia	United States	Philippines	Ukraine
		South Africa	Uruguay	Poland	United Kingdom
		Spain	Venezuela	Portugal	United States
		Sri Lanka	Vietnam	Romania	Uruguay
		St. Lucia	Yemen	Russia	Venezuela
		Swaziland	Zambia	Rwanda	Vietnam
		Sweden		Senegal	
		Taiwan		Singapore	
		Tanzania		Slovak Republic	
		Thailand		Slovenia	
		Tunisia		South Africa	
		Turkey		Spain	
		Uganda		Sri Lanka	
		Ukraine		Sweden	
		United Kingdom		Switzerland	
		United States		Syria	
		Uruguay		Taiwan	
		Uzbekistan		Tanzania	
		Venezuela		Thailand	
		Vietnam		Togo	
		Zambia		Tunisia	
		Zimbabwe		Turkey	
				Uganda	
				Ukraine	
				United Kingdom	
				United States	
				Uruguay	
				Uzbekistan	
				Venezuela	
				Vietnam	
				Yemen	
				Zambia	

III. Marriage Age Affects Educational Gender Inequality*

ABSTRACT

This chapter examines the effect of female age at marriage on female education and educational gender inequality. We provide empirical evidence that early female marriage age significantly decreases female education with panel data from 1980 to 2010. Socio-cultural customs serve as an exogenous identification for female age at marriage. We also show that effects of spousal age gaps between men and women significantly affect female education relative to male education. Each additional year between husband and wife reduces the female secondary schooling completion rate by 10 percentage points, the time women spend at university by one month, and overall affects female education significantly more negatively than male education. We also document that marriage age and conventional measures of gender discrimination do not act as substitutes.

III. 1. INTRODUCTION

Improving access to education for women is a central theme in economic development (United Nations, 2015). Not only are equal educational opportunities a pressing issue for the many disadvantaged women around the world; there is also evidence for the positive role of female human capital on economic development (Esteve-Volvart, 2000; Klasen 1999, 2002; Schultz, 2002; Sedgley & Elmslie, 2005; Todaro & Smith, 2014). Not only scholars and international institutions argue for positive overall effects of women's equality, even a report by McKinsey & Company (2015), a consultancy, suggests that $12 trillion could be added in 2025 to annual global GDP by bringing the gender parity level around the world "only" to the best-in-region country.

While we observe that the global educational gap is gradually shrinking, girls have still not caught up to boys and do not realize their full human capital potential (United Nations, 2015). The reasons for this educational gap may also relate to cultural customs and traditions that are not compatible with the idea of a highly educated female and male population. This work contributes

* A working paper of this chapter is circulating (Stimpfle & Stadelmann, 2016a) and has been presented at the 2016 Annual Conference of the German Economic Association (Verein für Socialpolitik), and the 2016 Annual Meeting of the European Public Choice Society (EPCS), where helpful feedback was provided by Alexander Fink. We are also grateful to Hartmut Egger, Mark Schelker, Elena Groß and Mario Larch for useful and encouraging comments at the Graduate Research Seminar of the University of Bayreuth.

to that field of research by linking gender inequality in education with marriage age. Marriage, marital customs and traditions of founding a family are a central cultural feature of different societies. The timing for marriage and, in particular, marriage age of women with respect to men can lead to significant economic ramifications for investments in education and human capital of the different sexes.[32] Differences in education and human capital in turn affect overall economic development (Barro, 2001; Lucas, 1988). There is only comparatively scarce international evidence on economic consequences of early female marital age in particular regarding effects of age of marriage on educational achievements.

Our analysis attempts to fill this gap in the literature, as we examine whether the marriage age for women, i.e. at which age the bride gets married and how that compares to her husband's age, matters for female educational prospects. We document that women get married at relatively young age in many countries and almost everywhere around the world at a considerably younger age than men, i.e. wives are usually younger than husbands and there often exists a considerable spousal age gap. The marriage age of women and the phenomenon of wives being on average younger than husbands impacts educational investment and we empirical identify it as a relevant factor for determining female education.

We provide a basic economic framework to conceptualize how a woman's timing of marriage affects her educational decision. Societal expectations of marriage age signal the timing for child-rearing, as marriage is usually the first and still the socially most accepted institution for conceiving children. Anticipated family offspring affects future female labor force participation, since, for given societal conventions, wives tend to be more often in charge of raising children which impacts their educational pay-off in the job market. The timing of marriage is related to exogenous socio-cultural customs but influence individual decisions on marriage. The earlier a women gets married, the shorter her anticipated pay-off to educational investments such that educational investments are lower for younger marriage ages than for older marriage ages. We thus hypothesize that for countries where women get married younger, their achieved level of education is likely to be lower.

---

[32] A recent article in The Economist, January 23rd 2016, mentions the story of Aisha Abduallai, a girl from Nigeria as an indicative example: "[…] so they [the parents] marry me off. He was 50 and I was 13. […] [Aisha's] education ended abruptly."

We employ a global panel data set from 1980-2010 to analyze the relationship between marriage age and educational achievements. Empirical results indicate that the absolute female age at marriage has a theory-consistent and highly significant effect on female education: Each year of marriage postponement for women is associated with a 3%-points higher female completion rate in secondary schooling, and to about three weeks, or 13% longer female tertiary education. To take account of endogeneity issues, we first employ fixed effects and different instrumentation strategies: We explain the culture-induced domestic female marriage age with a weighted average of the marriage age in adjacent countries and other instruments employed in the literature. Second, we investigate the effects of spousal age gaps, i.e. the female relative to the male marriage age. Introducing this variable tackles potential biases in case women in one part of the world get married earlier than somewhere else, irrespective of education levels. Our empirical results become even stronger: Each additional year difference between wife and husband reduces the female secondary schooling completion rate by 10%-points and the time women spend at university by one month. Finally, we employ a quasi difference-in-difference strategy to focus on differences between women and men regarding marriage age and educational achievement, i.e. we specifically examine spousal age gaps and educational gaps. This approach helps to eliminate potential confounding factors that affect the level of educational achievements jointly for women and men as we focus only on the differences between the two sexes. We show that spousal age gaps affect female education significantly more negatively than male education.

Numerous robustness tests support our main empirical findings. Further refinements and differential analyses suggest that gender parity in literacy, but not primary schooling, is affected by women's marriage age which is consistent with our theoretical considerations. The marriage age tradition of the parental generation also has an influence on current educational inequality. Importantly, we document that marriage age is no substitute measure for conventional indicators of gender discrimination, i.e. female marriage age consistently and significantly affects educational achievements of women independently of existing levels of other gender discrimination in society. The remainder of this chapter is organized as follows: Section 2 discusses the related literature and our conceptual framework. The data and identification strategy is presented in Section 3. We present our main empirical results and instrumental variable strategies in Section 4, and discuss a set of robustness tests and refinements in Section 5. Section 6 offers concluding remarks.

III. 2. Related Literature and Theoretical Considerations

Related Literature

This study relates to three strands of literature. First, we complement the literature on educational attainment of women and men. As the importance of human capital for economic development is ever more highlighted (Lucas, 2015), numerous studies have investigated the impact of educational gender inequality on growth and found generally negative effects. Esteve-Volart (2004) argues that educational gender inequality is an inefficient practice, as distortions in the allocation of talent and human capital investment lower economic growth. Some empirical analyses, however, tend to find opposite effects (Barro & Sala-i-Martin, 1995). In work by Barro and Lee (1994), higher male and female education positively affects a number of variables that indirectly foster development, such as life expectancy and child mortality. However, regression results suggest that female secondary schooling affects a country's growth rate negatively, while male secondary schooling has a significantly positive effect, an outcome for which they "do not have a convincing story" as explanation (ibid., p. 22). Easterly (2007), on the other hand, reaffirms that educational inequality represents a mechanism that hinders growth, and Klasen and Lamanna (2009) conclude from cross-country and panel regressions that gender gaps in education and employment significantly reduce economic growth. Knowles, Lorgelly, and Owen (2002) also find that that educational gender gaps are an impediment to economic development. Hill and King (1995) furthermore emphasize the beneficial aggregate social effects for growth when advancing female education. Apart from higher female productivity based on more equal education between genders, social benefits also promote growth indirectly, for example through extending the average life expectancy in a population or improving the political systems and processes.

Nevertheless, lower education for women in comparison to men remains a widespread phenomenon (Pekkarinen, 2012 surveys the literature). Alexander and Eckland (1974) have early documented gender gaps in education in a study on the U.S. Hyde, Fennema and Lamon (1990) summarize from American mathematics performance tests that gender differences emerge only in high school and college. A further study by Guiso, Monte, Sapienza, and Zingales (2008) finds that girls' underperformance in math relative to boys may be eliminated as a society becomes more gender-equal. Given such a societal trend, girls perform as well as boys in mathematics, and much better than boys in reading. Barro and Lee (2013) provide further encouraging empirical results, as significant progress has been made in developing countries towards reducing educational gender inequality. Yet, the authors emphasize that many challenges in making education more inclusive remain. An important driver for this persistent educational gender inequality is thought to relate to socio-cultural institutions. We show that educational attainment of women is systematically influenced by a core socio-cultural institution at the macro level, namely female age of marriage and the differences between female and male age of marriage. Morrison, Raju, and Sinha (2007) argue that the impact of gender equality on human development at the macro level is less well understood than at the individual level.

Second, our work adds to the literature on spousal age gaps, which suggests income prospects and fecundity, among others, as key explanatory variables.[33] Vella and Collins (1990) develop a model that proposes a link between income difference and age difference, as males and females are willing to trade youth for income. In general, the model suggests a positive age differential in favor of the husband due to biological differences, and the authors confirm their theoretical predictions with marital records from the U.S. H. Zhang (2014) develops a theoretical model which incorporates several variables to explain the husband-wife age gap. The model uses stochastic returns from human capital investments and emphasizes the interrelatedness between marriage timing, education levels, and labor market outcomes. It also builds in differential fecundity, i.e. depending on the expected number of children, women have more or less urgency to plan ahead and marry early.

The fecundity horizon is also the key explanatory variable in the model by Díaz-Giménez and Giolito (2013), who argue that the spousal age gap will persist even if gender wage gaps disappear. This is because, given regular reproductive conditions, shorter biological clocks force women to be less choosy than men of the same age, i.e. women are willing to marry older men because delaying marriage is too costly. These stable dynamics may only change through exogenous advances in reproductive technologies. X. Zhang (2014) confirms in his empirical study that the asymmetric fecundity horizon and the demand for children are driving forces for the spousal age gap. From cross-country regressions she also concludes that spousal age gaps are positively associated with the role children play in economic activities, the importance of the

[33] There are of course many more explanatory variables suggested in the literature. Edlund (1999), for example, lists unbalanced sex ratios, social status, and the functioning of capital markets (respectively ease of borrowing) as determinants of a spousal age gap. Li (2008) suggests that the gender life expectancy gap in favor of women tends to decrease the spousal age gap.

agricultural sector, and the share of rural population. Mansour and McKinnish (2014) suggest that lower occupational wages drive up the spousal age gap, i.e. negative selection causes the husband-wife gap to increase. The authors argue that better educated individuals with more economic opportunities rather marry similarly-aged spouses so that the ones "left behind" are more likely to marry someone with a larger age difference.

Anderberg, Hener, and Wilson (2014) conclude from individual data that more education leads to a smaller age gap, driven by the bride's later average age of marriage. They exploit the exogenous variation in qualifications induced by a legislative change to the minimum compulsory school leaving age in England and Wales in 1972 and find that while a spousal age gap is generally found desirable by individuals, that gap is attenuated by a higher level of education.

In a case study on Indonesia, Utomo (2014) reports a comparable association between a higher level of the wife's education and a decreasing spousal age gap, but limits her conclusion to a correlation. Similar results for the link between education and marriage age are found by Carmichael (2011) for a panel data set of 77 developing countries from 1950 to 2005[34], by Danziger and Neuman (1999) for a study on Muslims and Jews in Israel, by Gustafson and Fransson (2015) for Swedish marital data, by Glick, Handy, and Sahn (2015) for a case study on Madagascar, and by Garenne (2004) for sub-Saharan Africa. Issues of endogeneity have been pointed out in strands of the literature (Lise & Seitz, 2007; Matz, 2013; van der Vleuten, 2013), and causal links may be running from marriage age (due to established societal conventions) to female education outcomes. Casterline, Williams, and McDonald (1986, p. 354) note that "any systematic effort to test […] explanations is undermined by the fact that the [spousal] age difference can exert an independent influence on the forces suggested as its causes." Similarly, Mensch, Singh, and Casterline (2005) remark that there are nearly no studies that economists would consider acceptable in addressing the endogeneity problems around determinants of marriage timing.

Our contribution explicitly addresses the possibility that marriage age affects education, and our evidence suggests that societal conventions related to marriage timing can explain differences in female education. Moreover, analyzing the (absolute) age at marriage for women may overlook potential biases. If couples in a certain region habitually marry at younger ages than what the

[34] Moreover, she finds no clear evidence for the role of Islam on marriage age, while urbanization is significantly positive for marriage age and negative for spousal age gaps.

global average suggests, one might draw incorrect conclusions from examining absolute age levels only. We thus also systematically analyze the relative marriage age (spousal age gap) and show that the spousal age gap affects female education significantly more than male education.

Third, we also contribute to the literature relating marriage to differential gender outcomes and discrimination. Sociological literature in this field alludes to societal expectations and gender discrimination associated with female marriage age and husband-wife age gaps (Banks & Arnold, 2001; Blood & Wolfe, 1960; Freud, 1962; Lehmiller & Agnew, 2008), although Bhrolcháin (1992) cautions that trends in the age difference are not readily interpretable as reflecting change in the relative status of the sexes. Jensen and Thornton (2003) argue that women who marry young are more likely to experience domestic violence, whilst Lise and Seitz (2011) conclude that the spousal age gap has an economically and statistically insignificant effect on intra-household income allocation. The strand of literature presented here discusses the "internal complexity and the variety of social contexts" that shape husband-wife age differences (Pyke & Adams, 2010, p. 770), but does only partly support its conclusions and hypotheses with quantitative evidence. Our contribution adds to these analyses and provides international evidence which link societal expectations, resulting economic incentives, female age at marriage, and human capital investments.

Thereby, we also directly contribute to the economics of marriage. Economic perspectives on marriage have received attention in academia, following the seminal two-part articles by Becker (1973, 1974). The following papers represent no exhaustive overview but appear particularly suitable in our research context. Korenman and Neumark (1990) find no empirical support for the contention that marriage lowers women's wages. On the other hand, in a related paper (Korenman & Neumark, 1991), the authors find that wages for married men rise, even after controlling for human capital details such as actual labor market experience, and that such a marriage wage premium is persistent. Polachek and Xiang (2006) find more specifically that the husband-wife age gap increases the gender pay gap. In general, the literature suggests that in countries with larger spousal age gaps, women have lower incentives in the labor market since their older husbands are likely to have accumulated more wealth and higher wages. Matz (2013) estimates a negative impact of the spousal age gap on Ethiopian household incomes as cooperation between spouses may be impeded by large age differences. Spousal age gaps are also considered to affect fertility levels and, consequently, population growth (Casterline et al., 1986; Hajnal, 1965; van Zanden, 2011). Results from a study by Saardchom and Lemaire (2005), where marriage patterns between males and females are compared, indicate that social and cultural variables play a more important role than economic modernization variables. Fernández, Guner, and Knowles (2005) develop a model on marriage and inequality, and document empirically that the extent of marital sorting is negatively associated with per capita income levels, and greater gender discrimination is associated with more sorting. Rao (1993) suggests that spousal age gaps could be behind this century’s rise in dowries in South Asia, and Banerjee, Duflo, Ghatak, and Lafortune (2013) present a study on India that outlines a model on why economic forces have not been able to undermine the importance of in-caste marriage.

A stream of literature more explicitly connects gender inequality, human capital and marriage. Mincer and Polachek (1974) focus on the relation between human capital investments and observed market earnings of women within a family. Relatively shorter lifetime labor market participation of married women (mothers) implies lower returns on human capital investments than for married men (fathers), and hence yields persistent gender inequality. Goldin and Katz (2002) discuss how the introduction of the birth control pill enabled relatively more women in the U.S. to pursue higher education by indirectly lowering the cost of career investment. The reduced gender inequality due to the pill was additionally reinforced by its impact on delaying the average age at first marriage. Goldin, Katz, and Kuziemko (2006) echo the relationship between later age of marriage and a narrowing of the educational gender gap, in particular because the rising trend for the age at first marriage for both sexes has affected women overproportionately. In a careful historical analysis, Goldin (2006) tracks gender inequality in the U.S. especially in education and labor market participation throughout the twentieth century. She finds that for the last fifty years, in a “quiet revolution” women have continuously increased their investments in formal schooling and career building, since they have acquired longer planning horizons and an altered attitude towards marriage. A paper by Iyigun and Walsh (2007) argues that higher investment in female education raises women’s bargaining power within a marriage, which reduces gender inequality in households and ultimately makes marriage integral to the process of economic development. Finally, Chiappori, Iyigun, and Weiss (2009) propose a framework for the determination of pre-marital schooling and marriage patterns of men and women. Their approach suggests that investments in education are taken in anticipation of being married to a spouse with whom one can share consumption and coordinate work activities, a feature from which again women benefit disproportionately much.

Building on the general literature consensus of a positive association between female education and female marriage age, we specifically analyze causal effects of female age at marriage on female human capital in a global perspective, i.e. we investigate effects in a variety of development stages and economic settings. This helps examine whether such a relationship exists independent of national boundary conditions. We also aim to advance beyond a theoretical model, as we provide empirical evidence in addition to our conceptual considerations.

We are interested in female education levels, because educational gender discrimination matters not only for the directly affected girls, but also for the nation on a macroeconomic level (Kabeer & Natali, 2013). Studies show that investments in the education of young females can lead to outstanding returns (Psacharopoulos, 1988). A better trained women's labor force and higher labor force participation fuel growth through increased female productivity rates and earnings abilities (Dollar & Gatti, 1999). Dougherty (2005) finds that female education even has a double beneficial effect on women’s earnings, since besides higher productivity it also reduces the gender wage gap that is related to discrimination, tastes, and circumstances. A more equal distribution of human capital in the population also leads to more allocative efficiency of the work force as shown by Lagerlöf (2003). He develops a model that explains important changes in economic growth in Europe throughout economic history with developments in gender equality. Seguino and Floro (2003) find from panel data analyses for a set of semi-industrialized economies between 1975 and 1995, that an increase in the women’s wage share relative to that of men is associated with an increase in the domestic savings rate. Furthermore, beneficial generational effects have been proposed as children of more educated women display overall better well-being and higher productivity (Basu, 2002).[35]

Four received papers explore the causal effects of marriage age on education, and are hence closely linked to our contribution: In an early paper on the U.S., Marini (1978) concludes that women's earlier age at marriage is critical for limiting their educational attainment. Age at first marriage has no significant effect on the educational attainments of men but has a strong impact on female educational attainment. Foreman-Peck (2011) suggests from European historical

[35] For additional theoretical literature on this topic, see for example Lagerlöf (2003), Galor and Weil (1996), Dollar and Gatti (1999), Knowles et al. (2002)

evidence that later marriage sets up a virtuous cycle as it allows more female education and ultimately spurs economic growth. Field and Ambrus (2008) find for individuals in Bangladesh that each additional year that marriage is delayed for females is associated with an increase in years of their schooling and higher female literacy rate. Maertens (2013), based on her sample of three villages in India, argues similarly that educational gender inequality can be traced back to female marriage timing. However, this literature does not examine spousal age gaps, the case study character is relevant but does not provide international evidence, and the impact of other "regular" gender discrimination variables is not explicitly examined. Thus, to the best of our knowledge this work provides for the first time in the literature macro-evidence from international panel data via two alternative identification mechanisms, while also taking account of other gender discrimination variables. A final motivation to study how female education is affected by marriage age also lies in the fact that marriage age could potentially be regulated by age of marital consent laws and intensive public information campaigns may affect societal conventions in the long run too.

Conceptual Framework for the Gender-Specific Effect of Marriage Age

We consider a simple theoretical framework to better understand the economic rationale linking age at marriage and education of women. As argued in the literature (DiMaggio, 1994; Huntington, 1996; Inglehart & Welzel, 2005), we assume that cultural influences are robust, and cultural habits in a society adjust slowly over generations if at all, such that certain socio-cultural characteristics are given.[36]

For our purpose, the most relevant cultural dimension consists of the societal expectations and conventions on when to get married as a woman. These expectations belong to the decisive factors for the actual timing of a woman's marriage. Even if an individual young woman might not feel fully "ready" for marriage, established socio-cultural customs and resulting societal pressure may overrule personal sentiment. Societal norms and customs have also been repeatedly cited in the literature to explain observed female marriage age (Caldwell, Reddy, & Caldwell, 1983; Maertens, 2013; Mason & Smith, 2003; Mensch et al., 2005; Srinivas, 2000). Expected female marriage age due to societal expectations hence represents an important factor for female

[36] This is closely related to Cervellati and Sunde (2005), who develop a model for human capital accumulation based on expected lifetime.

life planning, and affects the incentives for individuals. The known expectations on marriage age allow for nearly perfect foresight planning, so that we may assume economically rational, and hence identical behavior of all family members (the girl and her parents).

Societal expectations and socio-cultural factors also apply to male marriage age. However, marriage age expectations affect a future wife much more than a future husband due to two important reasons: First, marriage represents the main socially accepted institution for conceiving children. We see that, globally and independent of cultural background, marriage is considered not an end in itself, but serves, as documented through a close temporal link, to begin childbearing.[37] For a global sample, Jensen and Thornton (2003) empirically document a continuous relationship for women between their marriage age and age at first birth, i.e. the older the bride is, the older she will become a mother. The authors explain this result as well with social norms, which emphasize the importance of child-birth taking place within marriage. Foreman-Peck (2011) also establishes the close link between marriage and the timing of first childbirth in a theoretical model. This pattern is even more reinforced still today in many cultures that are concerned with preserving a woman's virginity until marriage (Mensch et al., 2005), a behavior which is influenced also by economic rationale (Mariani, 2012). We hence assume that marriage has a signaling effect for conceiving children, which is particularly relevant for women.

Second, societal conventions differentiate between the time women and men invest into raising children, and usually dictate that mothers do most of the job (Bianchi, Milkie, Sayer, & Robinson, 2000; Levant, Slattery, & Loiselle, 1987; OECD, 2011; Sinno & Killen, 2009). Even in very gender-equal societies such as Sweden, women still use the majority of days for parental allowance (Statistics Sweden, 2014), and critical tasks such as giving birth are linked to females by nature. Wives hence tend to be more affected since global cultural customs put a higher emphasis on mothers to raise children than on fathers. This may in turn have a gender-specific effect on the labor market population. Mothers are not only the primary affected agent in the weeks and months before the date of delivery; the birth also indicates additional years of female work mostly dedicated to raising this child. The labor force of a man, in contrast, is less affected by becoming a father. Even in countries that may be regarded as most gender-egalitarian, the father's

[37] Malthus (1830) already observed that a prudential restraint on marriage, i.e. a later marriage age, would lead to lower birth rates and therefore act as demographic control, since the timing for the first child is pushed backwards.

involvement in raising children is usually confined to some weeks or a few months' time (Brandth & Kvande, 2016; Monna & Gauthier, 2008; OECD, 2011).

In essence, the foresight of expected marriage equaling expected offspring yields different incentive patterns for women versus men which should be empirically observable. The former know in advance that their labor force will be relatively more tied to raising children, whereas the latter expect less of an effect on their labor force trajectory. Out of economic rationale, girls (as well as their parents) know the expected marriage age affects the number of years they can be employed (before a child is conceived). If girls get married very young, the projected participation in the labor market approaches zero (child birth is "imminent"), and later entry less likely due to a missing previous job experience. In line with our theoretical reasoning, Miller (2011) documents that women's careers benefit from delaying the first child. In addition, the higher the socially expected number of children, the more unlikely would be a potential return into the labor market after the birth of the first child.

Finally, with this incentive scheme of labor market prospects in mind, we hypothesize that the socially expected female marriage age determines how much a girl initially intends to invest in education. Related work has suggested that the level of educational investments depends on the expected returns (Foster & Rosenzweig, 2007; Ngyuen 2007). As women will usually be married not before their first menstrual cycle, we define educational investments as forms of schooling beyond primary education. Also, we assume that individuals maximize their utility through efficient educational investment. Hence, only the amount of education is invested into which is required for expected successful labor market participation. Education obtained for "personal wisdom" that lacks economic pay-off is supposed to be irrelevant (or not differentially relevant between man and woman). An expected young age for getting married then means that educational investments are less likely to pay off because returns from joining the labor market are not sufficient. We argue that this effect holds in general, since there are always costs attached to schooling. In large parts of the world, families face even dual costs in the form of direct expenses for sending girls to school, as well as the opportunity cost for not having them as labor support in the household (Glick & Sahn, 2000; King & Hill, 1997). But also families in advanced economic countries with affordable schooling and university tuition will incur substantial costs in the form of associated expenses such as learning materials, commuting and new housing, extracurricular activities etc.

An agent therefore considers sending girls to school as foregone investment from an economic point of view, if expected future returns from their education are not adequately realized through subsequent labor force participation. Assuming (nearly) perfect anticipation through stable intra-generational cultural customs, a societal expectation of young female marriage age will prevent agents from investing in the girls' education already ex ante. This means that the decision on female education is caused by her expected marriage age, i.e. a socio-cultural convention. The mechanism is further reinforced by a negative relationship between female marriage age and the fertility rate.[38] Kalemli-Ozcan, Ryder, and Weil (2000), and Cervellati and Sunde (2007), among others, point to the fact that more educated parents face a higher opportunity cost of child-rearing. Given a low female marriage age environment and resulting little investments in female education, women are more likely to replace child quality with child quantity. This may prolong their expected time period outside the labor market and further reduces incentives to invest in their education; it contrasts to a high female marriage age environment that is associated with fewer children and an expected shorter break from work. In summary, we suggest that socio-cultural customs for marriage age signal the timing for child-rearing, which primarily affects female labor force participation and hence the expected pay-off to female educational investments. Qualitatively, this means the earlier a woman gets married, the lower we expect her education to be, which forms the hypothesis to be tested.[39]

III. 3. Data and Identification Strategy

Data

According to our theoretical considerations we expect that the age at marriage of women affects their level of human capital. For our empirical work, we measure our outcome variable, namely gender-specific education levels, via two measures that are common in the literature (see for example Barro & Lee, 2001, 2013; Castelló-Climent & Hidalgo-Cabrillana, 2012). First, we consider the accumulation of human capital measured via secondary school completion rates (Lorentzen, McMillan, & Wacziarg, 2008; Mankiw et al., 1992). Second, we measure the stock of human capital by the average years of tertiary schooling achieved. Both data stem from the set by

[38] In our panel data set, these two variables correlate significantly (r = -0.62).
[39] We summarize our conceptual framework schematically in figure 3.1 in the appendix.

Barro and Lee (2013). We hypothesize that marriage rather affects later schooling attainment, because this is when marriage decisions mostly interfere. Later we will examine further outcome variable alternatives as robustness checks, and we also compile both absolute female values and gender-relative values where male are divided by female values.

In order to test our hypothesis empirically, we apply a standard cross-country regression framework, as we posit that the economic mechanisms described hold independent of geography. Related studies have examined individual level data (Field & Ambrus, 2008; Maertens, 2013), which naturally have the advantage of capturing single household characteristics and socio-economic factors which aggregate variables can only proxy. However, we deliberately choose a different methodological approach to contribute with a macro perspective to this strand of literature. Moreover, including a larger set of geographies should increase the validity of results, as it allows to draw more general conclusions about the relationship under investigation. Hence, instead of individual information on marriage age we employ the 2012 World Marriage Data by the United Nations (2013), which provide cross-country singulate mean age at marriage (SMAM) data separately for males and females. This allows us to calculate average values for three decadal intervals per country from 1980 to 2010. Using the gender split, we can also create a spousal age gap which captures the female age at marriage relative to the male, so we can analyze differences between women and men. We simply refer to the SMAM for men and women as the *Female Marriage Age*, respectively *Male Marriage Age*, which is formally defined as the average length of single life expressed in years among those who marry before age 50 (United Nations, 2013). The SMAM represents the most common and natural measure for marriage age.[40] In total, we have 86 different countries with available data on gender-specific marriage age and educational outcomes for all three time intervals in our panel.[41]

A number of control variables enter our empirical analysis. These are linked to our theoretical considerations and they are also commonly used in related literature (Carmichael, 2011; Danziger & Neuman, 1999; Díaz-Giménez & Giolito, 2013; Field & Ambrus, 2008; Garenne, 2004). As we propose that marriage serves as an institution for conceiving children, we want to ensure our

[40] The marriage age data are generally not available for every year, hence we average data points per country in ten-year intervals. In cases of quick and dynamics changes, the SMAM variable might be prone to measurement errors (Preston et al., 2001). However, upon data inspection we find no evidence that this could be problematic in our context.
[41] In the appendix, see table 3.19 for a detailed variable description and sources table, and table 3.20 for detailed list of the countries employed for the panel. Including additional controls reduces the sample size due to lower data availability. The cross-section we estimate later in this paper has a larger sample of up to 135 countries.

estimates are not biased by fertility rates in a country. It is plausible that education for females differs across countries not because marriage happens in one country earlier, but because differences in the number of children per woman affects their educational decisions. By controlling for this variable, we eliminate cross-country differences in the average number of children a woman raises, which otherwise may have an unobservable effect on our core relationship between *Female Marriage Age* and female education. We also include the rate of urbanization in a country to capture socio-economic advancement, and the share of Muslim population as control for religious differences that may impact gender roles. Furthermore we employ four population gender ratios that potentially affect a balanced marriage market and could be a reason for age differences.[42] Finally, we consider effects from legal origin differences, from the share of women engaged in the labor market as well as from national income levels.[43]

Table 3.1 provides descriptive statistics for our key variables, broken down by decade. Differences in marriage age are profound: In the latest decade (2000-2010) for example, the *Female Marriage Age* in one country was more than double the age of another country (16 years in Niger versus 33 years in Jamaica). Over the last 30 years, both *Female* and *Male Marriage Age* have globally increased by on average two years. This goes hand in hand with a decline in fertility rates (Foreman-Peck, 2011); women have on average one child less in the 2000s than in the 1980s. The spousal age gap displays a stable pattern of around three and a half years, which is consolidating as reflected in the decreasing standard deviation. In all countries observed, men marry on average later than women. The age gap between husband and wife in the 2000s ranges from an average 1.1 years in Ireland to 8.8 years in Niger. The share of the Muslim population (based on McCleary & Barro, 2006), has remained rather constant and, for the aggregate sample, closely reflects actual shares of the world population. We also observe stable trends for the gender ratios (boys over girls) regarding birth and subsequent mortality rates: More boys than girls are born globally, but they also suffer from a higher mortality rate than girls in childhood and adolescence. Finally, labor participation rate rose in our time window from 40 to 50 percent, and

[42] These are the sex ratio at birth, the death rate of boys with ages 0-4, respectively 0-14, divided by the death rate of girls with ages 0-4, respectively 0-14, and the cumulative percentage of married men at the age of 40 over the cumulative percentage of married women at the age of 40.

[43] Legal origin has been found to affect gender inequality in Dollar, Fisman, and Gatti (2001) and Potrafke and Ursprung (2012), among others. We are aware that labor market characteristics as well as income levels are potentially endogenous. However, we want to ensure that cross-country differences in female labor force participation and per capita incomes do not bias our results. Not including these variables does not materially influence the coefficient of interest.

income (Feenstra, Inklaar, & Timmer, 2015) as well as urbanization levels have grown (UNESCO, 2013), in line with common findings.

Table 3.1: Descriptives

	1980s	1990s	2000s
Female Marriage Age	22.81	24.17	24.76
	(2.98)	(3.48)	(3.60)
Male Marriage Age	26.35	27.63	28.24
	(2.50)	(2.78)	(2.87)
Spousal Age Gap	3.54	3.46	3.48
	(1.60)	(1.51)	(1.38)
Spousal Age Gap Ratio	1.16	1.15	1.15
	(0.09)	(0.09)	(0.08)
Fertility	3.99	3.51	2.97
	(1.84)	(1.77)	(1.57)
Urbanization	50.09	53.41	55.36
	(23.52)	(23.99)	(23.45)
Share of Muslim population	0.19	0.21	0.22
	(0.34)	(0.34)	(0.34)
Sex ratio at birth	1.05	1.05	1.05
	(0.01)	(0.02)	(0.02)
Sex ratio under 5 mortality	1.09	1.08	1.08
	(0.16)	(0.15)	(0.18)
Sex ratio under 15 mortality	1.10	1.09	1.10
	(0.17)	(0.16)	(0.19)
Cum. pop. married at 40	0.98	0.98	0.98
	(0.36)	(0.04)	(0.05)
French Legal Origin	0.40	0.47	0.43
	(0.49	(0.50)	(0.50)
Female Labor Force Participation	40.52	42.28	50.24
	(17.50)	(14.59)	(15.69)
Log GDP per Capita	8.46	8.54	8.77
	(1.17)	(1.28)	(1.28)
Female Secondary Schooling Completion	13.94	17.32	24.19
	(11.36)	(12.25)	(15.21)
Secondary Schooling Completion Ratio	1.57	1.42	1.29
	(0.99)	(0.81)	(1.00)
Average Female Years Tertiary Education	0.17	0.24	0.39
	(0.17)	(0.23)	(0.34)
Average Years Tertiary Education Ratio	1.94	1.88	1.40
	(0.94)	(1.50)	(0.79)
Gender Parity Index in Literacy	0.81	0.83	0.89
	(0.21)	(0.21)	(0.16)
Primary Schooling Completion Ratio	1.33	1.16	1.38
	(0.85)	(0.49)	(2.66)
Number of observations	120	115	135

Notes: This table lists mean and (standard deviation) for the main variables of this paper, where each column shows the ten-year simple average value for the respective decade. The variables are: (i) the *Female Marriage Age* (SMAM); (ii) the *Male Marriage Age* (SMAM); (iii) the Spousal Age Gap, calculated as difference between *Male* and *Female Marriage Age*; (iv) the ratio of *Male* over *Female Marriage Age*; (v) the fertility rate; (vi) the level of urbanization in percent; (vii) the share of muslim population per country; (viii)-(x) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (xi) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (xii) dummy if the country's legal origin is based on French system; (xiii) the percentage of females in the national labor force; (xiv) the log of per capita GDP in PPP terms. (xv) the percentage of females that completed secondary attainment as highest school degree; (xvi) the percentage of males that completed secondary attainment as highest school degree divided by that percentage of females; (xvii) the average years of tertiary schooling for females; (xviii) the average years of tertiary schooling for males divided by that of females; (xix) the Gender Parity Index (females divided by males) for adult literacy; (xx) the percentage of males that completed primary attainment as highest school degree divided by that percentage of females. See the Appendix for more detailed variable definitions and sources.

Our outcome variables follow a common trend: Females have not only significantly improved their levels of education, but also reduced educational inequality relative to males.[44] All three categories that compare education along gender indicate that about 50 percent of the gap between boys and girls could be closed between the 1980s and the 2000s.[45] This general trend towards educational inequality reduction is well in line with literature findings (e.g., Todaro & Smith, 2014). Nevertheless, within that positive trend line large discrepancies remain: whereas in Benin still five times more men than women finish tertiary education and less than half of the women are literate relative to men, the Dominican Republic reports equal gender literacy and a one and a half times female to male ratio for average years of tertiary education.

Identification Strategy

We begin by estimating effects of the female age at marriage on education in country *c* and time period *t* with the following equation, using a regression control approach:

(3.1) $$\text{Educ}_{ct} = \mu + \alpha\,(\text{Female Marriage Age})_{ct} + \beta X_{ct} + \varepsilon_{ct}\,,$$

where Educ is our gender-specific educational outcome variable of interest, and Female Marriage Age denotes the *Female Marriage Age* variable as defined earlier. X is a vector of the control variables introduced before, which we include to mirror our conceptual framework. We then estimate effects in a panel setting with fixed effects over three ten-year timespans from 1980, which represents the earliest sensible data set available, up to 2010. Throughout this chapter, we will make modifications to equation 3.1, which will be presented in detail as they are introduced.

We noted endogeneity concerns when discussing the association of female education levels and female age at marriage in the literature. Three approaches are used in the following to mitigate this issue: (1) We include fixed effects regularly in our regression estimates as mentioned above. (2) We introduce a new instrumental variable that tries to identify the effect of socio-cultural customs, and (3) we apply a quasi difference-in-difference strategy as a complementary examination (starting p. 135).

Regarding point (2), we provide a new instrument related to socio-cultural customs to ensure effects of *Female Marriage Age* on education can be causally interpreted. Specifically, we

[44] Table 3.1 already includes summary statistics for the Gender Parity Index (GPI) in Adult Literacy for comprehensiveness reasons, as we will employ this variable at a later stage.
[45] In that time period, the gap between boys and girls shrank by 51 percent for completed secondary attainment, by 43 percent for average years of tertiary schooling, and by 58 percent for literacy rates.

instrument the domestic *Female Marriage Age* with an average *Female Marriage Age* of neighboring countries. This is the econometric mirroring of our conceptual framework presented earlier, which argues that shared socio-cultural expectations, which are stable within a generation and exogenously given, are closely associated with the actual timing of female marriage. This feature is suitable beyond national borders, since these expectations are much more determined through a common culture sphere than along formal borders. On the other hand, cross-country marriages that could cause endogeneity concerns represent exceptions. Hence, taking the neighboring *Female Marriage Age* from culturally closely associated nations constitutes a meaningful and relevant instrumental approach.

The exclusion restriction requires that the average *Female Marriage Age* from neighboring countries impacts domestic female education levels only through the average domestic *Female Marriage Age*. We find it indeed difficult to imagine how marriage age abroad would directly affect domestic education levels other than through the domestic marriage age. While we cannot directly test this assumption, we believe it is valid, especially since any formal domestic legislation on (female) schooling is confined to the national border, and thus should not impact our instrument. Similarly, foreign education should not have an effect on domestic marriage age, which would theoretically represent a case of reverse causality. For instance, we see little ground to argue that foreign education affects domestic education (and via that channel potentially domestic marriage age), since education levels even of adjacent countries differ on a substantial and persistent level, i.e. we consider cross-border effects as implausible. Education falls under governmental regulation, whereas decisions on marriage timing apart from minimum marriage age does not.[46] Hence, we believe that only the latter is primarily grounded in socio-cultural customs and can have an impact across borders.

We compile our basic instrument by following Correa, Jetter, & Agudelo (2016) in our base specification, i.e. we weigh all values of adjacent countries by the length of shared borders for an average "neighboring value". Adjacent countries with missing values are omitted for the weighted average calculation, which also means that islands are excluded from this sample altogether as they share no direct land border. Domestic *Female Marriage Age* and the *Female Marriage Age*

[46] For this reason we also test empirically if the years of compulsory schooling per country impact our proposed relationship between marriage age and education. While we do not report detailed results, inclusion of a compulsory schooling variable does not change or affect any of our empirical results in section IV.

of neighboring countries are highly correlated (r = 0.8 in our panel), which documents the relevance of our instrument. Related literature also proposes to use neighboring values as instrument to establish exogeneity. Maertens (2013) uses the stated ideal age for getting married of neighboring households within the same subcaste as instrument for an Indian case study. Similar to this micro-setting, here we employ neighboring values on a country level. This yields the following first-stage regression equation:

(3.2) $\text{Female Marriage Age}_{ct} = \gamma + \lambda \text{NEIGHBOR}_{ct} + \theta X_{ct} + \varepsilon_{FemaleMarriageAge_{ct}}$,

where NEIGHBOR_{ct} refers to the average weighted neighbor value of the *Female Marriage Age* in adjacent countries of country *c* in time *t*.

III. 4. EMPIRICAL RESULTS

Baseline Results

Figure 3.2 provides the central motivation for our study and shows the unconditional relationship between our two main outcome variables and *Female Marriage Age* for each time period. All scatterplots in the first three columns display a significant positive relationship between *Female Marriage Age* and female education levels, i.e. the later a woman gets married, the higher her education. As we move from the 1980s to the 2000s, the association becomes stronger for both secondary and tertiary education. The fourth column presents first differences, where for each country the changes in female education from the 1980s to the 2000s are plotted against the changes in *Female Marriage Age*, i.e. we evaluate whether changes in the age of marriage are associated with changes in female education over time, thus holding country characteristics constant. We observe that first differences in marriage age correlate most strongly with changes in advanced education. Changes in female secondary schooling show a weaker but still positive correlation with changes in women's marriage age over time.

Next we run OLS regressions based on equation (3.1) for our panel as summarized in table 3.2. Effects are estimated for secondary and tertiary schooling of women (for now, we focus on absolute female education levels), and we include step-wise additional controls as well as fixed effects. The empirical results support our theoretical predictions. *Female Marriage Age* has a positive and highly significant impact on female education, robust to inclusion of additional

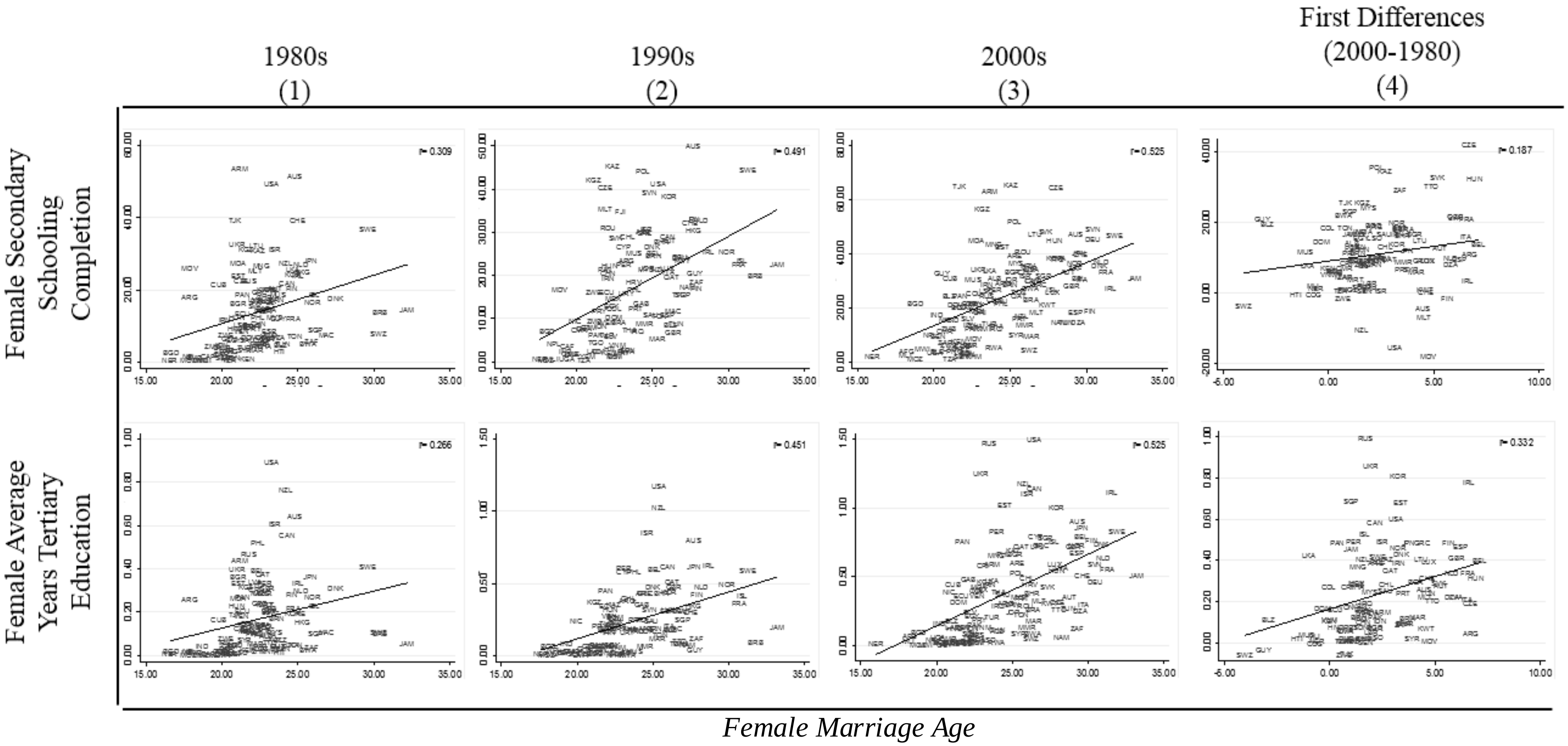


Figure 3.2: Scatter plots of Female Marriage Age against education variables

Table 3.2: Panel for level of female education, ordinary least squares (OLS)

1980-2010 Dependent variable =	Female Secondary Schooling Completion			Average Female Years Tertiary Education		
	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	86	85	81	86	85	81
Observations	258	255	243	258	255	243
Female Marriage Age	2.22	0.82	1.14	0.05	0.02	0.02
	(0.26)***	(0.34)***	(0.50)**	(0.01)***	(0.01)***	(0.01)***
Fertility		-4.03	-3.64		-0.02	0.01
		(1.11)***	(1.84)*		(0.01)**	(0.02)
Urbanization		0.08	-0.01		0.01	0.01
		(0.05)	(0.18)		(0.001)***	(0.01)
Share of Muslim population		3.04	46.41		-0.08	0.11
		(3.46)	(16.76)***		(0.06)	(0.29)
Sex ratio at birth		-3.67	-12.23		-0.89	-2.69
		(31.96)	(44.16)		(0.85)	(1.31)**
Sex ratio under 5 mortality		-38.47	-17.80		-2.75	-2.21
		(36.39)	(49.42)		(0.76)***	(0.81)***
Sex ratio under 15 mortality		35.54	3.18		2.73	2.07
		(38.24)	(49.10)		(0.82)***	(0.88)**
Cum. pop. married at 40		-41.92	-16.52		-1.55	-0.94
		(21.55)**	(26.25)		(0.35)***	(0.42)**
French Legal Origin		-4.69			0.04	
		(1.64)***			(0.04)	
Female Labor Force Participation			-0.08			0.01
			(0.09)			(0.001)**
Log GDP per Capita			6.12			0.17
			(2.40)***			(0.04)***
Continent dummies	no	yes	no	no	yes	no
Country fixed effects	no	no	yes	no	no	yes
R-squared	0.27	0.41	0.47	0.44	0.58	0.67

Notes: The dependent variable in column (1)-(3) is the percentage of the female population with a completed secondary education. Columns (4)-(6) estimate the average years of female tertiary schooling. The regressors are: (i) The female *Marriage Age* (SMAM); (ii) the total fertility rate; (iii) the level of urbanization in percent; (iv) the share of muslim population per country; (v)-(vii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (viii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (ix) dummy whether the country's legal origin is based on French system; (x) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xii) six continent dummies. French Legal Origin omitted in columns (3) and (6) because of inclusion of fixed effects. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

controls and fixed effects.[47] Taking the most stringent specification in columns (3) and (6), each additional year a woman delays marriage means that the share of all females in the country who complete secondary schooling, rises by 1.14 percentage points. Average years of tertiary education

[47] Based on highly significant p-values of a Hausman test for both our outcome variables, which serves to analyze whether the unique errors (ε_i) are correlated with the regressors (see Appendix, table 3.11), we continue with regularly controlling for fixed effects in our model.

are likewise increased by 0.02 years, which equals about one week more time spent at university or roughly a tenth of a semester. Among the additional controls, fertility rates and per capita income levels tend to matter for both outcome variables.[48] Per capita income levels and female labor force participation are likely endogenous in our model, but should still be taken into consideration to ensure that other covariates are not biased from omitted variables.

For this reason, we estimate one model without these two controls, and compare it to the specification with all controls: results are very similar. We also include a control variable for imbalances in the marriage market (*cum. pop. married at 40*, i.e. the cumulative ratio of married males over married females at age 40). It might be that the bridal marriage age is determined not so much based on socio-cultural expectations, but is more directly related to the widely debated issue of "missing women" (Anderson & Ray, 2010; Sen, 1990), which affects marriage market characteristics.

If there are women missing on the marriage market due to unbalanced overall population ratios, men would likely have to choose and marry young females as long as they are still available. This would bias the female marriage age downwards. Such a missing women effect could also result in a "marriage squeeze" (Edlund, 2002; Grossbard-Shechtman, 1993), as relatively more men do not marry because the unbalanced marriage market leaves them without a matching partner. Against this backdrop, we proxy for a distorted marriage market with this ratio as we would expect a balanced ratio for a population with normal distribution of sexes.[49] Other covariates are either non-significant, or they have no consistent effect on the two outcome variables under investigation, i.e. they are rather erratic.

Lastly, we test specifications where we additionally control for the average age of a woman when giving birth to her first child (see table 3.14 in the appendix). This variable has no large effects and *Female Marriage Age* remains robust, thus the overall outcome remains unchanged. It seems that indeed our key variable of interest, the *Female Marriage Age,* exerts a crucial influence on education, while the timing of first birth has an evidently more subordinate impact. This

[48] Given the high correlation between fertility and marriage age, we may also interpret the sex ratio at birth as de-facto instrument for fertility, following Becker, Cinnirella & Woessmann (2010). Based on a reviewer's suggestion we also exclude fertility in an alternative specification (not reported here). This leads to even more significant results for our core variable of interest.

[49] We also tested an alternative specification with the ratio of cumulative percentage of married men over that of married women at the age of 30 as control variable, and we omitted this variable as well in another specification, based on a reviewer's suggestion. Results turn out to be very similar.

confirms our conceptual framework in which we derived that marriage age constitutes the more fundamental explanatory variable, i.e. age at first birth is an outcome of this regressor. Related studies tend to emphasize the role of maternal age at first birth for female education, but according to our theoretical logic and empirical findings, incorrectly so since both variables are ultimately impacted by female marriage age. If the timing of marriage shifts, so does the maternal age at first birth as a consequence (not vice versa). Hence, we believe our considerations reflect better the causal sequence that leads to educational gender inequality.

Instrumental Variables and Spousal Age Gaps

Although the initial association between female age at marriage and female education levels appears statistically and economically relevant, we try to ensure that the relationship between the two variables can be causally interpreted. Hence, in table 3.3 (Panel A) we proceed to instrumental variables estimates, where the average of the *Female Marriage Age* of adjacent countries, weighted by shared land border, serves as our instrument.[50] The domestic *Female Marriage Age* is strongly affected by established societal expectations, which a country shares with its neighbors through common cultural heritage. Thereby, we account for concerns of causality and endogeneity between marriage age and education.

We then take an additional step to tackle omitted variable bias. So far we have analyzed our key variable of interest, the marriage age for women, as absolute values through the *Female Marriage Age*. We accounted for national differences by the inclusion of country fixed effects as well as cultural variables such as Muslim population shares. However, our absolute perspective might still overlook particular social customs regarding marriage age that would bias results. Also, *Male Marriage Age* correlates highly with *Female Marriage Age* so that results so far could capture effects simply from marrying young per se, rather than from *Female Marriage Age*. Hence, we take the difference between the *Marriage Age* of husband and wife to obtain a spousal age gap (SAG) per country. Introducing this new variable is by and large due to methodological considerations, as our overall focus in this research remains with the age at marriage. Yet,

---

[50] In a comment on the related working paper version of this chapter, concerns were raised regarding the limited number of countries in the following regressions and resulting potential selection bias. However, we do not believe this to be an issue. The sample still features countries from all six continents and represents simultaneously the large majority of the world population since basically all highly populous nations are included. Furthermore we employ continent dummies to address bias of potentially over-represented regions, and we note that related literature operates on a similar empirical data basis.

examining SAG effects should even be better suited to reduce biases in case women in one part of the world get married earlier than somewhere else, irrespective of education levels. It likewise considers how men behave in terms of timing for marriage. As all of our observations display a higher average age for men than for women at marriage, spousal age gaps are consistently positive values. For our regression analyses in table 3.3, we examine spousal age gap effects in a simple OLS regression (Panel B), and then again instrument the domestic SAG with neighboring values, where for consistency reasons we now use the SAG of adjoining countries weighted by shared border length (Panel C).

Finally, we seek further evidence by employing Generalized Method of Moments (GMM) techniques. These allow for a more flexible estimation than least squares methods and further address potential endogeneity issues. The specific method chosen is the one-step Arellano-Bond dynamic panel estimator, which is based on the idea that our instrumental variables approach so far does not exploit all of the information available in the sample, and which also recognizes fixed effects (Arellano & Bond, 1991). In this GMM context, we may construct more efficient estimates of the dynamic panel data model, especially for our panel which is characterized by few time periods with many individual cases (countries). However, as the model needs to employ all available lags of the specified variables in levels dated t-1 or earlier, our panel is reduced from three to two independent observation points per country.

All results confirm a strong and significant relationship according to our theoretical considerations: In panel A, the IV procedure increases the coefficients increase in comparison to before, i.e. *Female Marriage Age* explained by societal conventions displays a greater impact on female education. Estimated effects more than double in size in comparison to the OLS results. Hence, estimates that do not explicitly address causality (table 3.2) tend to bias down marriage age effects. First stage F-test statistics in the linear model indicate very good identification of our instrument, which is always a highly significant predictor for spousal age gaps.[51] The IV results also remain robust to the inclusion of a set of control variables: in our most comprehensive GMM specification, each year a woman postpones marriage translates into a 3.2 percentage points higher female completion rate in secondary schooling. For tertiary education, one year of marriage delay equals nearly three weeks longer tertiary schooling for women. Note that the sample average in the 2000s amounts to only one semester of total university attendance for women; reducing those

[51] See table 3.12 in the Appendix for detailed first stage results.

twenty weeks by three because of one year earlier marriage means corresponds to a cut by 13 percent.[52]

Moving on to panel B, the relative M*arriage Age* effects (spousal age gap) appear to have a slightly more robust impact on tertiary than on secondary education. The latter is no longer significant in either linear or GMM specification if all controls and fixed effects are added to the model. In any case, coefficients are now consistently negative which is in accordance with our theoretical considerations: The larger the spousal age gap, i.e. the earlier a woman gets married relative to the husband, the lower her education. In absolute size the effects are smaller than before. This likely reflects the fact that effects from simply "marrying young" are now removed, since we only analyze the gender-specific impact from women marrying relatively younger than the husband. Also, the lack of instrumentation in panel B is likely to again bias coefficients downwards due to reverse causality.

We therefore employ our described instrument (in this case the average spousal age gap of neighboring countries weighted by length of border shared) in panel C. Revisiting the strength of our instrument, the first stage F-test values display mostly robustness with regards to common threshold levels (Staiger & Stock, 1997). However, we note that once all controls and fixed effects are added, the coefficients of spousal age gap in the second stage can no longer be estimated precisely enough to maintain significance levels. This suggests that fixed effects capture a lot of the rather invariant spousal age gap regressor, which does not fluctuate much over time. Still the direction remains unambiguous and coefficients are larger than in panel B. We attribute this again to better identification and resolved reverse causality issues, consistent with the observations we made when contrasting OLS and IV estimates for absolute *Female Marriage Age*.

In our preferred GMM specification with only exogenous controls and fixed effects (columns 10 and 14), one additional standard deviation of the spousal age gap variable (1.4 years) leads to a 14 percentage point lower secondary schooling completion rate for girls. It analogously translates into about six weeks less of female tertiary education. In the 2000s, Germany and Albania displayed roughly such a gap in female educational levels, and correspondingly a spousal age gaps of approximately 1.4 years (4.0 years in Albania versus 2.7 years in Germany).

[52] Our estimates tend to be smaller compared to findings for Bangladesh and India: Field and Ambrus (2008) estimate that one additional year of delayed marriage results in 2.6 more months of total education in Bangladesh, while Maertens (2013) estimates up to 8.5 months for rural India. We would reconcile the numbers such around three weeks of that total additional time can be attributed to tertiary education.

Table 3.3: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age

	Linear Regression								*Generalized Method of Moments (GMM Arellano-Bond)*					
1980-2010 Dependent variable =	Female Secondary Schooling Completion				Average Female Years Tertiary Education				Female Secondary Schooling Completion			Average Female Years Tertiary Education		
Panel A: Female Marriage Age (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Number of countries	71	70	66	66	71	70	66	66	71	70	66	71	70	66
Observations	213	210	198	198	213	210	198	198	142	140	132	142	140	132
Female Marriage Age	3.21	2.09	2.12	3.88	0.07	0.04	0.04	0.04	3.61	3.62	3.23	0.07	0.06	0.05
	(0.05)***	(0.67)***	(0.78)***	(1.17)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.58)***	(1.25)***	(1.23)***	(0.01)***	(0.02)***	(0.02)**
Fertility		-3.98	-4.04	-4.01		-0.01	-0.01	-0.01		-3.29	-3.59		-0.01	-0.01
		(0.88)***	(0.96)***	(1.54)***		(0.01)	(0.02)	(0.02)		(1.42)**	(1.39)***		(0.03)	(0.02)
Urbanization		0.03	-0.04	-0.22		0.01	0.01	-0.01		-0.12	-0.15		0.01	-0.01
		(0.07)	(0.09)	(0.18)		(0.001)***	(0.01)	(0.01)		(0.22)	(0.22)		(0.01)	(0.01)
Share of Muslim population		4.04	4.44	63.01		-0.05	0.04	0.09		46.09	37.21		0.17	0.01
		(3.84)	(4.52)	(22.58)***		(0.07)	(0.08)	(0.30)		(17.47)	(17.18)**		(0.30)	(0.29)
Sex ratio at birth		-3.67	-27.29	1.70		-0.12	-0.11	-1.09		14.64	-29.19		-0.39	-1.01
		(58.43)	(62.03)	(100.74)		(1.00)	(1.01)	(1.34)		(75.29)	(78.79)		(0.68)	(0.69)
Sex ratio under 5 mortality		28.68	48.73	112.85		-1.56	-1.20	-1.38		52.42	54.82		-1.19	-0.92
		(42.45)	(43.76)	(60.77)*		(0.72)**	(0.69)*	(0.81)*		(51.32)	(48.52)		(0.98)	(0.87)
Sex ratio under 15 mortality		-32.01	-51.86	-140.22		1.52	1.14	1.27		-69.57	-68.22		0.99	0.83
		(45.98)	(47.34)	(70.74)**		(0.78)**	(0.76)	(0.94)		(59.38)	(55.63)		(1.19)	(1.06)
Cum. pop. married at 40		-48.13	-42.67	11.83		-1.64	-1.59	-1.19		8.21	9.81		-0.33	-0.42
		(25.46)*	(28.48)	(42.52)		(0.43)***	(0.46)***	(0.57)**		(46.61)	(47.50)		(0.82)	(0.75)
French Legal Origin		-6.23	-5.98			-0.01	0.02							
		(2.27)***	(2.51)**			(0.04)	(0.04)							
Female Labor Force Participation			0.05	-0.10			0.01	-0.01			-0.08			0.01
			(0.06)	(0.10)			(0.001)***	(0.001)**			(0.12)			(0.01)
Log GDP per Capita			1.42	3.06			0.05	0.11			4.28			0.11
			(1.92)	(3.05)			(0.03)*	(0.04)***			(2.17)**			(0.03)***
Continent dummies	no	yes	yes	no	no	yes	yes	no	no	no	no	no	no	no
Country fixed effects	no	no	no	yes	no	no	no	yes	yes	yes	yes	yes	yes	yes
First stage F-test statistics	274.96	52.93	41.05	19.07	274.96	52.93	41.05	19.07						
R-squared	0.33	0.52	0.50	0.40	0.34	0.59	0.61	0.69						

Table 3.3 continued: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age

	Linear Regression								*Generalized Method of Moments (GMM Arellano-Bond)*					
1980-2010 Dependent variable =	Female Secondary Schooling Completion				Average Female Years Tertiary Education				Female Secondary Schooling Completion			Average Female Years Tertiary Education		
Panel B: Spousal Age Gap (OLS)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Number of countries	86	85	81	81	86	85	81	81	86	85	81	86	85	81
Observations	258	255	243	243	258	255	243	243	172	170	162	172	170	162
Spousal Age Gap	-3.51	-0.46	-0.60	-0.66	-0.09	-0.02	-0.01	-0.04	-0.76	-0.13	-0.31	-0.04	-0.03	-0.04
	(0.65)***	(0.23)**	(0.79)	(1.11)	(0.01)***	(0.01)**	(0.01)	(0.02)**	(0.39)**	(0.08)*	(1.05)	(0.01)***	(0.01)***	(0.01)***
Fertility		-4.70	-4.14	-4.29		-0.04	-0.01	-0.01		-4.59	-4.06		-0.05	0.02
		(0.66)***	(0.76)***	(1.05)***		(0.01)***	(0.01)	(0.02)		(1.54)***	(1.55)***		(0.01)***	(0.01)
Urbanization		0.12	0.03	-0.01		0.01	0.01	0.01		0.17	0.14		0.01	0.01
		(0.05)**	(0.07)	(0.16)		(0.001)	(0.001)**	(0.01)		(0.17)	(0.16)		(0.001)**	(0.01)
Share of Muslim population		2.78	2.77	44.21		-0.06	0.02	0.25		15.74	11.96		-0.16	-0.16
		(3.40)	(3.84)	(20.65)**		(0.07)	80.07)	(0.32)		(12.14)	(14.02)		(0.22)	(0.25)
Sex ratio at birth		-18.51	-8.22	-28.22		-1.03	-0.42	-3.09		35.67	-1.87		-2.32	-2.91
		(53.79)	(57.34)	(86.71)		(1.00)	(1.02)	(1.35)**		(42.59)	(50.66)		(1.70)	(1.68)*
Sex ratio under 5 mortality		-68.89	-42.53	-55.03		-3.43	-2.58	-2.87		-93.37	-64.88		-3.46	-2.43
		(30.96)**	(33.43)	(40.72)		(0.56)***	(0.58)***	(0.64)***		(33.67)**	(36.77)*		(0.88)***	(0.77)***
Sex ratio under 15 mortality		69.76	42.95	48.79		3.52	2.62	2.84		93.98	64.28		3.57	2.51
		(32.03)**	(34.51)	(43.28)		(0.58)***	(0.60)***	(0.68)***		(35.56)**	(38.72)*		(0.93)***	(0.83)***
Cum. pop. married at 40		-54.55	-50.25	-33.61		-1.94	-1.76	-1.22		-61.23	-44.79		-1.66	-1.14
		(19.77)***	(20.37)***	(25.19)		(0.36)***	(0.35)***	(0.39)***		(24.25)**	(22.18)**		(0.37)	(0.37)***
French Legal Origin		-5.18	-4.66			0.04	0.08							
		(2.11)***	(2.26)**			(0.04)	(0.04)*							
Female Labor Force Participation			0.04	-0.06			0.01	0.01			-0.06			0.01
			(0.06)	(0.09)			(0.001)***	(0.001)**			(0.08)			(0.01)
Log GDP per Capita			3.49	7.60			0.10	0.21			5.38			0.16
			(1.47)**	(2.26)***			(0.03)***	(0.04)***			(2.21)**			(0.04)***
Continent dummies	no	yes	yes	no	no	yes	yes	no	no	no	no	no	no	no
Country fixed effects	no	no	no	yes	no	no	no	yes	yes	yes	yes	yes	yes	yes
R-squared	0.20	0.52	0.50	0.11	0.26	0.59	0.59	0.44						

Table 3.3 continued: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age

	Linear Regression								*Generalized Method of Moments (GMM Arellano-Bond)*					
1980-2010 Dependent variable =	Female Secondary Schooling Completion				Average Female Years Tertiary Education				Female Secondary Schooling Completion			Average Female Years Tertiary Education		
Panel C: Spousal Age Gap (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Number of countries	71	70	66	66	71	70	66	66	71	70	66	71	70	66
Observations	213	210	198	198	213	210	198	198	142	140	132	142	140	132
Spousal Age Gap	-9.15	-0.83	-0.54	112.30	-0.19	-0.09	-0.09	1.14	-6.59	-9.77	-22.51	-0.14	-0.07	-0.25
	(1.25)***	(0.40)**	(2.95)	(255.31)	(0.02)***	(0.05)**	(0.06)*	(2.66)	(1.07)***	(4.64)**	(15.69)	(0.03)***	(0.04)**	(0.19)
Fertility		-3.76	-3.56	-26.96		-0.01	0.01	-0.24		-4.27	-4.08		-0.05	-0.08
		(0.90)***	(0.91)***	(46.59)		(0.02)	(0.02)	(0.49)		(2.52)*	(6.00)		(0.03)	(0.06)
Urbanization		0.15	0.11	1.45		0.01	0.01	0.02		-0.12	-0.88		0.01	-0.01
		(0.07)**	(0.09)	(3.89)		(0.01)***	(0.01)	(0.04)		(1.35)	(0.45)		(0.01)	(0.01)
Share of Muslim population		2.50	4.07	-704.20		0.03	0.09	-7.67		78.79	169.92		-0.02	1.52
		(3.78)	(3.91)	(1687.95)		(0.07)	(0.07)	(17.62)		(307.24)	(114.21)		(1.46)	(1.37)
Sex ratio at birth		-25.76	-36.83	532.05		0.06	0.17	4.27		63.76	42.77		-0.40	-2.40
		(52.16)	(56.34)	(1481.59)		(0.99)	(1.06)	(15.47)		(145.18)	(134.47)		(1.14)	(2.13)
Sex ratio under 5 mortality		-46.08	-17.95	89.21		-1.87	-1.31	-1.61		-80.65	-1.44		-3.48	-1.86
		(34.88)	(39.28)	(425.02)		(0.66)***	(0.74)*	(4.43)		(76.01)	(98.41)		(0.98)***	(1.23)
Sex ratio under 15 mortality		50.77	22.21	93.39		2.12	1.57	3.63		63.63	-34.15		3.70	1.69
		(35.23)	(39.18)	(408.90)		(0.67)***	(0.74)	(4.27)		(156.10)	(123.42)		(1.04)***	(1.54)
Cum. pop. married at 40		-71.89	-67.66	-471.59		-2.86	-2.78	-6.07		-79.00	-53.58		-1.97	-1.43
		(27.92)***	(28.91)**	(910.76)		(0.53)***	(0.55)***	(9.51)		(77.95)	(54.81)		(0.46)***	(0.64)**
French Legal Origin		-6.40	-6.44			0.02	0.03							
		(2.03)***	(2.06)***			(0.04)	(0.04)							
Female Labor Force Participation			0.09	-0.17			0.01	0.01			-0.06			0.01
			(0.06)	(0.75)			(0.001)**	(0.01)			(0.17)			(0.01)
Log GDP per Capita			1.52	-6.55			0.04	0.02			12.71			0.23
			(1.56)	(38.88)			(0.03)	(0.41)			(7.58)*			(0.10)**
Continent dummies	no	yes	yes	no	no	yes	yes	no	no	no	no	no	no	no
Country fixed effects	no	no	no	yes	no	no	no	yes	yes	yes	yes	yes	yes	yes
First stage F-test statistics	76.06	19.68	15.54	8.91	76.06	19.68	15.54	8.91						
R-squared	0.24	0.55	0.54	0.03	0.25	0.57	0.57	0.03						

Notes: The dependent variable in columns (1)-(4) and (9)-(11) is the percentage of the female population with a completed secondary education. Columns (5)-(8) and (12)-(14) estimate the average years of female tertiary schooling. The regressors are: (i) The *Female Marriage Age* (SMAM) in Panel A, and the Spousal Age Gap (*Male* minus *Female Marriage Age*) in Panels B and C; (ii) the total fertility rate; (iii) the level of urbanization in percent; (iv) the share of muslim population per country; (v)-(vii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (viii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (ix) dummy whether the country's legal origin is based on French system; (x) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xii) six continent dummies. French Legal Origin omitted in columns (4), (8), and (9)-(14) because of inclusion of fixed effects, respectively first differences. The GMM estimator in Panel B uses as instrument the regressor itself to mirror the linear OLS scenario. Panels A and C use as instrument the weighted average of the absolute, respectively relative Female Marriage Age of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Effects on Female Relative to Male Education

Our research so far has identified negative effects of young female marriage age as well as of large spousal age gaps on female education levels. Our theoretical considerations attribute a negative impact on females only, since men's career prospects are generally not constrained by children and founding a family. The following analysis tests more explicitly if indeed male and female education is affected differently.

For this purpose, we go back to equation (3.1), but employ a re-coded variable as regressor, namely the ratio of *Male* over *Female Marriage Age* (spousal age gap ratio). Analogously, our outcome variable for secondary and tertiary education is now coded as male over female values per country. We hence have a quasi diff-in-diff specification, employing relative levels on both sides of the equation, as gender ratios in education levels are explained with gender ratios in marriage age. Estimating gender-relative education serves to identify the real educational gender gap independent of the many other characteristics that potentially influence a country's level of education. Nonetheless, we also include a set of control variables to strengthen our findings, and we address endogeneity in the usual way by taking neighboring countries' values as instruments.[53]

Table 3.4 reports econometric results of how gender differences in marriage age translate into gender differences in education. There is strong evidence for a gender-discriminating effect of marriage age gaps, since larger age gaps also increase the educational gap between men and women. This lends strong support to our hypothesis that women's education is relatively more influenced by marriage timing than men's. Consistent with previous results, estimated effects using instruments are again considerably larger when we compare all significant spousal age gap ratio regressors of the OLS with the equivalent IV specifications. Our preferred and highly significant IV specification with key controls (columns (3) and (7)) suggests that switching from the same age of both husband and wife to a scenario where the husband is twice as old (a switch in the ratio from one to two) leads to a more than three times higher completion rate for boys than girls in secondary schooling, and a nearly eleven times longer time spent at university for male than for female adolescents. Results indicate that robustness cannot be held at conventional significance levels when including fixed effects on top of other controls. This is likely again due

[53] See table 3.13 in the Appendix for detailed first stage results.

Table 3.4: Panel for relative male-to-female education levels (ratio)

1980-2010								
Dependent variable =	Secondary Schooling Completion Ratio				Average Years Tertiary Education Ratio			
Panel A: Spousal Age Gap Ratio (OLS)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	86	85	81	81	86	85	81	81
Observations	258	255	243	243	258	255	243	243
Spousal Age Gap Ratio	5.32	4.51	1.64	0.85	7.52	4.47	5.38	4.98
	(1.32)***	(2.74)*	(0.88)*	(1.75)	(1.97***	(2.26)**	(2.67)**	(3.43)
Fertility		0.16	0.18	0.25		0.25	0.21	0.23
		(0.07)**	(0.05)***	(0.07)***		(0.07)***	(0.08)***	(0.11)**
Urbanization		-0.01	0.01	-0.01		-0.01	0.01	-0.01
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.02)
Share of Muslim population		-0.22	0.15	-1.66		-0.64	-0.85	-2.67
		(0.27)	(0.15)	(1.20)		(0.53)	(0.57)	(1.61)*
Sex ratio at birth		2.23	3.34	-4.39		-0.39	0.13	-6.21
		(2.00)	(1.29)***	(5.64)		(3.74)	(3.45)	(3.43)*
Sex ratio under 5 mortality		-0.85	-0.14	0.64		2.52	0.33	0.88
		(1.68)	(1.28)	(1.75)		(1.89)	(1.99)	(2.63)
Sex ratio under 15 mortality		1.10	0.60	-0.49		-2.71	-0.52	-1.55
		(1.67)	(1.28)	(1.84)		(1.91)	(1.98)	(2.72)
Cum. pop. married at 40		-0.45	-0.68	0.08		0.45	-0.16	1.68
		(1.32)	(0.69)	(0.95)		(1.57)	(1.61)	(1.75)
French Legal Origin		0.23	0.05			0.13	0.07	
		(0.18)	(0.10)			(0.21)	(0.23)	
Female Labor Force Participation			0.01	-0.01			-0.01	-0.02
			(0.01)	(0.01)*			(0.01)*	(0.01)
Log GDP per Capita			-0.10	0.10			-0.11	-0.03
			(0.06)	(0.11)			(0.13)	(0.22)
Continent dummies	no	yes	yes	no	no	yes	yes	no
Country fixed effects	no	no	no	yes	no	no	no	yes
R-squared	0.20	0.30	0.54	0.28	0.24	0.33	0.38	0.23
Panel B: Spousal Age Gap Ratio (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	71	70	66	66	71	70	66	66
Observations	213	210	198	198	213	210	198	198
Spousal Age Gap Ratio	7.34	5.24	3.62	-54.17	8.47	9.29	10.71	-36.53
	(1.29)***	(2.84)*	(1.49)***	(118.00)	(0.90)***	(2.75)***	(2.64)***	(91.95)
Fertility		0.27	0.20	1.21		0.14	0.12	0.91
		(0.09)***	(0.04)***	(1.96)		(0.07)**	(0.08)	(1.39)
Urbanization		-0.01	0.01	-0.03		0.01	0.01	-0.03
		(0.01)	(0.01)	(0.08)		(0.01)	(0.01)	(0.08)
Share of Muslim population		-0.22	0.18	15.69		-0.59	-0.66	11.19
		(0.35)	(0.18)	(37.20)		(0.33)*	(0.32)**	(29.21)
Sex ratio at birth		4.87	4.45	-20.09		-2.07	1.87	-18.91
		(5.19)	(2.58)*	(38.84)		(4.38)	(4.46)	(27.37)
Sex ratio under 5 mortality		-3.82	-1.96	3.83		2.43	0.07	4.11
		(3.51)	(1.81)	(12.92)		(2.84)	(3.11)	(9.22)
Sex ratio under 15 mortality		4.75	2.62	-10.30		-2.35	-0.01	-9.53
		(3.56)	(1.83)	(23.76)		(2.89)	(3.14)	(17.21)
Cum. pop. married at 40		1.35	0.01	16.40		1.05	0.30	14.34
		(2.48)	(1.19)	(35.60)		(1.92)	(2.02)	(26.45)
French Legal Origin		0.27	0.04			-0.17	-0.17	
		(0.18)	(0.09)			0.16	(0.15)	
Female Labor Force Participation			0.01	0.01			-0.01	0.01
			(0.01)*	(0.03)			(0.01)	(0.03)
Log GDP per Capita			-0.14	0.37			-0.02	-0.02
			(0.07)	(0.82)			(0.13)	(0.49)
Continent dummies	no	yes	yes	no	no	yes	yes	no
Country fixed effects	no	no	no	yes	no	no	no	yes
First stage F-test statistics	216.92	54.75	48.60	5.18	216.92	54.75	48.60	5.18
R-squared	0.21	0.34	0.58	0.01	0.30	0.40	0.44	0.01

Notes: The dependent variable in column (1)-(4) is the ratio of the male over female population share with a completed secondary education. Columns (5)-(8) estimate the ratio of the male over female average years of tertiary schooling. The regressors are: (i) The Spousal Age Gap Ratio (*Male* over *Female Marriage Age*); (ii) the total fertility rate; (iii) the level of urbanization in percent; (iv) the share of muslim population per country; (v)-(vii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (viii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (ix) dummy whether the country's legal origin is based on French system; (x) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xii) six continent dummies. French Legal Origin omitted in columns (4) and (8) because of inclusion of fixed effects. Panel B uses as instrument the weighted average of the Spousal Age Gap Ratio of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

to the rather invariant spousal age gap ratio which is consequently absorbed into fixed effects so that its stand-alone explanatory power decreases.[54]

Altogether, the estimates again support the conclusion that educational gender gaps are exacerbated by marriage age. Spousal age gaps do not only have a negative effect on female education, but this effect is also significantly larger than what we observe for male education. Women's education suffers not only in absolute levels, but also relatively more than men's when an age gap between husband and wife opens up.

III. 5. Robustness Tests and Refinements

Alternative Instruments

In this section we run a number of additional specifications to further check the robustness of our results. We begin by employing two alternative instruments to confirm that the conclusions drawn so far do not depend on a particular choice of instrumental variable.

On the one hand, we use as instrument the average female marriage age of five neighboring countries. If a country has less than five neighboring countries with a shared land border, we include countries in the same (cultural) region to always obtain a balanced group of five. While this is a mathematically less deterministic approach than our base instrument, it has the advantage of including more countries (especially islands), and of shaping the exposure to neighboring values towards countries that have similar cultural customs.

In a second check we proceed to another alternative instrument to corroborate our findings through an entirely different channel. The origin of different gender roles and gender treatment as a result of agricultural practices and physiological differences between men and women has recently received renewed interest. Following Boserup (1970), Alesina, Giuliano, and Nunn (2013) present findings that link the suitability of a location for cultivating crops that benefitted from plow adoption to unequal gender norms today. As these "plow-positive" crops required more body strength and force, men assumed the natural role of performing such work. Women, in return, adopted primarily roles within the domestic household. The authors argue that these labor division

[54] We also tested the impact of three dummies for different female marriage age groups (18-23, 23-28, and 28-32 years); see table 3.15 in the appendix. The two younger age groups tend to have a marginal negative impact, which seems plausible as it reflects that the education of very young women is most adversely affected by entering marriage. Nevertheless, the spousal age gap variable keeps robustness in terms of coefficient size and significance, hence spousal age gaps maintain their significant impact irrespective of the absolute female age at marriage.

practices transformed into norms about the natural role of each gender.[55] These persist as cultural beliefs even in today's time, i.e. after the economy has advanced from a traditional agricultural structure. Assuming that such views on gender roles are also mirrored in today's marriage age patterns, plow-positive and plow-negative environments represent relevant instruments for our analysis.

Traditionally, education and schooling were a privilege reserved to very few in society, who were not representative of the agricultural workforce. It is implausible to think that girls and boys, respectively their parents, would have decided on marriage age back then as a consequence of schooling, simply because there was virtually no schooling. Few cases of mandatory schooling were limited to basic primary education, which would not interfere with adolescence and related marriage decision. Widespread education only appeared in Europe and North America when industrialization began – hence precisely once the economy moved away from its traditional agricultural character (Barnard, 1969; Cordasco, 1976; Maynes, 1985; Mulhern, 1959; Too, 2001). In other, less developed parts of the world, general schooling has only emerged in the second half of the twentieth century (Lockheed & Verspoor, 1992). In summary, education appeared only relatively recently in time, and hence could not have possibly influenced traditionally anchored gender roles and the associated female marriage age. In contrast, the cultural heritage of agricultural economic reasons on gender roles is likely important in the decision-making of when a woman should get married. Thus, we believe that plow-positive and plow-negative environments could serve as valid instruments. Note that we cannot run this model with fixed effects since our invariant instrument variables would be omitted from the analysis. We therefore estimate a random effects model, but include continent dummies as additional control to proxy geographical fixed effects.

Table 3.5 summarizes the empirical results for spousal age gap effects on female education, when the former is instrumented with our two alternative variables presented. Coefficients closely correspond to the preferred instrument as used in tables 3.3 and 3.4, but the estimates are comparatively less precise. Given the known strength of our other control variables for impacting female education, full robustness is difficult to achieve when employing this relatively weaker instrument and including all covariates (see table 3.16 in the appendix for first stage results).

[55] The line or argument is supported by Gimbutas (2007) who finds that prior to the invention of the plow, societies tended to be matriarchal and more equal. Hodder (2005) argues similarly.

Overall, the additional IV results using our alternative instruments are nonetheless economically meaningful. In particular, the sign of coefficients is as expected in 14 out of 16 cases (negative for spousal age gaps, and positive for spousal age gap ratio), suggesting that spousal age gaps causally and negatively affect female education.

Table 3.5: Panel for level of female education using alternative instruments

Instrument employed	*Average of 5 neighboring countries*				*Plow-environment*			
Panel A: Spousal Age Gap (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1980-2010 Dependent variable =	Female Secondary Schooling Completion		Average Female Years Tertiary Education		Female Secondary Schooling Completion		Average Female Years Tertiary Education	
Number of countries	86	81	86	81	86	81	86	81
Observations	258	243	258	243	258	243	258	243
Spousal Age Gap	-6.79	7.77	-0.15	-0.10	-7.76	-0.19	-0.17	-0.12
	(1.03)***	(11.37)	(0.02)***	(0.16)	(1.77)***	(0.29)	(0.04)***	(0.05)**
Additional Controls	no	yes	no	yes	no	yes	no	yes
Country fixed effects	no	yes	no	yes	no	no	no	no
First stage F-test statistics	161.20	2.02	161.20	2.02	25.43	5.98	25.43	5.98
R-squared	0.20	0.32	0.26	0.33	0.20	0.49	0.26	0.51
Panel B: Spousal Age Gap Ratio (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1980-2010 Dependent variable =	Secondary Schooling Completion Ratio		Average Years Tertiary Education Ratio		Secondary Schooling Completion Ratio		Average Years Tertiary Education Ratio	
Number of countries	86	81	86	81	86	81	86	81
Observations	258	243	258	243	258	243	258	243
Spousal Age Gap Ratio	11.85	20.64	20.73	200.37	5.66	2.47	6.61	-1.49
	(3.64)***	(36.03)	(5.81)***	(1558.84)	(1.90)***	(2.83)	(2.13)***	(7.51)
Additional Controls	no	yes	no	yes	no	yes	no	yes
Country fixed effects	no	yes	no	yes	no	no	no	no
First stage F-test statistics	17.03	0.44	17.03	0.44	29.02	5.77	29.02	5.77
R-squared	0.21	0.01	0.25	0.01	0.20	0.54	0.24	0.30

Notes: The dependent variable in panel A, column (1)-(2) and (5)-(6) is the percentage of the female population with a completed secondary education. Columns (3)-(4) and (7)-(8) in panel A estimate the average years of female tertiary schooling. The dependent variable in Panel B column (1)-(2) and (5)-(6) is the ratio of the male over female population share with a completed secondary education. Columns (3)-(4) and (7)-(8) in panel B estimate the ratio of the male over female average years of tertiary schooling. The Spousal Age Gap is the *Male* minus *Female Marriage Age* and the Spousal Age Gap Ratio the *Male* over *Female Marriage Age*. Additional Controls are i) the total fertility rate; (ii) the level of urbanization in percent; (iii) the share of muslim population per country; (iv)-(vi) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (vii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (viii) dummy whether the country's legal origin is based on French system; (ix) the percentage of females in the national labor force; (x) the log of per capita GDP in PPP terms; and (xi) six continent dummies for columns (6) and (8). Columns (1)-(4) use as instrument the average of the spousal age gap, respectively the spousal age gap ratio, of five neighboring countries; columns (5)-(8) use as instrument plow-positive and plow-negative environment (Alesina et al., 2013). See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Alternative measures for educational gender inequality

We introduce additional outcome variables designed to reaffirm that we correctly measure educational gender inequality. We first adopt the gender parity in literacy (GPI) index by the UNESCO (2013). Literacy rates also serves as category of human capital beyond formal educational attainment measures. The latter are criticized for ignoring human capital accumulated outside schooling (Barro & Lee, 2013). As this variable is not available for our panel data, we move to a cross-section analysis, taking the latest time interval averaging the years 2000-2010. We also look at primary schooling completion rates of boys relative to girls. So far we analyzed secondary and tertiary schooling, since this is when we would expect marriage decisions to interfere most. Put differently, we hypothesize boys and girls to plan primary schooling independent of anticipated marriage age so that its explanatory power should be low in this case.

Table 3.6 lists the results for the GPI, respectively primary schooling outcome variable, where OLS, IV, and the quasi diff-in-diff method are each summarized in panels A through C. For the GPI we adopt the UNESCO coding, which calculates this variable as female over male literacy rates and is inverse to the ratios of primary schooling in columns (4) to (6) (and likewise inverse to the secondary and tertiary schooling ratios we analyzed before). Therefore, we would expect a negative sign for the GPI, as a relatively younger women (i.e. a larger spousal age gap variable) should reduce literacy equality by lowering the numerator. The findings indicate indeed clear evidence for the negative impact of spousal age gaps on gender parity in literacy. Coefficients are always significant at the one percent level, and also large in absolute terms. The estimates suggest for example that one additional year between husband and wife has a larger effect on equal literacy levels between men and women than an increase in income levels by one log unit. In the IV specification with full controls, an increase of the spousal age gap by one year leads to a reduction of the gender parity in literacy by ten percentage points – roughly the difference between Italy and Iran. Lastly, also the spousal age gap ratio is highly significant and negative. This means that a one unit change in the spousal age gap ratio causes a roughly two times inverse effect on the female-to-male literacy ratio.

Marriage age effects on the pure attainment measure of primary schooling are mixed, which is actually in line with our expectations: primary schooling is basic education and not yet systematically influenced by marriage decisions. The findings also indicate that primary schooling does not fully explain differences in literacy rates as measure by the GPI. This suggests that more time is required, i.e. enrollment into secondary schooling, to reach proper literacy. We have documented earlier that marriage age does affect female secondary schooling.

In a further approach towards alternative outcome variables, we examine effects on the female average years of total schooling per country, both in absolute terms, and relative to the male average years of schooling (see table 3.17 in the appendix). This allows to consider the impact on primary, secondary, and tertiary education in an integrated perspective. Results indicate that the marriage age for women remains a largely robust and significant determinant for their overall education, both in absolute and in relative perspective. For instance, each year of marriage delay translates into about four months of longer total schooling. Hence, we are able to show that the total length of education is affected by marriage timing, consistent with our theoretical considerations.

Table 3.6: Cross-section for alternative education outcome variables

2000s: Dependent variable =	Gender Parity in Literacy (GPI)			Primary Schooling Completion Ratio		
Panel A: Spousal Age Gap (OLS)	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	135	129	111	133	131	119
Spousal Age Gap	-0.08	-0.02	-0.03	0.44	0.07	-0.04
	(0.01)***	(0.01)***	(0.01)***	(0.17)***	(0.26)	(0.04)
Fertility		-0.06	-0.05		0.97	0.17
		(0.01)***	(0.01)***		(0.27)***	(0.04)
Urbanization		0.01	0.01		0.01	-0.01
		(0.01)	(0.01)		(0.01)	(0.01)
Share of Muslim population		-0.07	-0.04		-0.18	0.29
		(0.03)**	(0.04)		(0.93)	(0.15)*
Sex ratio at birth		-0.74	-0.57		5.95	1.54
		(0.42)*	(0.48)		(13.68)	(1.87)
Sex ratio under 5 mortality		-0.87	-0.22		-7.23	-0.29
		(0.46)*	(0.59)		(10.99)	(1.65)
Sex ratio under 15 mortality		0.85	0.26		8.36	0.44
		(0.46)*	(0.60)		(11.07)	(1.65)
Cum. pop. married at 40		0.21	0.39		0.07	-0.88
		(0.23)	(0.31)		(7.33)	(0.99)
French Legal Origin		-0.04	-0.03		0.82	0.08
		(0.02)**	(0.02)		(0.57)	(0.08)
Female Labor Force Participation			0.01			-0.01
			(0.01)			(0.01)
Log GDP per Capita			0.03			0.02
			(0.02)			(0.06)
Continent dummies	no	yes	yes	no	yes	yes
R-squared	0.43	0.68	0.65	0.04	0.10	0.32

Panel B: Spousal Age Gap (IV)	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	109	106	96	114	113	103
Spousal Age Gap	-0.12	-0.11	-0.10	0.27	-1.76	-0.02
	(0.01)***	(0.04)***	(0.03)***	(0.27)	(0.93)*	(0.07)
Fertility		-0.02	-0.03		1.54	0.12
		(0.02)	(0.02)*		(0.43)***	(0.04)***
Urbanization		-0.01	-0.01		-0.04	-0.01
		(0.01)	(0.01)		(0.03)	(0.01)
Share of Muslim population		0.11	-0.10		2.54	0.22
		(0.08)	(0.07)		(1.86)	(0.16)
Sex ratio at birth		-0.19	0.03		28.57	2.39
		(0.64)	(0.62)		(18.85)	(1.67)
Sex ratio under 5 mortality		-1.23	-0.25		-1.20	0.45
		(0.66)*	(0.80)		(14.58)	(1.62)
Sex ratio under 15 mortality		1.41	0.41		3.70	-0.32
		(0.67)**	(0.78)		(14.52)	(1.59)
Cum. pop. married at 40		-0.54	-0.22		-11.67	-0.53
		(0.53)	(0.46)		(11.75)	(1.02)
French Legal Origin		-0.06	-0.04		1.35	0.09
		(0.03)	(0.03)		(0.74)*	(0.07)
Female Labor Force Participation			0.01			0.01
			(0.01)			(0.01)
Log GDP per Capita			0.03			0.01
			(0.02)			(0.05)
Continent dummies	no	yes	yes	no	yes	yes
First stage F-test statistics	101.94	10.88	15.98	111.40	14.04	19.96
R-squared	0.35	0.52	0.58	0.04	0.01	0.41

Table 3.6 continued: Cross-section foralternative education outcome variables

2000s: Dependent variable =	Gender Parity in Literacy (GPI)			Primary Schooling Completion Ratio		
Panel C: Spousal Age Gap Ratio (IV)	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	109	106	96	132	130	118
Spousal Age Gap Ratio	-2.05	-1.90	-1.59	3.28	-51.46	-4.02
	(0.21)***	(0.68)***	(0.48)***	(4.30)	(24.87)**	(1.83)**
Fertility		-0.01	-0.03		2.08	0.22
		(0.02)	(0.02)*		(0.64)***	(0.05)***
Urbanization		-0.01	-0.01		-0.05	-0.01
		(0.01)	(0.01)		(0.03)	(0.01)
Share of Muslim population		0.01	0.01		3.49	0.48
		(0.06)	(0.05)		(2.09)*	(0.19)***
Sex ratio at birth		0.10	0.04		39.49	3.43
		(0.64)	(0.56)		(24.18)*	(2.18)
Sex ratio under 5 mortality		-0.42	0.41		8.43	1.07
		(0.60)	(0.67)		(16.60)	(1.86)
Sex ratio under 15 mortality		0.43	-0.36		-6.17	-0.92
		(0.60)	(0.67)		(16.48)	(1.86)
Cum. pop. married at 40		-0.01	0.27		-17.15	-1.63
		(0.30)	(0.34)		(12.71)	(1.12)
French Legal Origin		-0.04	-0.02		1.15	0.10
		(0.03)	(0.03)		(0.80)	(0.08)
Female Labor Force Participation			0.01			-0.01
			(0.01)			(0.01)
Log GDP per Capita			0.01			-0.01
			(0.02)			(0.06)
Continent dummies	no	yes	yes	no	yes	yes
First stage F-test statistics	89.34	9.54	18.86	106.78	7.94	16.82
R-squared	0.39	0.51	0.60	0.04	0.01	0.25

Notes: The dependent variable in columns (1)-(3) is the Gender Parity in Literacy Index (GPI), and in (4)-(6) the ratio of male over female primary schooling completion rates. The regressors are: (i) Spousal Age Gap (*Male* minus *Female Marriage Age*), respectively Spousal Age Gap Ratio (*Male* over *Female Marriage Age*); (ii) the total fertility rate; (iii) the level of urbanization in percent; (iv) the share of muslim population per country; (v)-(vii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (viii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of tota female population); (ix) dummy whether the country's legal origin is based on French system; (x) the percentage of females in the national labor force; (xi) th log of per capita GDP in PPP terms; (xii) six continent dummies. Panels B and C use as instrument the weighted average of the spousal age gap (ratio) of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level,

Additional Controls and Test on Unobservables

Next we examine whether effects remain robust if we include a set of additional controls. We begin by considering potential issues around the country selection in our data sample. As our objective is to provide conclusions that are valid independent of geographies, we already include continent dummies regularly in our regressions. We also noted earlier that we include countries from all world regions so that selection bias seems unlikely. Here, we revisit the question whether our proposed relationship between marriage age and education depends on the state of development in a country. We systematically control for income levels and state of urbanization in our estimates, which should capture this cross-country variation reasonably well. In addition, when including fixed effects, the country-specific "poverty" versus "wealth" effects should be removed as well.

Nonetheless, we now test more explicitly if marriage age effects on education are present in developing countries only. We include a dummy in our panel data for all OECD countries, which indicates an advancement from a development state to more sustained prosperity. The empirical estimates show that the results obtained so far are robust to the inclusion of this control (see table 3.18 in the appendix). The OECD dummy obviously has a positive effect on female education, but not at the expense of our core variable of interest, be it *Female Marriage Age* or the Spousal Age Gap. We also employ an alternative dummy as cross-check, which codes all developing countries with one; consistently, the coefficient for this dummy turns negative but again our core variable of interest remains robust.[56] Thus we correctly attribute educational gender differences to marriage age, as the simple distinction of developing versus developed country is not able to better explain our dependent variable.[57]

The second dimension of additional controls relates to gender discrimination variables. These, however, are not available for the entire timeframe of our panel so that data restrictions suggest a cross-section, averaging the years 2000-2010. We summarize in table 3.7 the estimates for the quasi diff-in-diff specification, extended by a comprehensive list of covariates which have been proposed in the literature for gender discrimination in education (Barro & Lee, 2013; Cohen & Soto, 2007; Ross, 2008). The objective is to check whether the effect of spousal age gaps can in fact be closely replicated by employing commonly used variables of gender discrimination.

First we add as variable the legal minimum marriage age for women. It might be that our proposed causal relationship is biased from legal restrictions on marriage age which are different for each country. This could be particularly relevant towards the later stages of education when minimum marriage age laws potentially act as additional barrier against marring young and quitting education. The empirical results cannot support such a hypothesis. However, the fact that the inclusion of this control does not render the spousal age gap ratio insignificant strengthens the main results of this work. There is additional potential for educational gender equality through

[56] The same effect can be observed when coding all African countries with one, i.e. an African regional dummy does not change the core results.

[57] We note that the OECD dummy has a more significant effect on tertiary than on secondary education. Also, we experimented with regressions for a sample that excludes all OECD countries. However, we consider the remaining number of developing countries (ca. 40) as too small, and it likewise overly limits the degrees of freedom when including our sizeable set of regular controls. Hence, estimates appear unreliable and we do not pursue them further.

Table 3.7: Cross-section for educational gender inequality with discrimination controls

Panel A

2000s: Dependent variable =	Secondary Schooling Completion Ratio							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	118	113	113	105	105	105	105	90
Spousal Age Gap Ratio	4.30	3.44	4.29	5.02	4.34	4.37	4.29	4.60
	(2.16)**	(2.02)*	(2.27)**	(2.27)**	(2.24)**	(2.27)**	(2.22)**	(2.36)**
Minimum Legal Marriage Age		-0.04						-0.02
		(0.03)						(0.04)
Gender Inequality Index (GII)			0.26					0.51
			(0.50)					(0.83)
Women Treated with Respect				-0.01				-0.01
				(0.01)				(0.01)
Women's economic rights					0.04			-0.15
					(0.10)			(0.17)
Women's political rights						0.11		0.24
						(0.07)		(0.13)*
Women's social rights							0.06	0.13
							(0.08)	(0.15)
Additional Controls	yes	yes	yes	yes	yes	yes	yes	yes
First stage F-test statistics	10.73	10.98	9.39	9.73	10.68	10.23	10.50	6.90
R-squared	0.24	0.34	0.25	0.16	0.24	0.24	0.25	0.28

Panel B

2000s: Dependent variable =	Average Years Tertiary Education Ratio							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	118	113	113	105	105	105	105	90
Spousal Age Gap Ratio	10.44	10.28	10.67	12.15	10.65	10.93	10.80	14.91
	(5.11)**	(5.24)**	(5.34)**	(5.38)**	(5.09)**	(5.22)**	(5.12)**	(6.54)**
Minimum Legal Marriage Age		0.11						0.25
		(0.07)						(0.10)***
Gender Inequality Index (GII)			-1.23					-1.68
			(1.03)					(1.77)
Women Treated with Respect				0.01				0.01
				(0.01)				(0.01)
Women's economic rights					0.27			0.04
					(0.19)			(0.42)
Women's political rights						0.20		0.21
						(0.13)		(0.23)
Women's social rights							0.10	0.01
							(0.15)	(0.31)
Additional Controls	yes	yes	yes	yes	yes	yes	yes	yes
First stage F-test statistics	10.73	10.98	9.39	9.73	10.68	10.23	10.50	6.90
R-squared	0.29	0.30	0.25	0.10	0.27	0.25	0.25	0.04

Table 3.7 continued: Cross-section for educational gender inequality with discrimination controls

Panel C								
2000s: Dependent variable =	Gender Parity in Literacy (GPI)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	110	103	101	93	99	99	99	75
Spousal Age Gap Ratio	-1.59	-1.46	-1.69	-1.69	-1.54	-1.60	-1.55	-1.40
	(0.61)***	(0.66)**	(0.80)**	(0.71)**	(0.60)***	(0.63)***	(0.61)***	(0.94)
Minimum Legal Marriage Age		0.01						0.01
		(0.01)						(0.01)
Gender Inequality Index (GII)			0.01					-0.13
			(0.17)					(0.27)
Women Treated with Respect				-0.01				0.01
				(0.01)				(0.01)
Women's economic rights					-0.03			-0.01
					(0.04)			80.07)
Women's political rights						-0.02		-0.09
						(0.03)		(0.05)*
Women's social rights							-0.02	-0.03
							(0.02)	(0.05)
Additional Controls	yes	yes	yes	yes	yes	yes	yes	yes
First stage F-test statistics	16.45	13.50	9.49	12.79	15.75	14.03	12.88	5.26
R-squared	0.60	0.63	0.57	0.60	0.61	0.60	0.61	0.67

Notes: The dependent variable in Panel A is the ratio of the male over female population share with a completed secondary education. Panel B estimates the ratio of the male over female average years of tertiary schooling, and Panel C estimates the Gender Parity in Literacy Index (GPI). The regressors are: (i) The Spousal Age Gap ratio (*Male* over *Female Marriage Age*); (ii) the minimum legal age of marriage for women without parental consent, taken from the United Nations Statistics Division ; (iii) the Gender Inequality Index from the United Nations Development Programme (UNDP); (iv) the percentage of female respondents answering yes to the question, "Do you believe that women in this country are treated with respect and dignity, or not?", and taken from Gallup World Poll; (v)-(vii) three female discrimination measures as taken from the CIRI Human Rights Data Project; (viii) is a vector of additional controls: the total fertility rate; the level of urbanization in percent; the share of muslim population per country; the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); dummy whether the country's legal origin is based on French system; the percentage of females in the national labor force; the log of per capita GDP in PPP terms; six continent dummies.Spousal Age Gap Ratio is instrumented by the weighted average of the spousal age gap ratio of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

reducing spousal age gaps, on top of what current marriage age legislation already attempts to mitigate. We then consider the Gender Inequality Index (UNDP, 2013) which measures gender inequality along three dimensions and reaches values of zero for total equality. This also leaves our core relationship between spousal age gaps and educational inequality unaffected.

A further gender discrimination variable tested deviates from an aggregate index perspective. We control for the percentage of agreement among females to a question from the 2014 Gallup World Poll.[58] Women are asked whether they "believe that women in this country are treated with respect and dignity". This survey variable on gender equality sentiment per country represents a different perspective on discrimination than aggregate indices. Nonetheless, our results remain robust to the inclusion of this variable as well.

[58] We take the values from the dataset comprising the 2014 Social Progress Index (Porter, Stern, & Green, 2014).

Finally, we test three control variables for female discrimination as measured through the Cingranelli-Richards (CIRI) Human Rights Dataset. These have no material impact on our variable of interest. Even when all discrimination variables are included simultaneously (column 8 per panel), our core relationship remains very stable, and significant in two out of three cases. However, in those specifications we note serious finite sample bias as reflected in low first stage F-values. Overall, results remain robust, and the coefficients for secondary and tertiary education are also similar to the panel estimates presented earlier (table 3.4, panel B).

Robustness is examined along a further dimension, namely potential endogeneity arising from omitted variable bias. Following Oster (2014), we run a test which exploits the variation of the coefficient β_1 of our key variable of interest (we take here the spousal age gap ratio) when including observed control variables in the regression. The objective is to minimize potential omitted variable bias. The method by Oster suggests that if the inclusion of observed covariates increases the explanatory power (R-squared) of the model substantially, but changes β_1 only marginally, then potential unobserved variables should not much affect the coefficient either, since the included controls capture already what the researcher considered as most relevant for a potential bias. This allows to determine the so-called identified set, which we calculate under the most conservative assumptions of equal selection between observed and unobserved controls ($\widetilde{\delta}$ = 1) and a maximum potential value for R-squared (Rmax = 1). If this set excludes zero, we may conclude that results are robust to potential omitted variable bias. We run the analysis for the cross-section of the 2000s.

Table 3.8: Oster (2014) tests: Potential bias from unobservables in cross-section

Dependent variable =	Secondary Schooling Completion Ratio	Average Years Tertiary Education Ratio	GPI Adult Literacy
Uncontrolled $\dot{\beta}_1$	2.848	8.462	-2.054
Controlled $\tilde{\beta}_1$	4.297	10.439	-1.588
Uncontrolled $\dot{R}^2$	0.078	0.277	0.391
Controlled $\widetilde{R^2}$	0.244	0.288	0.604
Identified set $[\tilde{\beta}_1, \beta_1{*}']$	[4.297, 10.934]	[10.439, 138.404]	[-1.588, -0.72019]
Zero excluded from identified set?	yes	yes	yes

Notes: This procedure of assessing potential bias from unobserved variables by looking at movements in coefficients for spousal age gap, and the R-squared when including observed covariates has been developed by Oster (2014). It is based on previous work by Altonji, Elder, and Taber (2005, 2008). The uncontrolled $\dot{\beta}_1$ is calculated without including any additional controls, while the controlled $\tilde{\beta}_1$ employs all *regular* control variables as seen in table 3.7 (column (1) in panels A through C). $\beta_1{*}'$ is calculated with an assumed value of $\widetilde{\delta}$ = 1, and Rmax = 1.

Results, presented in table 3.8, are reassuring for all three outcome variables. The identified set remains fully robust, as zero is excluded for all three specifications. We conclude that effects of unobserved controls are highly unlikely to have a sizeable impact on the observed effect of spousal age gaps on gender inequality in secondary schooling, tertiary education, and adult literacy.

Generational Effects of Marriage Age

Based on an intra-generational perspective, we have previously argued that societal expectations at a given period regarding the "ideal timing" for marriage act as underlying mechanism. Now we want to investigate whether there might be an alternative channel to current societal expectations, namely the traditional legacy of parental marriage age. This would imply intergenerational effects from spousal age gaps on education which we could not capture thus far. In that line of thought, the hypothesis here tests whether the age gap between husband and wife also affects their children's education by disadvantaging the daughters. This is because in addition to current societal expectations, female children would be expected to time their marriage also based on the parents' tradition, so that returns from educational investments in girls would depend on that factor as well. This would ultimately suggest that children's education is partly pre-determined by the relative marriage age of their parents.

For our inter-generational extension, individual-level data which would directly link the spousal age gap of the actual parents to their daughters' education levels are not available. We have to resort to a 20-year lag of the average national spousal age gap ratio as explanatory variable, which proxies a one generation timeframe. We then measure effects on female education levels, keeping our regular instrumentation specification in addition to the built-in time lag. Data availability for marriage age allows to estimate this specification only for education levels in the 2000s, using parental marriage age data from the 1980s.

Results are given in table 3.9. The estimates indicate a clearly significant effect from the parental generation marriage age pattern to current educational gender inequality. Doubling the age gap between "mother and father" is associated with a three times higher inequality between "son and daughter" in secondary schooling, and even ten times higher inequality in tertiary education; literacy inequality would be increased by a factor of 1.2. The figures confirm findings by Sekhri and Debnath (2014) on the role of parental marriage age for children's education, but we add the perspective that parental marriage age also impacts educational equality between boys and girls. In summary, spousal age gaps do not only affect the education within a generation, but also potentially impact educational gender equality of the children's generation.[59]

Table 3.9: Cross-section for educational gender inequality effects from parental generation

2000s Dependent variable =	Secondary Schooling Completion Ratio		Average Years Tertiary Education Ratio		Gender Parity in Literacy (GPI)	
Spousal Age Gap Ratio (IV)	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	87	79	87	79	76	68
Spousal Age Gap Ratio Parental Generation	4.22	2.98	6.69	9.78	-1.60	-1.21
	(1.12)***	(1.60)*	(1.67)***	(4.43)**	(0.17)***	(0.68)*
Additional Controls	no	yes	no	yes	no	yes
First stage F-test statistics	88.98	16.86	88.98	16.86	98.44	11.04
R-squared	0.10	0.37	0.20	0.09	0.54	0.72

Notes: The dependent variable in column (1)-(2) is the ratio of the male over female population share with a completed secondary education. Columns (3)-(4) estimate the ratio of the male over female average years of tertiary schooling, and columns (5) - (6) estimate the Gender Parity in Literacy Index (GPI). The regressors are: (i) The Spousal Age Gap ratio (*Male* over *Female Marriage Age*), and Additional Controls: (ii) the total fertility rate; (iii) the level of urbanization in percent; (iv) the share of muslim population per country; (v)-(vii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (viii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (ix) dummy whether the country's legal origin is based on French system; (x) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xii) six continent dummies. Spousal Age Gap Ratio is instrumented by the weighted average of the spousal age gap ratio of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Marriage Age Effects on Further Aspects of Gender Inequality

Finally, in table 3.10, we investigate whether spousal age gaps affect further spheres of gender inequality. We first re-visit educational gender inequality and address the notion that measurement issues of human capital focusing on school attainment have received criticism (Hanushek, 2015). The quality of education has been proposed as another important and complementary variable to quantitative enrollment rates, as it measures the effectiveness of accumulating human capital (Castelló-Climent & Hidalgo-Cabrillana, 2012; Hanushek & Woessmann, 2009).[60] We hence employ the PISA 2006 data set for our latest time interval as alternative outcome variable. We then attempt to estimate the ratio of national test scores of boys over girls. Results suggest that there is no robust relationship between test scores and spousal age gaps. Thus, spousal age gaps are an important determinant for how long a young female attends

[59] We also ran an alternative specification using ten year lags of spousal age gaps, with similar effects. In general, previous spousal age gaps predict spousal age gaps for the next decade highly significantly, but with a coefficient of ca. 0.6, so effects decrease in absolute terms over time.

[60] Barro and Lee (2013) note that educational attainment and human capital quality measures show high correlation, but still the latter appears more diverse for countries with similar levels of educational attainment. Guiso et al. (2008) explicitly relate the achievement gap between boys and girls in PISA text exams with indicators of a gender-equal culture. However, Fryer and Levitt (2010) show that results are not robust to including a group of Middle Eastern countries.

schooling in the first place. Once this tollgate of (potential) gender discrimination has been passed, there is no empirical evidence that the subsequent quality of education received should differ between genders depending on the marriage age. Admittedly though, the sample size is limited, and results may show selection bias since the test scores are only collected for girls who could enroll to school, i.e. girls who suffered less from educational gender inequality in the first place.

Table 3.10: Further gender inequality effects from spousal age gaps

2000s Dependent variable =	Quality of Schooling (PISA scores) male relative to female		Adolescent births		Women in Politics	
Spousal Age Gap Ratio (IV)	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	49	48	150	133	139	120
Spousal Age Gap Ratio	-0.33	-0.07	474.15	228.38	-56.39	-54.28
	(0.15)**	(0.31)	(42.52)***	(79.63)***	(12.99)***	(27.04)**
Additional Controls	no	yes	no	yes	no	yes
First stage F-test statistics	36.06	8.69	69.68	22.39	80.16	24.20
R-squared	0.01	0.54	0.51	0.84	0.07	0.36

Notes: The dependent variable in column (1)-(2) is the ratio of the male over female student performance on the science scale for the PISA test 2006. Columns (3)-(4) estimate the number of births to women with age 15–19 per 1,000 women with age 15–19. Columns (5)-(6) are the proportion of seats in parliament held by women, measured in 2000. The variable ranges from 0 to 100. The regressors are: (i) The Spousal Age Gap ratio (*Male* over *Female Marriage Age*), and Additional Controls: (ii) the total fertility rate; (iii) the level of urbanization in percent; (iv) the share of muslim population per country; (v)-(vii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (viii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (ix) dummy whether the country's legal origin is based on French system; (x) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xii) six continent dummies. Spousal Age Gap Ratio is instrumented by the weighted average of the spousal age gap ratio of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

For our remaining two outcome variables that measure further aspects of gender inequality and female discrimination, we find strong and robust effects. Spousal age gaps are associated with a significant increase in teenage pregnancies that is also large in absolute size and robust to fertility levels in a country. We interpret this as supporting evidence for our proposed conceptual framework: marrying (relatively) earlier as a woman means a significantly higher likelihood to also conceive children early, i.e. a higher teenage pregnancy rate. Consequently, because of young marriage age and subsequent child-rearing, women drop out of the labor market at a comparatively younger age, which reduces their expected return on education investments. Similarly, larger husband-wife age gaps decrease female participation in politics as measured through the proportion of seats in parliament. Based on this additional evidence of spousal age gaps affecting gender inequality we are confident that the core argument of this study – the causal link from marriage age to educational gender inequality – is no accidental empirical finding, but part of a broader robust pattern.

III. 6. Concluding Remarks

Bridging the gender gap in education is a key challenge in the world, as women on average still receive less schooling than men, which affects their livelihoods as well as the growth prospects of a country. Equal opportunities for women and economic development are closely interrelated, but the interrelationships are often too weak to be self-sustaining so that public policies are needed (Duflo, 2012). One specific societal aspect that falls under potential policy intervention relates to the minimum marriage age. Entering marriage early as a girl or young woman is widely associated with a lack of gender equality.

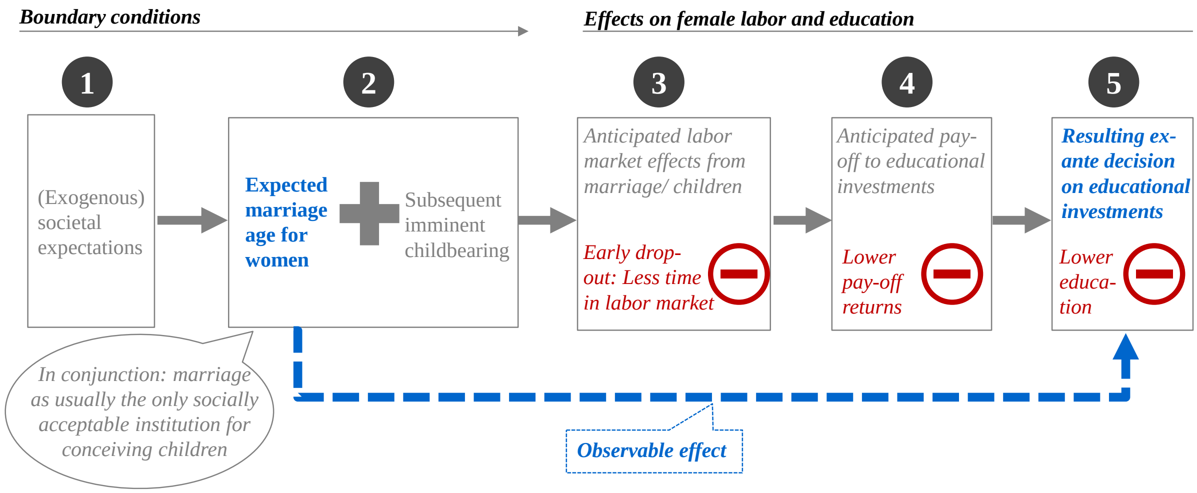
This chapter investigated how marriage age influences female education and educational gender inequality. Specifically, we answered the question whether getting married younger as a woman, both in absolute perspective and in comparison to the husband's age, leads to worse female education. Since marriage is still widely the only socially accepted institution for conceiving children, the anticipated marriage age also proxies the expected age of first birth for women. Hence, the earlier a woman on average expects to get married, the shorter will be her anticipated pay-off to educational investments. Acting rationally, that investment will be adjusted downwards already ex ante, so that in essence lower female marriage age leads to lower female education.

We proceeded to empirically support our hypothesis by estimating marriage age effects on various measures of female education. Specifically, we applied a global cross-country panel data set from 1980-2010, in which we instrumented the domestic female marriage age with an average of the marriage age in adjacent countries weighted by shared land border. The absolute female age at marriage has indeed a highly significant effect on female education: In our preferred specification, each year of marriage postponement for women leads to a three percentage points higher female completion rate in secondary schooling, and to about three weeks longer female tertiary education.

We then moved to examining spousal age gap effects, i.e. female relative to male marriage age, to address potential biases in case women in one part of the world get married earlier than somewhere else, irrespective of education levels. We obtained similarly robust effects in our panel, where each additional year between husband and wife lowers the female secondary schooling completion rate by 10 percentage points, and cuts the time women spend at university by one month. Finally, quasi diff-in-diff specifications indicated that spousal age gaps affect female education significantly more negatively than male education. Numerous robustness checks confirmed our findings, and in a cross-section we conducted a number of refinement analyses. Estimates suggested that gender parity in literacy is strongly affected by spousal age gaps, but pure primary schooling attainment is not. Moreover, the marriage pattern of the parental generation also influences the children's educational gender inequality. Importantly, we documented that marriage age and conventional measures of gender discrimination do not act as substitutes.

Building up human capital through sufficient education is a key mechanism to empower women. In addition to compulsory schooling laws, our results suggest that governments have the regulation of minimum marriage age at their disposal to influence how much schooling young females receive. A lower barrier for marriage age in our framework would set a minimum threshold of "guaranteed" pay-off to educational investments, since the likelihood of women becoming mothers at a young age and thus dropping out of the labor market from early on decreases. In contrast to legal boundaries for fertility or maternal age at first birth, both of which raise strong ethical concerns especially with regards to implementation enforcements, marriage age as the underlying cause lends itself much better to formal regulation. In essence we believe that pregnancy rates at very young ages, which likely lead to female drop-out of education and thus gender gaps in education, could be effectively reduced by enforcing adequate minimum marriage age laws.

We recognize that our macro-perspective does not allow to account for individual differences in skills and ability, which may make further education unsuitable independent of marriage plans. Furthermore, there are of course more dimensions than human capital where gender equality is critical, such as access to markets and decision-making power within the household, political empowerment, health, and many more. But with the documented gender-specific negative impact of a young female marriage age on female education, the case for more rigorous minimum marriage age laws around the world appears to be justified.



Appendix Figure 3.1: Conceptual framework for female marriage age effects on female education

Appendix Table 3.11: Hausman Test

1980-2010								
Dependent variable =	Female Secondary Schooling Completion				Average Female Years Tertiary Education			
	Fixed	Random	Difference	St. Error	Fixed	Random	Difference	St. Error
Female Marriage Age	1.1477	0.8334	0.3143	0.2322	0.0238	0.0230	0.0008	0.0029
Fertility	-3.5572	-3.3160	-0.2412	0.7745	0.0132	0.0136	-0.0004	0.0105
Urbanization	-0.0043	0.0268	-0.0311	0.1438	0.0023	0.0024	-0.0001	0.0021
Share of Muslim population	46.0170	2.4950	43.5220	18.3068	0.0924	0.0266	0.0659	0.2853
Sex ratio at birth	-4.4022	-6.9957	2.5935	2.0934	-0.0861	-0.0879	0.0018	<0.0001
Sex ratio under 5 mortality	-18.0639	-13.2141	-4.8498	25.3503	-2.2736	-1.7060	-0.5676	0.3036
Sex ratio under 15 mortality	3.3904	3.6200	-0.2300	29.9149	2.0977	1.3733	0.7244	0.3811
Cum. pop. married at 40	-17.0570	-30.9666	13.9097	14.4322	-1.0577	-1.2918	0.2341	0.1691
French Legal Origin	*n/a (omitted because of collinearity)*							
Female Labor Force Participation	-0.0798	0.0375	-0.1173	0.0617	0.0031	0.0042	-0.0011	0.0008
Log GDP per Capita	5.9773	1.9485	4.0288	1.6563	0.1564	0.0827	0.0737	0.0220
Chi-Square (p-value)	26.50 (0.003)				57.12 (<0.001)			

Appendix Table 3.12: First stage of panel for level of female education, absolute and relative

1980-2010 (IV First Stage)								
Dependent variable =	Female Marriage Age (Panel A)				Spousal Age Gap (Panel C)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	71	70	66	66	71	70	66	66
Observations	213	210	198	198	213	210	198	198
Female Marriage Age of Bordering Countries	0.79	0.51	0.46	0.42				
	(0.05)***	(0.07)***	(0.07)***	(0.10)***				
Spousal Age Gap of Bordering Countries					-0.25	-0.15	-0.13	0.01
					(0.03)***	(0.03)***	(0.03)***	(0.03)
Fertility		-0.28	-0.29	-0.42		0.09	0.02	0.19
		(0.16)*	(0.17)*	(0.26)*		(0.07)	(0.08)	(0.09)**
Urbanization		0.04	0.02	0.01		-0.02	-0.02	-0.01
		(0.01)***	(0.02)	(0.04)		(0.01)***	(0.01)***	(0.01)
Share of Muslim population		-0.28	-0.21	-4.05		1.26	0.94	6.69
		(0.79)	(0.90)	(4.14)		(0.25)***	(0.28)***	(1.41)***
Sex ratio at birth		-4.91	1.17	-4.84		1.65	-4.21	-4.89
		(11.95)	(12.35)	(19.57)		(4.65)	(4.78)	(6.64)
Sex ratio under 5 mortality		-19.71	-16.25	-17.84		-2.03	-0.40	-0.41
		(7.36)***	(7.61)**	(9.80)*		(3.25)	(3.40)	(3.32)
Sex ratio under 15 mortality		22.39	18.86	23.24		3.91	2.35	-1.28
		(7.75)***	(8.00)**	(10.73)**		(3.36)	(3.49)	(3.64)
Cum. pop. married at 40		-10.22	-13.04	-15.74		-7.18	-6.56	3.76
		(4.69)**	(4.74)***	(6.39)**		(2.14)***	(2.16)***	(2.17)
French Legal Origin		-0.54	-0.48			0.56	0.45	
		(0.47)	(0.50)			(0.14)***	(0.15)***	
Female Labor Force Participation			-0.01	-0.01			-0.01	0.01
			(0.01)	(0.02)			(0.01)	(0.01)
Log GDP per Capita			0.84	0.79			-0.06	0.11
			(0.33)***	(0.53)			(0.13)	(0.18)
Continent dummies	no	yes	yes	no	no	yes	yes	no
Fixed effects	no	no	no	yes	no	no	no	yes
F-test	274.96	52.93	41.05	19.07	76.06	19.68	15.54	8.91
Angrist-Pischke F-statistics (p-value)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.64
Anderson LM test (p-value)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.64
Cragg-Donald Wald F statistic	274.96	52.93	41.05	19.07	76.06	19.68	15.54	8.91
Stock-Yogo critical values 10%					16.38			
Stock-Yogo critical values 25%					5.53			
R-squared	0.57	0.64	0.66	0.60	0.27	0.69	0.68	0.24

Notes: The dependent variable in column (1)-(4) is the *Female Marriage Age*. Columns (5)-(8) estimate the Spousal Age Gap (*Male* minus *Female Marriage Age*). The regressors are: (i) The weighted average of the *Female Marriage Age* of adjoining countries with a common border, where weights are according to relative length of shared border; (ii) the weighted average of the Spousal Age Gap of adjoining countries with a common border, where weights are according to relative length of shared border; (iii) the total fertility rate; (iv) the level of urbanization in percent; (v) the share of muslim population per country; (vi) the sex ratio at birth (males over females); (vii) the mortality rate of boys over the mortality rate of girls under 5, and (viii) under 15 years, respectively; (ix) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (x) a dummy whether the country's legal origin is based on French system; (xi) the percentage of females in the national labor force; (xii) the log of per capita GDP in PPP terms; (xiii) six continent dummies. See the Appendix for more detailed variable definitions and sources. Robust Standard

Appendix Table 3.13: First stage of panel for relative male-to-female education levels (ratio)

1980-2010 (IV First Stage)

Dependent variable =	Spousal Age Gap Ratio (Panel B)			
	(1)	(2)	(3)	(4)
Number of countries	71	70	66	66
Observations	213	210	198	198
Spousal Age Gap Ratio of Bordering Countries	0.84	0.54	0.50	-0.06
	(0.06)***	(0.07)***	(0.07)***	(0.13)
Fertility		0.01	-0.01	0.02
		(0.01)	(0.01)	(0.01)***
Urbanization		-0.01	-0.01	-0.01
		(0.001)***	(0.001)***	(0.01)
Share of Muslim population		0.06	0.04	0.32
		(0.01)***	(0.02)***	(0.08)***
Sex ratio at birth		-0.06	-0.40	-0.27
		(0.25)	(0.25)	(0.40)
Sex ratio under 5 mortality		0.16	0.16	0.09
		(0.16)	(0.17)	(0.19)
Sex ratio under 15 mortality		-0.11	-0.09	-0.21
		(0.17)	(0.17)	80.21)
Cum. pop. married at 40		-0.20	-0.12	0.30
		(0.11)	(0.11)	(0.12)**
French Legal Origin		0.02	0.02	
		(0.01)***	(0.01)**	
Female Labor Force Participation			-0.01	0.01
			(0.01)	(0.01)
Log GDP per Capita			-0.01	0.01
			(0.01)	(0.01)
Continent dummies	no	yes	yes	no
Fixed effects	no	no	no	yes
F-test	216.92	54.75	48.60	5.18
Angrist-Pischke F-statistics (p-value)	<0.001	<0.001	<0.001	0.61
Anderson LM test (p-value)	<0.001	<0.001	<0.001	0.61
Cragg-Donald Wald F statistic	216.92	54.75	48.60	5.18
Stock-Yogo critical values 10%	16.38			
Stock-Yogo critical values 25%	5.53			
R-squared	0.52	0.75	0.74	0.28

Notes: The dependent variable is the Spousal Age Gap Ratio (*Male* over *Female Marriage Age*). The regressors are: (i) The weighted average of the Spousal Age Gap Ratio of adjoining countries with a common border, where weights are according to relative length of shared border; (ii) the total fertility rate; (iii) the level of urbanization in percent; (iv) the share of muslim population per country; (v) the sex ratio at birth (males over females); (vi) the mortality rate of boys over the mortality rate of girls under 5, and (vii) under 15 years, respectively; (viii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (ix) a dummy whether the country's legal origin is based on French system; (x) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xii) six continent dummies. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 3.14: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age with Age at First Birth Control

1980-2010								
Dependent variable =	Female Secondary Schooling Completion				Average Female Years Tertiary Education			
Panel A: Female Marriage Age (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	56	56	52	52	56	56	52	52
Observations	168	168	156	156	168	168	156	156
Female Marriage Age	3.36	2.05	1.66	7.11	0.05	0.05	0.04	0.10
	(1.64)**	(0.98)**	(0.82)**	(10.74)	(0.02)**	(0.02)**	(0.03)*	(0.05)**
Mother Age at First Child	-0.64	-0.98	-0.19	-3.38	0.01	-0.01	0.01	-0.07
	(1.60)	(1.74)	(1.69)	(9.18)	(0.03)	(0.03)	(0.03)	(0.13)
Fertility		-3.47	-3.86	-0.50		0.07	0.09	0.04
		(1.76)**	(1.83)**	86.08)		(0.03)**	(0.03)***	(0.09)
Urbanization		0.08	0.05	0.24		0.01	0.01	-0.01
		(0.10)	(0.13)	(0.55)		(0.001)***	(0.001)***	(0.01)
Share of Muslim population		5.16	5.82	99.33		-0.06	0.03	0.32
		(5.66)	(5.97)	(78.03)		(0.09)	(0.09)	(1.11)
Sex ratio at birth		21.44	-45.84	-256.53		2.86	4.07	-2.27
		(132.38)	(139.75)	(478.37)		(2.22)	(2.18)*	(6.81)
Sex ratio under 5 mortality		-10.41	19.14	-9.14		-2.00	-1.04	-4.09
		(56.08)	(63.97)	(179.28)		(1.00)**	(1.01)	(2.55)
Sex ratio under 15 mortality		10.24	-19.72	-70.70		2.07	1.15	4.60
		(54.99)	(62.96)	(133.16)		(0.97)**	(1.00)	(1.89)**
Cum. pop. married at 40		-34.65	-22.46	119.34		-0.88	-0.53	0.74
		(45.97)	(57.70)	(197.66)		(0.78)	(0.90)	(2.81)
French Legal Origin		-9.12	-9.95			-0.06	-0.03	
		(3.26)***	(3.58)***			(0.05)	(0.06)	
Female Labor Force Participation			0.09	-0.06			0.01	0.01
			(0.09)	(0.16)			(0.001)***	(0.01)
Log GDP per Capita			-0.55	-2.09			0.06	0.12
			(3.98)	(16.17)			(0.06)	(0.23)
Continent dummies	no	yes	yes	no	no	yes	yes	no
Country fixed effects	no	no	no	yes	no	no	no	yes
R-squared	0.32	0.57	0.57	0.07	0.43	0.68	0.76	
Panel B: Spousal Age Gap (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	56	56	52	52	56	56	52	52
Observations	168	168	156	156	168	168	156	156
Spousal Age Gap	-11.82	-1.63	-0.87	5.85	-0.05	-0.16	-1.22	0.08
	(3.39)***	(0.57)**	(0.51)*	(6.90)	(0.02)***	(0.07)**	(5.86)	(0.12)
Mother Age at First Child	-1.07	-0.65	6.36	3.25	0.04	-0.01	-0.15	0.02
	(3.31)	(1.65)	(16.43)	(1.21)***	(0.04)	(0.03)	(0.84)	(0.02)
Fertility		-3.64	-6.17	-4.23		0.06	0.12	-0.01
		(1.76)**	(6.30)	(3.00)		(0.03)*	(0.28)	(0.05)
Urbanization		0.15	0.62	0.08		0.01	-0.02	-0.01
		(0.20)	(2.01)	(0.36)		(0.01)	(0.10)	(0.01)
Share of Muslim population		2.91	-40.37	27.92		0.04	1.31	-0.67
		(8.95)	(145.78)	(45.89)		(0.17)	86.33)	(0.80)
Sex ratio at birth		11.95	-464.37	-162.27		4.07	15.79	-0.96
		(163.10)	(1292.27)	(351.77)		(3.07)	(59.25)	(6.16)
Sex ratio under 5 mortality		-32.72	349.98	133.50		-2.79	-10.72	-2.12
		(93.93)	(1011.01)	(104.37)		(1.78)	(47.46)	(1.83)
Sex ratio under 15 mortality		49.66	-392.15	-168.07		3.02	12.37	3.25
		(103.59)	(1138.56)	(105.99)		(1.96)	(54.45)	(1.86)*
Cum. pop. married at 40		-98.74	-2.61	-23.41		-2.91	-4.96	-1.24
		(89.13)	(175.21)	(49.87)		(1.70)	(17.83)	(0.87)
French Legal Origin		-9.06	-27.07			0.01	0.47	
		(6.27)	(50.37)			(0.12)	(2.59)	
Female Labor Force Participation			0.09	-0.08			0.01	0.01
			(0.26)	(0.13)			(0.02)	(0.01)
Log GDP per Capita			6.75	8.92			-0.03	0.27
			(17.73)	(5.20)*			(0.69)	(0.09)***
Continent dummies	no	yes	yes	no	no	yes	yes	no
Country fixed effects	no	no	no	yes	no	no	no	yes
R-squared	0.26	0.60	0.09	0.10	0.50	0.65	0.10	0.36

Appendix Table 3.14 continued: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age with Age at First Birth Control

1980-2010								
Dependent variable =	Secondary Schooling Completion Ratio				Average Years Tertiary Education Ratio			
Panel C: Spousal Age Gap Ratio	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	56	56	52	52	56	56	52	52
Observations	168	168	156	156	168	168	156	156
Spousal Age Gap Ratio	7.41	5.27	2.52	-0.42	0.27	1.32	1.68	-0.32
	(2.14)***	(2.32)**	(1.46)*	(0.31)	(0.14)**	(0.33)***	(1.87)	(1.53)
Mother Age at First Child	0.03	0.06	0.21	-0.11	-0.05	-0.01	0.12	-0.24
	(0.05)	(0.08)	(0.44)	(0.06)*	(0.04)	(0.06)	(0.32)	(0.28)
Fertility		0.23	0.17	0.19		0.19	0.33	0.05
		(0.18)	(0.23)	(0.14)		(012)	(0.29)	(0.71)
Urbanization		-0.01	0.03	-0.01		-0.01	0.04	-0.07
		(0.01)	(0.06)	(0.02)		(0.01)	(0.04)	(0.09)
Share of Muslim population		-0.84	-1.21	2.20		0.19	-1.92	-0.55
		(0.57)	(2.89)	(2.09)		(0.60)	(1.60)	(10.53)
Sex ratio at birth		10.16	-5.61	-25.01		4.07	-9.98	-19.14
		(10.14)	(30.96)	(16.00)		(7.58)	(23.90)	(82.78)
Sex ratio under 5 mortality		0.25	8.30	-7.82		3.51	6.80	-7.10
		(7.30)	(23.60)	(4.75)*		(5.01)	(17.81)	(24.64)
Sex ratio under 15 mortality		0.45	-9.73	8.84		-2.84	-9.50	6.93
		(7.01)	(27.18)	(4.82)*		(4.83)	(20.26)	(24.95)
Cum. pop. married at 40		-0.70	7.39	0.94		-0.12	8.06	-0.38
		(4.47)	(14.19)	(2.27)		(2.52)	(17.41)	(11.91)
French Legal Origin		0.38	-0.56			-0.22	-0.49	
		(0.32)	(1.38)			(0.26)	(1.13)	
Female Labor Force Participation			0.01	0.01			-0.01	-0.04
			(0.02)	(0.01)			(0.02)	(0.03)
Log GDP per Capita			-0.02	-0.11			0.19	1.12
			(0.46)	(0.24)			(0.58)	(1.28)
Continent dummies	no	yes	yes	no	no	yes	yes	no
Country fixed effects	no	no	no	yes	no	no	no	yes
R-squared	0.14	0.22	0.05	0.26	0.27	0.37	0.19	0.02

Notes: The dependent variable in column (1)-(4) is the percentage of the female population with a completed secondary education. Columns (5)-(8) estimate the average years of female tertiary schooling. The regressors are: (i) The *Female Marriage Age* (SMAM) in Panel A, and the Spousal Age Gap (*Male* minus *Female Marriage Age*) in Panels B (absolute) and C (relative); (ii) the average age of a woman at the birth of her first child; (iii) the total fertility rate; (iv) the level of urbanization in percent; (v) the share of muslim population per country; (vi)-(viii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (ix) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (x) dummy whether the country's legal origin is based on French system; (xi) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xiii) six continent dummies. French Legal Origin omitted in columns (4), and (8) because of inclusion of fixed effects. All panels use as instrument the weighted average of the absolute, respectively relative Female Marriage Age of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 3.15: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age with Age Dummies

1980-2010								
Dependent variable =	Female Secondary Schooling Completion				Average Female Years Tertiary Education			
Panel A: Female Marriage Age (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	71	70	66	66	71	70	66	66
Observations	213	210	198	198	213	210	198	198
Female Marriage Age	4.45	3.79	3.61	5.78	0.10	0.05	0.05	0.05
	(0.87)***	(1.46)***	(1.63)**	(2.53)**	(0.01)***	(0.02)**	(0.02)**	(0.03)*
Dummy_18	14.61	10.89	10.05	20.40	0.34	0.15	0.10	0.13
	(7.31)**	(8.87)	(9.46)	(13.93)	(0.12)***	(0.15)	0.15	(0.17)
Dummy_23	8.78	4.87	4.19	9.54	0.17	0.06	(0.05)	0.06
	(4.50)**	(5.25)	(5.53)	(7.88)	(0.08)**	(0.09)	0.09)	(0.10)
Dummy_28								
Fertility		-4.11	-4.57	-4.15		-0.01	-0.01	-0.01
		(1.04)***	(1.11)***	(1.78)**		(0.02)	(0.02)	(0.02)
Urbanization		-0.06	-0.13	-0.22		0.01	0.01	-0.01
		(0.09)	(0.11)	(0.21)		(0.001)***	(0.01)	(0.01)
Share of Muslim population		8.41	7.13	68.30		-0.04	0.04	0.12
		(5.62)	(6.51)	(26.74)***		(0.07)	(0.08)	(0.33)
Sex ratio at birth		27.57	4.95	52.01		0.05	0.01	-0.75
		(69.63)	(72.84)	(119.23)		(1.06)	(1.04)	(1.46)
Sex ratio under 5 mortality		52.54	61.22	103.21		-1.66	-1.28	-1.44
		(44.11)	(44.03)	(64.67)		(0.70)**	(0.67)**	(0.79)*
Sex ratio under 15 mortality		-62.74	-70.37	-136.94		1.59	1.20	1.29
		(49.28)	(49.15)	(77.28)*		(0.77)**	(0.73)*	(0.95)
Cum. pop. married at 40		-30.80	-28.39	11.74		-1.61	-1.59	-1.19
		(28.64)	(30.49)	(47.12)		(0.44)***	(0.46)***	(0.58)**
French Legal Origin		-6.92	-7.02			-0.01	0.02	
		(3.44)**	(3.89)			(0.04)	80.04)	
Female Labor Force Participation			-0.03	-0.13			0.01	0.01
			(0.07)	(0.11)			(0.001)***	(0.001)***
Log GDP per Capita			1.97	2.36			0.05	0.11
			(2.29)	(3.63)			(0.03)	(0.04)***
Continent dummies	no	yes	yes	no	no	yes	yes	no
Country fixed effects	no	no	no	yes	no	no	no	yes
R-squared	0.30	0.44	0.42	0.26	0.31	0.57	0.60	

Appendix Table 3.15 continued: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age with Age Dummies

1980-2010								
Dependent variable =	Female Secondary Schooling Completion				Average Female Years Tertiary Education			
Panel B: Spousal Age Gap (IV)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	71	70	66	66	71	70	66	66
Observations	213	210	198	198	213	210	198	198
Spousal Age Gap	-7.45	-2.07	-1.74	32.40	-0.18	-0.12	-0.06	0.29
	(2.42)***	(0.86)**	(0.78)**	(27.14)	(0.05)***	(0.07)**	(0.03)**	(0.29)
Dummy_18	-7.02	-11.64	-9.73	-29.66	-0.07	-0.10	-0.03	-0.31
	(5.44)	(5.84)**	(6.33)	(17.57)*	(0.11)	(0.11)	(0.15)	(0.19*
Dummy_23	-6.93	-8.30	-7.03	-11.54	-0.13	-0.07	-0.05	-0.13
	(3.15)**	(3.65)**	(3.43)**	(6.53)*	(0.07)**	(0.07)	(0.07)	(0.07)**
Dummy_28								
Fertility		-4.22	-3.92	-10.07		-0.01	-0.01	-0.06
		(0.92)***	(0.85)***	(4.10)***		(0.02)	(0.02)	(0.04)
Urbanization		0.20	0.10	0.25		0.01	0.01	0.01
		(0.08)***	(0.11)	(0.52)		(0.001)***	(0.01)	(0.01)
Share of Muslim population		-2.26	2.91	-166.17		-0.01	0.17	-1.97
		(5.25)	(5.68)	(178.09)		(0.10)	(0.16)	(1.93)
Sex ratio at birth		-32.20	-37.99	68.99		0.01	-0.08	-0.60
		(51.63)	(56.14)	(233.28)		(0.95)	(1.24)	(2.53)
Sex ratio under 5 mortality		-16.11	9.55	118.59		-1.63	-1.82	-1.30
		(36.89)	(39.47)	(131.24)		(0.68)**	(0.90)**	(1.43)
Sex ratio under 15 mortality		12.52	-8.73	-73.82		1.81	2.05	1.85
		(39.95)	(42.68)	(120.54)		(0.73)***	(0.99)**	(1.31)
Cum. pop. married at 40		-28.76	-47.40	-130.29		-2.48	-2.66	-2.46
		(36.77)	(36.75)	(92.17)		(0.67)***	(0.78)***	(1.00)***
French Legal Origin		-8.22	-6.79			0.01	0.05	
		(2.33)***	(2.43)***			(0.04)	(0.06)	
Female Labor Force Participation			0.09	-0.05			0.01	0.01
			(0.07)	(0.21)			(0.01)*	(0.01)*
Log GDP per Capita			1.52	1.49			0.04	0.10
			(1.54)	(7.38)			(0.03)	(0.08)
Continent dummies	no	yes	yes	no	no	yes	yes	no
Country fixed effects	no	no	no	yes	no	no	no	yes
R-squared	0.27	0.57	0.56	0.05	0.26	0.61	0.49	0.12

1980-2010								
Dependent variable =	Secondary Schooling Completion Ratio				Average Years Tertiary Education Ratio			
Panel C: Spousal Age Gap Ratio	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	71	70	66	66	71	70	66	66
Observations	213	210	198	198	213	210	198	198
Spousal Age Gap Ratio	6.95	5.72	3.83	-20.81	0.55	12.57	16.83	-11.71
	(1.39)***	(3.31)*	(1.71)**	(19.29)	(0.07)***	(5.87)**	(6.89)**	(14.96)
Dummy_18	-0.22	-0.23	-0.05	0.99	-0.22	-0.54	-0.65	0.63
	(0.29)	(0.41)	(0.20)	(0.93)	(0.24)	(0.41)	(0.45)	(0.73)
Dummy_23	-0.20	-0.07	0.08	0.33	-0.06	-0.07	-0.17	0.16
	(0.24)	(0.31)	(0.14)	(0.35)	(0.21)	(0.21)	(0.24)	(0.28)
Dummy_28								
Fertility		0.28	0.21	0.56		0.24	0.17	0.47
		(0.09)***	(0.04)***	(0.25)**		(0.07)***	(0.08)**	(0.18)***
Urbanization		-0.01	0.01	-0.01		0.01	0.01	-0.01
		(0.01)	(0.01)	(0.02)		(0.01)	(0.01)	(0.02)
Share of Muslim population		-0.27	0.16	4.85		-1.13	-1.48	3.09
		(0.37)	(0.18)	(6.03)		(0.61)*	(0.67)	(4.70)
Sex ratio at birth		4.87	4.39	-8.72		-4.94	-3.02	-11.29
		(5.23)	(2.59)*	(10.19)		(4.99)	(5.48)	(7.85)
Sex ratio under 5 mortality		-2.86	-1.58	-1.65		4.36	3.20	0.70
		(3.67)	(1.87)	(5.11)		(2.98)	(3.49)	(4.02)
Sex ratio under 15 mortality		3.70	2.24	-0.50		-4.77	-3.58	-3.11
		(3.79)	(1.91)	(5.16)		(3.09)	(3.62)	(4.02)
Cum. pop. married at 40		1.44	-0.28	4.47		1.83	1.30	6.06
		(2.73)	(1.31)	(5.05)		(1.92)	(2.14)	(3.75)
French Legal Origin		0.28	0.06			-0.19	-0.25	
		(0.18)	(0.09)			(0.23)	(0.25)	
Female Labor Force Participation			0.01	0.01			-0.01	-0.01
			(0.01)**	(0.01)**			(0.01)	(0.01)
Log GDP per Capita			-0.14	0.29			-0.04	-0.02
			(0.07)**	(0.30)			(0.16)	(0.22)
Continent dummies	no	yes	yes	no	no	yes	yes	no
Country fixed effects	no	no	no	yes	no	no	no	yes
R-squared	0.20	0.33	0.58	0.58	0.26	0.35	0.35	0.05

Notes: The dependent variable in column (1)-(4) is the percentage of the female population with a completed secondary education. Columns (5)-(8) estimate the average years of female tertiary schooling. The regressors are: (i) The *Female Marriage Age* (SMAM) in Panel A, and the Spousal Age Gap (*Male* minus *Female Marriage Age*) in Panels B (absolute) and C (relative); (ii)-(iv) three dummies for female age at marriage (18-23, 23-28, 28-32 years); (v) the total fertility rate; (vi) the level of urbanization in percent; (vii) the share of muslim population per country; (viii)-(x) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (xi) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (xii) dummy whether the country's legal origin is based on French system; (xiii) the percentage of females in the national labor force; (xiv) the log of per capita GDP in PPP terms; (xv) six continent dummies. French Legal Origin omitted in columns (4), and (8), because of inclusion of fixed effects, and Dummy_28 because of collinearity. All panels use as instrument the weighted average of the absolute, respectively relative Female Marriage Age of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 3.16: First stage of panel of female education using alternative instruments

1980-2010 (IV First Stage)								
Dependent variable =	Spousal Age Gap (Panel A)				Spousal Age Gap Ratio (Panel B)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of countries	86	81	86	81	86	81	86	81
Observations	258	243	258	243	258	243	258	243
Spousal Age Gap of 5 Neighboring Countries	0.81	-0.15						
	(0.06)***	(0.11)						
Spousal Age Gap Ratio of 5 Neighboring Countries					0.03	0.01		
					(0.01)***	(0.01)		
Plow-positive environment			-0.20	-0.11			-0.02	-0.01
			(0.30)	(0.24)			(0.02)	(0.01)
Plow-negative environment			2.72	1.48			0.15	0.08
			(0.49)***	(0.45)***			(0.03)***	(0.03)***
Fertility		0.14		0.09		0.01		0.01
		(0.08)*		(0.07)		(0.001)***		(0.01)***
Urbanization		-0.03		-0.01		-0.01		-0.01
		(0.01)***		(0.001)**		(0.01)*		(0.001)**
Share of Muslim population		8.03		0.81		0.36		0.05
		(1.48)***		(0.28)***		(0.07)***		(0.02)***
Sex ratio at birth		-3.97		1.87		-0.07		0.16
		(6.36)		(4.63)		(0.32)		0.26
Sex ratio under 5 mortality		4.44		4.37		0.25		0.25
		(3.00)		(2.86)		(0.16)		(0.16)
Sex ratio under 15 mortality		-6.33		-4.15		-0.35		-0.26
		(3.17)**		(2.93=		(0.17)**		(0.16)
Cum. pop. married at 40		2.68		-2.60		0.27		-0.06
		(1.85)		(1.77)		(0.10)***		(0.10)
French Legal Origin				0.25				0.01
				(0.16)				(0.01)
Female Labor Force Participation		-0.01		-0.01		0.01		-0.01
		(0.01)		(0.01)		(0.01)		(0.01)
Log GDP per Capita		0.39		-0.01		0.01		-0.01
		(0.17)		(0.12)		(0.01)		(0.01)
Continent dummies	no	yes	no	yes	no	yes	no	yes
Fixed effects	no	yes	no	no	no	yes	no	no
F-test	161.20	2.02	25.43	5.98	17.03	0.44	29.02	5.77
Sargan-Hansen statistic (p-value)			0.08	0.14			0.97	0.16
Angrist-Pischke F-statistics (p-value)	<0.001	0.14	<0.001	0.002	<0.001	0.49	<0.001	0.002
Anderson LM test (p-value)	<0.001	0.14	<0.001	0.002	<0.001	0.50	<0.001	0.003
Cragg-Donald Wald F statistic	161.20	2.02	25.43	5.98	17.03	0.44	29.02	5.77
Stock-Yogo critical values 10%	16.38	16.38	19.93	19.93	16.38	16.38	19.93	19.93
Stock-Yogo critical values 25%	5.53	5.53	7.25	7.25	5.53	5.53	7.25	7.25
R-squared	0.39	0.26	0.16	0.53	0.07	0.37	0.21	0.57

Notes: The dependent variable in column (1)-(4) is the Spousal Age Gap (*Male* minus *Female Marriage Age*). Columns (5)-(8) estimate the Spousal Age Gap Ratio (*Male* over *Female Marriage Age*). The regressors are: (i) The average of the Spousal Age Gap of five neighboring countries, respectively the (ii) the Spousal Age Gap Ratio of five neighboring countries; (iii)-(iv) plow-positive and plow-negative environment based on Alesina et al. (2013); (v) the total fertility rate; (vi) the level of urbanization in percent; (vii) the share of muslim population per country; (viii) the sex ratio at birth (males over females); (ix) the mortality rate of boys over the mortality rate of girls under 5, and (x) under 15 years, respectively; (xi) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (xii) a dummy whether the country's legal origin is based on French system; (xiii) the percentage of females in the national labor force; (xiv) the log of per capita GDP in PPP terms; (xv) six continent dummies. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 3.17: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age with alternative outcome variable

1980-2010				
Dependent variable =	Average Female Years of Education			
Panel A: Female Marriage Age (IV)	(1)	(2)	(3)	(4)
Number of countries	71	70	66	66
Observations	213	210	198	198
Female Marriage Age	0.69	0.30	0.26	0.34
	(0.06)***	(0.08)***	(0.09)***	(0.11)***
Fertility		-0.89	-0.88	-0.85
		(0.11)	(0.11)***	(0.15)***
Urbanization		0.03	0.02	0.02
		(0.01)***	(0.01)**	(0.02)
Share of Muslim population		-0.29	-0.05	4.81
		(0.62)	(0.69)	(2.13)**
Sex ratio at birth		-5.10	-6.89	-4.75
		(7.45)	(7.39)	(9.51)
Sex ratio under 5 mortality		-8.94	-7.22	-4.60
		(5.05)*	(4.83)	(5.74)
Sex ratio under 15 mortality		7.82	6.23	2.13
		(5.59)	(5.34)	(6.68)
Cum. pop. married at 40		-2.67	-3.03	0.16
		(3.15)	(3.23)	(4.01)
French Legal Origin		-1.65	-1.48	
		(0.38)***	(0.41)***	
Female Labor Force Participation			0.01	0.01
			(0.01)*	(0.01)
Log GDP per Capita			0.30	0.43
			(0.23)	(0.29)
Continent dummies	no	yes	yes	no
Country fixed effects	no	no	no	yes
R-squared	0.45	0.79	0.80	0.27

Panel B: Spousal Age Gap (IV)	(1)	(2)	(3)	(4)
Number of countries	71	70	66	66
Observations	213	210	198	198
Spousal Age Gap	-2.59	-1.19	-0.36	9.75
	(0.28)***	(0.37)***	(0.43)	(22.38)
Fertility		-0.80	-0.72	-2.84
		(0.13)***	(0.13)***	(4.08)
Urbanization		0.05	0.04	0.17
		(0.01)***	(0.01)***	(0.34)
Share of Muslim population		-1.02	-0.87	-61.80
		(0.57)*	(0.57)	(147.94)
Sex ratio at birth		-12.84	-15.34	41.30
		(7.79)*	(8.26)*	(129.85)
Sex ratio under 5 mortality		-15.58	-10.34	-6.65
		(5.21)***	(5.76)*	(37.25)
Sex ratio under 15 mortality		14.87	10.03	22.42
		(5.26)***	(5.74)*	(35.83)
Cum. pop. married at 40		-7.32	-6.56	-41.82
		(4.17)*	(4.24)	(79.82)
French Legal Origin		-1.71	-1.52	
		(0.30)***	(0.30)***	
Female Labor Force Participation			0.01	0.01
			(0.01)	(0.07)
Log GDP per Capita			0.38	-0.41
			(0.23)*	(3.41)
Continent dummies	no	yes	yes	no
Country fixed effects	no	no	no	yes
R-squared	0.40	0.82	0.81	0.18

Appendix Table 3.17 continued: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age with alternative outcome variable

1980-2010				
Dependent variable =	Average Years of Education Ratio			
Panel C: Spousal Age Gap Ratio	(1)	(2)	(3)	(4)
Number of countries	71	70	66	66
Observations	213	210	198	198
Spousal Age Gap	0.26	0.16	0.01	-1.66
	(0.03)***	(0.06)***	(0.06)	(3.88)
Fertility		0.16	0.18	0.51
		(0.03)***	(0.03)***	(0.71)
Urbanization		-0.01	-0.01	-0.02
		(0.001)***	(0.01)	(0.06)
Share of Muslim population		0.24	0.21	9.96
		(0.10)**	(0.15)	(25.68)
Sex ratio at birth		5.30	2.30	-16.56
		(1.84)***	(2.13)	(22.68)
Sex ratio under 5 mortality		0.58	0.01	-0.37
		(1.32)	(0.148)	(6.47)
Sex ratio under 15 mortality		-0.16	0.44	-2.29
		(1.29)	(1.48)	(6.22)
Cum. pop. married at 40		-2.19	-2.16	5.05
		(0.89)***	(1.09)**	(13.86)
French Legal Origin		0.09	0.16	
		(0.07)	(0.08)**	
Female Labor Force Participation			-0.01	-0.01
			(0.01)	(0.01)
Log GDP per Capita			-0.13	0.38
			(0.06)**	(0.59)
Continent dummies	no	yes	yes	no
Country fixed effects	no	no	no	yes
R-squared	0.29	0.56	0.58	0.05

Notes: The dependent variable is the average years of female schooling (absolute in Panels A and B, and relative to male average years of schooling in Panel C). The regressors are: (i) The *Female Marriage Age* (SMAM) in Panel A, and the Spousal Age Gap (*Male* minus *Female Marriage Age*) in Panels B (absolute) and C (relative); (ii) the total fertility rate; (iii) the level of urbanization in percent; (iv) the share of muslim population per country; (v)-(vii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (viii) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (ix) dummy whether the country's legal origin is based on French system; (x) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xii) six continent dummies. French Legal Origin omitted in column (4) because of inclusion of fixed effects. All panels use as instrument the weighted average of the absolute, respectively relative Female Marriage Age of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 3.18: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age with OECD dummy

1980-2010						
Dependent variable =	Female Secondary Schooling Completion			Average Female Years Tertiary Education		
Panel A: Female Marriage Age (IV)	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	71	70	66	71	70	66
Observations	213	210	198	213	210	198
Female Marriage Age	3.09	2.22	2.21	0.07	0.05	0.04
	(0.36)***	(0.71)***	(0.83)***	(0.01)***	(0.01)***	(0.01)***
OECD Country Dummy	1.95	2.15	1.18	0.06	0.18	0.17
	(2.57)	(4.08)	(4.48)	(0.05)	(0.07)***	(0.07)**
Fertility		-3.91	-4.01		-0.01	0.01
		(0.91)***	(1.00)***		(0.02)	(0.02)
Urbanization		0.01	-0.05		0.01	0.01
		(0.08)	(0.09)		(0.001)*	(0.01)
Share of Muslim population		4.45	4.71		-0.03	0.07
		(3.96)	(4.69)		(0.07)	(0.08)
Sex ratio at birth		-1.96	-25.82		-0.03	0.01
		(59.21)	(62.74)		(1.01)	(1.01)
Sex ratio under 5 mortality		34.41	51.92		-1.22	-0.94
		(43.99)	(45.01)		(0.75)*	(0.72)
Sex ratio under 15 mortality		-39.39	-56.02		1.06	0.78
		(48.27)	(49.37)		(0.82)	(0.79)
Cum. pop. married at 40		-45.38	-40.82		-1.51	-1.46
		(26.19)*	(29.29)		(0.45)***	(0.47)***
French Legal Origin		-6.31	-6.04		-0.01	0.01
		(2.33)***	(2.58)**		(0.04)	(0.04)
Female Labor Force Participation			0.05			0.01
			(0.06)			(0.001)***
Log GDP per Capita			1.37			0.04
			(1.99)			(0.03)
Continent dummies	no	yes	yes	no	yes	yes
Country fixed effects	no	no	no	no	no	no
R-squared	0.34	0.51	0.50	0.37	0.60	0.63

Appendix Table 3.18 continued: Panel for level of female education, absolute and relative (Spousal Age Gap) Female Marriage Age with OECD dummy

1980-2010						
Dependent variable =	Female Secondary Schooling Completion			Average Female Years Tertiary Education		
Panel B: Spousal Age Gap (IV)	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	71	70	66	71	70	66
Observations	213	210	198	213	210	198
Spousal Age Gap	-8.49	-0.71	-0.40	-0.17	-0.13	-0.14
	(1.83)***	(2.61)	(3.05)	(0.04)***	(0.05)***	(0.06)**
OECD Country Dummy	1.84	-0.45	-0.51	0.07	0.17	0.17
	(2.96)	(2.77)	(2.82)	(0.06)	(0.06)***	(0.06)***
Fertility		-3.80	-3.60		0.01	0.02
		(0.93)***	(0.93)***		(0.02)	(0.02)
Urbanization		0.15	0.11		0.01	0.01
		(0.08)**	(0.09)		(0.01)	(0.01)
Share of Muslim population		2.34	3.88		0.10	0.15
		(3.93)	(4.06)		(0.08)	(0.08)
Sex ratio at birth		-26.14	-36.57		0.20	0.08
		(52.27)	(56.43)		(1.07)	(1.16)
Sex ratio under 5 mortality		-46.65	-18.64		-1.66	-1.08
		(35.09)	(39.51)		(0.72)**	(0.81)
Sex ratio under 15 mortality		51.55	23.12		1.82	1.27
		(35.59)	(39.55)		(0.73)***	(0.81)
Cum. pop. married at 40		-71.51	-67.20		-3.00	-2.94
		(28.06)***	(29.04)**		(0.57)***	(0.60)***
French Legal Origin		-6.43	-6.46		0.03	0.04
		(2.04)***	(2.06)***		(0.04)	(0.04)
Female Labor Force Participation			0.09			0.02
			(0.06)			(0.01)**
Log GDP per Capita			1.55			0.03
			(1.57)			(0.03)
Continent dummies	no	yes	yes	no	yes	yes
Country fixed effects	no	no	no	no	no	no
R-squared	0.25	0.56	0.54	0.29	0.53	0.52

1980-2010						
Dependent variable =	Secondary Schooling Completion Ratio			Average Years Tertiary Education Ratio		
Panel C: Spousal Age Gap Ratio	(1)	(2)	(3)	(4)	(5)	(6)
Number of countries	71	70	66	71	70	66
Observations	213	210	198	213	210	198
Spousal Age Gap	0.41	0.21	0.18	0.52	0.26	0.77
	(0.07)***	(0.16)	(0.14)	(0.07)***	(0.26)	(0.40)**
OECD Country Dummy	0.09	0.17	0.31	0.25	1.06	0.44
	(0.17)	(0.24)	(0.13)**	(0.14)*	(0.24)***	(0.35)
Fertility		0.24	0.21		0.26	0.24
		(0.09)***	80.04)***		(0.06)***	(0.12)**
Urbanization		-0.01	0.01		-0.01	0.01
		(0.01)	(0.01)		(0.01)	(0.01)
Share of Muslim population		-0.09	0.34		-0.10	-0.85
		(0.27)	(0.19)*		(0.38)	80.53)
Sex ratio at birth		8.45	5.14		0.68	2.78
		(4.84)*	(2.60)**		(4.10)	(6.81)
Sex ratio under 5 mortality		-2.60	-1.53		3.49	-0.66
		(3.46)	(1.82)		(2.32)	(4.75)
Sex ratio under 15 mortality		3.26	1.83		-4.24	-0.36
		(3.54)	(1.82)		(2.37)*	(4.76)
Cum. pop. married at 40		-0.71	0.67		2.26	0.99
		(2.38)	(1.34)		(1.53)	(3.47)
French Legal Origin		0.22	-0.01		-0.27	-0.04
		(0.18)	(0.10)		(0.23)	(0.27)
Female Labor Force Participation			0.01			0.01
			(0.001)**			(0.01)
Log GDP per Capita			-0.17			-0.06
			(0.07)**			(0.20)
Continent dummies	no	yes	yes	no	yes	yes
Country fixed effects	no	no	no	no	no	no
R-squared	0.19	0.28	0.57	0.27	0.43	0.33

Notes: The dependent variable in column (1)-(3) is the percentage of the female population with a completed secondary education. Columns (4)-(6) estimate the average years of female tertiary schooling. The regressors are: (i) The *Female Marriage Age* (SMAM) in Panel A, and the Spousal Age Gap (*Male* minus *Female Marriage Age*) in Panels B (absolute) and C (relative); (ii) a Dummy for a country being member of the OECD; (iii) the total fertility rate; (iv) the level of urbanization in percent; (v) the share of muslim population per country; (vi)-(viii) the sex ratio at birth (males over females), and the mortality rate of boys over the mortality rate of girls under 5, and under 15 years, respectively; (ix) the cumulative percentage of married men at the age of 40 (out of total male population) divided by the cumulative percentage of married women at the age of 40 (out of total female population); (x) dummy whether the country's legal origin is based on French system; (xi) the percentage of females in the national labor force; (xi) the log of per capita GDP in PPP terms; (xiii) six continent dummies. Specification with country fixed effects not included since this would eliminate the country dummy of interest. All panels use as instrument the weighted average of the absolute, respectively relative Female Marriage Age of adjoining countries with a common border, where weights are according to relative length of shared border. See the Appendix for more detailed variable definitions and sources. Robust Standard Errors are reported in parentheses. *** ,** and * denote statistical significance at the 1, 5 and 10% level, respectively.

Appendix Table 3.19: Data and Sources

Variable Name	Description	Source	Data restrictions and remarks
Country	Name of country	Feenstra, Robert C., Inklaar, R. & Timmer, M. P. (2015). The Next Generation of the Penn World Table. American Economic Review, forthcoming, available for download at www.ggdc.net/pwt	
CCode	ISO 3166-1 alpha-3 country codes	Defined in ISO 3166-1, part of the ISO 3166 standard published by the International Organization for Standardization (ISO)	
Male Marriage Age (SMAM)	Singulate mean age at marriage (SMAM) for men, measured as the average length of single life expressed in years among those who marry before age 50.		
Female Marriage Age (SMAM)	Singulate mean age at marriage (SMAM) for women, measured as the average length of single life expressed in years among those who marry before age 50.	United Nations, Department of Economic and Social Affairs, Population Division (2013). World Marriage Data 2012 (POP/DB/Marr/Rev2012).	for each decade (1980s, 1990s, 2000s), simple averages are taken from all available values per country on male and female SMAM
First Difference Female Marriage Age	smam_female in the 2000s (average 2000-2009) minus smam_female in the 1980s (average 1980-1989)		
Spousal Age Gap (SAG)	Male Marriage Age minus Female Marriage Age		
Spousal Age Gap Ratio	Male Marriage Age over Female Marriage Age		
plow_negative_environment	The average fraction of ancestral land that was suitable for growing barley, rye, and wheat divided by the fraction that was suitable for any crops.	Alesina, A., Giuliano, P., & Nunn, N. (2013). On the Origin of Gender Roles: Women and the Plough. *Quarterly Journal of Economics*, *128* (2): 469-530. Data taken from http://scholar.harvard.edu/nunn/pages/data-0	
plow_positive_environment	The average fraction of ancestral land that was suitable for growing foxtail millet, pearl millet, and sorghum divided by the fraction that was suitable for any crops.		
LOG GDP per Capita	The log of GDP per capita per time period (Output-side real GDP at current PPPs)	Feenstra, Robert C., Inklaar, R. & Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review*, forthcoming, available for download at www.ggdc.net/pwt	for each decade (1980s, 1990s, 2000s), simple averages are taken from all available values on income per country
Secondary Schooling Completion (Diff)	Secondary Schooling Completion rate of males minus the rate of females		
Secondary Schooling Completion (Ratio)	Secondary Schooling Completion rate of males over the rate of females		
Average Years Tertiary Education (Diff)	Average years of tertiary education of males minus the years of females		
Average Years Tertiary Education (Ratio)	Average years of tertiary education of males over the years of females	Own construction based on dataset by Barro and Lee (Barro, R., & Lee, J.-L. (2010). A New Data Set of Educational Attainment in the World, 1950-2010. *Journal of Development Economics*, *104* , 184-198.) www.barrolee.com	for each decade (1980s, 1990s, 2000s), simple averages are taken from all available values for the respective education variable per country
Female Secondary Schooling Completion	Secondary Schooling Completion rate of females		
First Difference Female Secondary Schooling Completion	Female secondary schooling in the 2000s (average 2000-2009) minus female secondary schooling in the 1980s (average 1980-1989)		
Female Average Years Tertiary Education	Average years of tertiary education of females		
Primary Schooling Completion (Ratio)	Primary Schooling Completion rate of males over the rate of females		
Gender Parity Index (GPI) for Adult Literacy	Ratio of female to male aduty literacy rates	UNESCO Institute for Statistics (UIS) Data Centre (http://data.uis.unesco.org).	
Fertility	Total fertility (children by women)	United Nations, Department of Economic and Social Affairs, Population Division (2015). World Population Prospects: The 2015 Revision, DVD Edition.	
Female Labor Force Participation	Female labour force as a percent of the female working age population.	International Labour Organization (ILO); data online available.	for each decade (1980s, 1990s, 2000s), simple averages are taken from all available values
Urbanization	Urban population (as % of total population)	Own construction based on "Rural population" indicator from UNESCO Institute for Statistics (UIS) Data Centre (http://data.uis.unesco.org).	
Muslim	Muslim population (as % of total population)	Religion adherence data from Barro, R. (2003). http://scholar.harvard.edu/barro/publications/religion-adherence-data	No data for 1980s and 1990s available. Took religion adherence from the 1970 as proxy for 1980s, and adherence from the 2000 as proxy for 1990s
French Legal Origin	Dummy for country with French legal origin / tradition	Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. Journal of economic growth, 9(2), 131-165.	
Cum. pop. married at 40	Cumulative percentage of married men at the age of 40 (out of total male population) divided by cumulative percentage of married women at the age of 40 (out of total female population)	Own construction based on United Nations, Department of Economic and Social Affairs, Population Division (2013). World Marriage Data 2012 (POP/DB/Marr/Rev2012).	For each decade (1980s, 1990s, 2000s), simple averages are taken from all available values per country
Asia			
Europe			
Africa			
North America	Dummy variable taking value 1 if country is located on given continent, 0 otherwise	Own construction based on World Bank definition of World regions	
South America			
Oceania			
Mother Age at First Child	Median age at first birth among women	United Nations, Department of Economic and Social Affairs, Population Division. World Fertility Report 2012	
Women's economic rights (wecon)	Women's Economic Rights index ranging from 0 (no economic rights) to 3 (all or nearly all of women's economic rights were guaranteed by law and the government fully and vigorously enforces these laws in practice)		
Women's political rights (wopol)	Women's Political Rights index ranging from 0 (no political rights) to 3 (political rights are guaranteed in both law and practice)	Cingranelli, D. L., Richards, D. L., & Clay, K. C. (2014). The CIRI Human Rights Dataset. http://www.humanrightsdata.com. Version 2014.04.14.	
Women's social rights (wosoc)	Women's Social Rights index ranging from 0 (no social rights for women) to 3 (all or nearly all of women's social rights were guaranteed by law and the government fully and vigorously enforced these laws in practice)		

Appendix Table 3.19 continued: Data and Sources

Variable Name	Description	Source	
Sex ratio at birth	Sex ratio at birth by decade and country (males over females)		
Sex ratio under 5 mortality	Deaths of boys with age 0-4 (as % of all boys in that age) divided by the deaths of girls with age 0-4 (as % of all girls in that age)	United Nations, Department of Economic and Social Affairs, Population Division (2013). World Population Prospects: The 2012 Revision, DVD Edition.	
Sex ratio under 15 mortality	Deaths of boys with age 0-14 (as % of all boys in that age) divided by the deaths of girls with age 0-14 (as % of all girls in that age)		For each decade (1980s, 1990s, 2000s), simple averages are taken from all available values on the respective variable per country
Female SMAM / SAG Border Neighbors	Weighted average of female SMAM / SAG of adjoining countries with a common border, where weights are according to relative length of shared border	Own construction based on United Nations, Department of Economic and Social Affairs, Population Division (2013). World Marriage Data 2012 (POP/DB/Marr/Rev2012), and for the weighting by border length from CEPII, following Correa, E. A, Jetter, M., & Agudelo, A. M. (2016). Corruption: Transcending borders, Kyklos, 69(2), 183-207.	
Female SMAM / SAG 5 neighboring countries	Simple average of the female SMAM / SAG of five neighboring countries	Own construction based on United Nations, Department of Economic and Social Affairs, Population Division (2013). World Marriage Data 2012 (POP/DB/Marr/Rev2012), and own selection of 5 neighboring countries	
Gender Inequality Index (GII)	Index for measurement of gender disparity along three dimensions (reproductive health, gender empowerment, economic status) scaled from 0 (total equality) to 1 (total inequality)	United Nations Development Programme (UNDP) 2013, for download at http://hdr.undp.org/en/content/table-4-gender-inequality-index	Only available since 2010, which we take as proxy for the 2000s time interval
Women Treated With Respect	Percentage of female respondents answering yes to the question, "Do you believe that women in this country are treated with respect and dignity, or not?"	Data originally from Gallup World Poll; data here taken from respective category in Social Progress Index 2014 (www.socialprogressimperative.org)	Took 2014 values as proxy for the 2000s time interval
Minimum Legal Marriage Age for Women	Minimum legal age for marriage without parental consent as defined in terms of the laws of the individual country	United Nations Statistics Division (DYB 2011: http://unstats.un.org/unsd/demographic/products/dyb/dyb2011/notes/notes24.pdf)	
Spousal Age Gap Ratio Parental Generation	Male Marriage Age over Female Marriage Age per country (average 1980-1989)	Own construction based on United Nations, Department of Economic and Social Affairs, Population Division (2013). World Marriage Data 2012 (POP/DB/Marr/Rev2012)	Due to lags only available for 2000s time period
Average Female Years of Education	Average Years of Schooling Attained	Barro, R., & Lee, J.-L. (2010). A New Data Set of Educational Attainment in the World, 1950-2010. Journal of Development Economics, 104, 184-198. www.barrolee.com	
Female Primary Schooling	Percentage of female population with a completed primary education		
Quality of Schooling male relative to female	Gender difference (male over female) in student performance on the science scale for the PISA test 2006	OECD 2007. PISA 2006: Science Competencies for Tomorrow's World. http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/pisa2006results.htm	Took 2006 PISA test as proxy for the 2000s time period
Adolescent births	Number of births to women with age 15-19 per 1,000 women with age 15-19	Taken as sub-dimension from the Gender Inequality Index (GII); United Nations Development Programme (UNDP) 2013, for download at http://hdr.undp.org/en/content/table-4-gender-inequality-index	Took 2010 values as proxy for the 2000s time interval
Women in Politics	Proportion of seats in parliament held by women, measured in 2000	United Nations' Women's Indicators and Statistics Database.	

Appendix Table 3.20: Overview of countries per time period with data on marriage age

Panel 1980-2010 (tables 3.2-3.5)	Cross-section 2000s (table 3.6)
86	*135*
Argentina	Afghanistan
Australia	Albania
Austria	Algeria
Bahrain	Antigua and Barbuda
Bangladesh	Argentina
Belgium	Armenia
Belize	Aruba
Botswana	Azerbaijan
Brazil	Bahrain
Brunei Darussalam	Bangladesh
Cameroon	Belarus
Canada	Benin
Chile	Bhutan
China	Bolivia (Plurinational State of)
Colombia	Bosnia and Hercegovina
Czech Republic	Botswana
Denmark	Brazil
Dominican Republic	Brunei Darussalam
Ecuador	Bulgaria
Egypt	Burkina Faso
Finland	Burundi
France	Cambodia
Greece	Cameroon
Guyana	Cape Verde
Haiti	Chad
Hungary	Chile
Iceland	China
India	Colombia
Indonesia	Congo
Iran (Islamic Republic of)	Costa Rica
Iraq	Croatia
Ireland	Cuba
Israel	Cyprus
Italy	Democratic People's Republic of of Korea
Jamaica	Democratic Republic of the Congo
Japan	Dominican Republic
Kazakhstan	Ecuador
Kenya	Egypt
Kuwait	El Salvador
Kyrgyzstan	Estonia
Luxembourg	Ethiopia
Malaysia	Gabon
Maldives	Georgia
Mali	Ghana
Malta	Greece
Mauritania	Guinea
Mauritius	Guyana
Mexico	Haiti
Morocco	Honduras
Mozambique	Hungary
Myanmar	India
Nepal	Indonesia
Netherlands	Iran (Islamic Republic of)
New Zealand	Iraq
Niger	Israel
Norway	Italy
	Jamaica
	Jordan
	Kazakhstan
	Kenya
	Kuwait
	Kyrgyzstan
	Lao People's Democratic Republic
	Latvia
	Lebanon
	Lesotho
	Liberia
	Libya
	Lithuania

Appendix Table 3.20 continued: Overview of countries per time period with data on marriage age

Panel 1980-2010 (tables 3.2-3.5)	Cross-section 2000s (table 3.6)
Pakistan	Madagascar
Panama	Malawi
Paraguay	Malaysia
Peru	Maldives
Philippines	Mali
Poland	Malta
Portugal	Mauritania
Qatar	Mauritius
Republic of Korea	Mexico
Saudi Arabia	Mongolia
Singapore	Morocco
Slovakia	Mozambique
South Africa	Myanmar
Spain	Namibia
Sudan	Nepal
Sweden	New Caledonia
Switzerland	Nicaragua
Thailand	Niger
Tonga	Nigeria
Trinidad and Tobago	Oman
Tunisia	Pakistan
Turkey	Palau
United Arab Emirates	Panama
United Kingdom	Papua New Guinea
United Republic of Tanzania	Paraguay
United States of America	Peru
Venezuela (Bolivarian Republic of)	Philippines
Vietnam	Poland
Zambia	Portugal
Zimbabwe	Puerto Rico
	Qatar
	Republic of Moldova
	Romania
	Russian Federation
	Rwanda
	Samoa
	Sao Tome and Principe
	Saudi Arabia
	Senegal
	Seychelles
	Sierra Leone
	Singapore
	Slovenia
	Solomon Islands
	South Africa
	Spain
	Sri Lanka
	Sudan
	Suriname
	Swaziland
	Syrian Arab Republic
	Tajikistan
	Thailand
	Tonga
	Trinidad and Tobago
	Tunisia
	Turkey
	Uganda
	Ukraine
	United Arab Emirates
	United Republic of Tanzania
	Venezuela (Bolivarian Republic of)
	Vietnam
	Yemen
	Zambia
	Zimbabwe

IV. Does Central Europe Import the Missing Women Phenomenon?*

ABSTRACT

This chapter examines whether immigrants have brought the missing women (at birth) phenomenon to Germany and Switzerland. Using a range of micro data since 1990, we find no systematic gender selection of foreigners collectively, but a group of Balkan, Chinese and Indian immigrants tend to display substantially elevated sex ratios at birth. Employing different estimation methods we consistently calculate around 1,500 missing girls in Germany (2003-2014) and Switzerland (1990-2014) combined from these selected Balkan and Asian immigrant groups. A Germany-specific measure of cultural adaptation has no substantial effect on the level of sex selection observed, and Swiss-specific data indicate a skewed ratio for Asian higher parity births. With household survey data we attempt to identify underlying reasons for sex selection in Germany, but find no robust associations for any socio-economic variable employed. However, the sex of older siblings tends to matter, and again Balkan, Chinese and Indian immigrants increase the boy-birth likelihood whereas immigrants collectively do not.

IV. 1. INTRODUCTION

Migration is a pressing global development issue and the United Nations (2016) estimate a current 244 million international migrants worldwide, of which 76 million are hosted by European countries. Economic prospects and labor market opportunities have recently turned Central Europe into the continent's gravity center for immigration. Migration flows to Germany and Switzerland have been of sizeable dimension in the last decades, in particular from countries located in Southeast Europe.[61]

---

* A working paper version of this chapter is circulating (Stimpfle & Stadelmann, 2016b) and has been presented at the Graduate Research Seminar of the University of Bayreuth. We cordially thank Marco Portmann, Mario Larch, and Hartmut Egger for helpful comments.

[61] In the 2014 German micro census, the individually listed groups from Bosnia-Herzegovina, Serbia, and former Yugoslavia alone amount to more than two million people, or twelve percent of all migrants in Germany (Federal Statistical Office, 2015a). Migrants from that region have also begun to pursue alternative migration channels: about 30 percent of all asylum applications in Germany in 2015 were submitted from Balkan people (BAMF, 2016). Likewise in Switzerland, the 2014 national statistics indicate that citizens of Serbia, Montenegro, Kosovo, Bosnia-Herzegovina, and Macedonia account for 14 percent of all foreigners in the country, totaling around 275,000 people BFS (2015b).

The socio-economic consequences of these migration waves have influenced the public discourse: changes of the gender composition in European countries due to migration received media attention (Ritter, 2016), and migrants may carry a range of traditional practices that differ from their new environment.[62] Such practices include different gender attitudes and potentially a relative preference of sons over daughters. Sen (1990) famously dubbed this form of female discrimination the "missing women" phenomenon. Although it is typically rather associated with India and China (Croll, 2001; World Bank, 2011), there is evidence that even within Europe gender ratios at birth vary among countries (Anderson & Ray, 2010; Guilmoto, 2015; Instat, 2014). Previous literature has focused on Asian immigrants and suggests continued gender selection practices (Abrevaya, 2009; Dubuc & Coleman, 2007). We are not aware, however, of any work that has examined to what extent Southeast European immigrants have affected sex selection at birth in Central European countries.[63] As migrants in Germany and Switzerland tend to come from the Balkans rather than from Asia, this chapter aims to fill the research gap in the economics of migration and gender. We search for evidence whether the missing women (at birth) issue was "imported" by migrants from Southeast Europe to Central Europe.[64]

The map in figure 4.1 reflects our research motivation. It gives an overview of the average sex ratios at birth (2003-2013) of all European countries from Germany and Switzerland to the East,[65] as well as India and China. Countries are highlighted in red whose sex ratio has been at least 1.07, i.e. for 100 girl births there are on average 107 boy births or more. We observe a geographical concentration in Southeast Europe that deviates from what is considered a biologically normal sex ratio at birth of 1.04-1.07 (Chahnazarian, 1988). Albania, for example, reports an average ratio of 1.11 which is as high as the widely discussed national Indian figures. This significant regional outlier motivates our focus on the (Western) Balkans, and this work is the first to provide a deeper explorative analysis of migrant groups from that region to Germany and Switzerland.

[62] Kountouris and Remoundou (2016), for example, document how cultural background is a significant determinant of migrants' individual environmental attitudes.
[63] We proxy Central Europe with the two countries Germany and Switzerland since other countries that may be classified as part of Central Europe have a much lower number of migrants (even less so from Asia and from the (West) Balkans), and thus will unlikely materially affect results for the overall region.
[64] Specifically we focus on excess female mortality at birth and refrain from examining unbalanced sex ratios overall, acknowledging that the "missing women" phenomenon comprises additional aspects. We revisit the precise scope of our research in further detail later in this chapter.
[65] Data for all European countries except Kosovo, are from the World Health Organization Europe (2015); for Kosovo, from the Kosovo Agency of Statistics (2014); for India and China, from United Nations Population Division (2015).

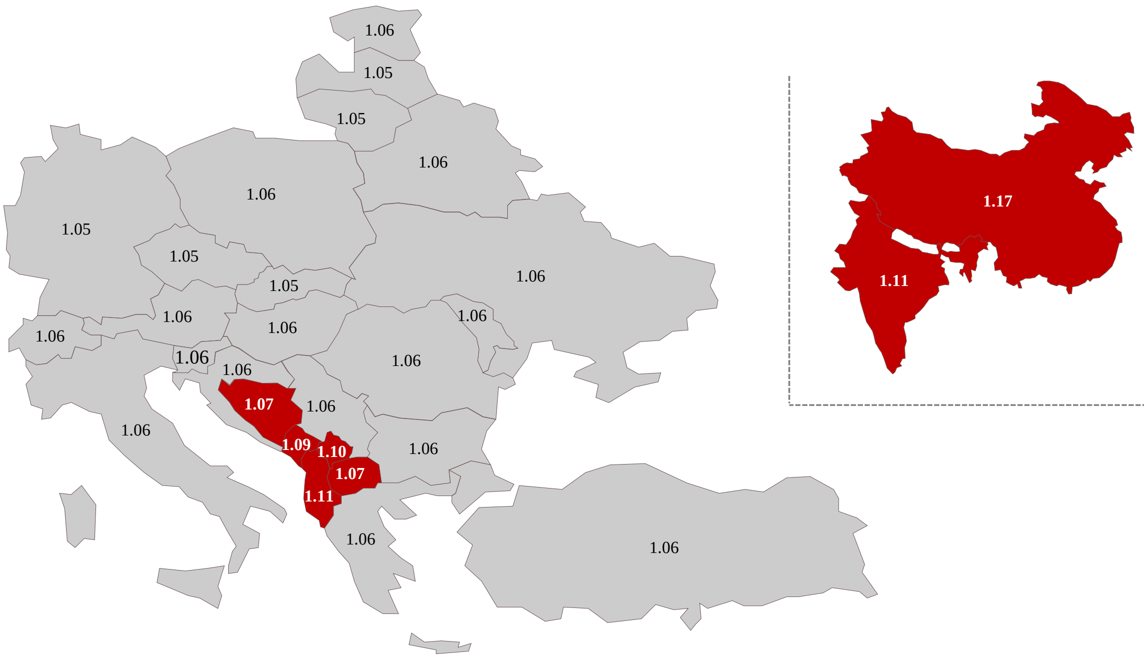


Figure 4.1: Average Sex Ratio at Birth (boys over girls) from 2003-2013 in Central and Eastern Europe, India, and China

To examine this complex topic of migration and gender discrimination, we take a "forensic" approach, using different empirical techniques and three distinct micro data sources. We initially resort to abortion statistics with the aim to discover sex-selective behavior among parents. These data offer a first indication on the extent of abortion practices, and their dynamics with respect to timing and child parity. We then analyze more than a decade of complete national birth registries along maternal citizenship and sex of the born child. Comparing Germany and Switzerland, which are otherwise similar both in terms of culture and socio-economic position, serves to investigate the consistency of our findings. In addition, we are able to infer from the German data a proxy for the parental level of cultural adaptation. Swiss birth records in return document birth parity, which also allows for a more granular perspective on prenatal sex selection practices.

Lastly, we employ the socio economic panel study (SOEP), a representative annual survey of households in Germany, to explore potential determinants for different sex ratios at birth. Thereby, we try to obtain a more refined picture of individual household characteristics.

Results from abortion statistics over recent years in Germany and Switzerland offer no conclusive evidence and only allow for speculative conjectures about existing sex selection. From the birth registries we find that foreigners collectively show a slightly elevated sex ratio at birth compared to natives. The ratios, however, are still within biologically normal ranges and provide no evidence of a systematic issue of “missing women” at birth among migrants. Yet, sub-groups of Balkan, Chinese, and Indian immigrants display substantially elevated sex ratios. Based on these findings we conclude that gender selection is likely present among these sub-groups, which we document nonetheless to be small in absolute numbers.

Further results indicate that sex ratios at birth do not vary much if the father is German or from the same country as the mother, and also the number of years the parents spent in Germany has no substantial differentiating effect. The birth data for Switzerland reveal a significantly skewed ratio for Asian higher parity births. Independent of the estimation method chosen, we then consistently calculate around 1,500 missing girls (missing women at birth) in Germany (2003-2014) and Switzerland (1990-2014) combined from the highlighted Balkan and Asian immigrant groups. For foreigners collectively, in return, no robust evidence for missing girls is found.

Finally, we exploit the German SOEP and regress the proportion of male children as well as the boy-birth likelihood per household on demographic and socio-economic variables. Results again suggest that a higher likelihood for boys is associated only with selected (Balkan and Asian) migrant groups, i.e. not migrants collectively, which is consistent with the findings from birth registries. Furthermore, existing female siblings play an important role, raising the odds for a boy birth significantly if the first child was a girl. Lastly, religiosity seems to be the only socio-economic variable that matters, but effects depend on the population group examined.

The chapter is structured as follows. Section 2 reviews the related literature and points at the research gaps regarding Central Europe and migrants from Balkan countries. Section 3 presents our three data sources and section 4 discusses the empirical results. Section 5 offers concluding remarks.

IV. 2. LITERATURE REVIEW

This chapter contributes to three strands of the “missing women” literature. First it adds to the research quantifying excess female mortality rates. It secondly aims to improve the understanding of the mechanisms for effectively implementing gender selection at birth. Finally, it seeks to explain the genuine “underlying motives for gender selection” (Abrevaya, 2009, p.29). We will review the literature related to each strand in the following.

Quantifying Excess Female Mortality Rates

A number of estimates have been made on how many women are globally missing due to this form of gender inequality. Apart from the challenge of obtaining valid population statistics from all countries, the academic discussion mostly revolves around the correct counterfactual, i.e. an "unbiased" biological sex ratio. After the famous "100 million" figure by Sen (1990), Coale (1991) proposes that in fact rather 60 million females are missing. Klasen (1994) and Klasen and Wink (2003) provide further methodological refinements, where the latter study estimates again more than 100 million missing women. Bongaarts and Guilmoto (2015) provide a historical perspective and calculate that this figure grew from 61 million in 1970 to 126 million in 2010 and will further rise to 150 million in 2035 before declining to 142 million by 2050. Yet another contribution by Anderson and Ray (2010) calculates the flow of annually missing women and offers a more granular view on mortality rates by age groups.

Studies have examined this phenomenon in various corners of the world. Findings by UNFPA (2012b), and by Basu and Das Gupta (2001) provide a global perspective, while the majority of studies examine the two most prominent country cases, namely China and India (among others, see Agnihotri, 1996; Das Gupta, Chung, & Shuzhuo, 2009; Das Gupta & Shuzhuo, 1999; Mayer, 1999; Shepherd, 2008; Sudha & Rajan, 1999). Recent studies expanded the geographical focus. Dahl and Moretti (2008) analyze the U.S., China, Mexico, Colombia, Kenya and Vietnam. Meslé, Vallin, and Badurashvili (2007) document biased sex ratios in the Caucasus, and Yount (2001) examines excess female mortality in the Middle East.

Some papers have specifically looked at Southeast Europe, which also displays excess female mortality rates as documented in figure 4.1. Studies on general demographics and gender inequality in that region (Gjonca, 2004; Gjonca, Aasve, & Mencarini, 2008, 2009; Hall, 1994) have been complemented by focused research on the sex ratio at birth. Polasek, Kolcic, Kolaric, and Rudan (2005) investigate the sex ratio at birth in Croatia before, during and after the Balkan war of the 1990s, but find no significant deviations in any period. Guilmoto (2010) points at the biased sex ratios at birth in a "regional block" (Albania, Kosovo, Macedonia, Montenegro), and calls out the substantial knowledge gaps around that issue.[66] A recent report by the United Nations

[66] Also, the issue of "missing women" is often unnoticed or not yet acknowledged in the region itself, as described by Guilmoto and Duthé (2013). For instance, Stump (2011), a member of the European Council, notes that the Albanian authorities do not consider the skewed birth ratio as a nation-wide problem, but a sporadic phenomenon limited to some remote areas.

(UNFPA, 2012a) on Albania takes up this challenge and provides rich statistical evidence of excess prenatal female mortality, as well its likely underlying causes. Anecdotal evidence by Durham (1909) describes how the strong symbolic need for a son led women in Albania (so-called "sworn virgins") even to renounce marriage and acquire a masculine identity in order to assume the role of a virtual son in families with no male offspring. The dramatic bias of the sex ratio at birth most likely began only in the 1990s, since beforehand no contraception was widely available and abortion was prohibited (see also David, 1970). Nonetheless, the reliability of official figures is discussed in the UN report (UNFPA, 2012a), as the sex ratio at birth from the Albanian census differs substantially from the ratio calculated via birth registries (though both are clearly male biased).

Figures by the United Nations Population Fund (UNFPA, 2012a) as well as national statistical bureaus (for example Instat, 2012) also emphasize the enormous migration waves from Balkan countries to Western and Central Europe: in Albania alone, 24 percent of the population emigrated during the 1990s. It has been pointed out that migration flows may substantially affect the level of gender preference and the sex ratio at birth in both the original and the receiving country (Attané & Guilmoto, 2007; Dyson, 2012). Recent work therefore analyzed sex ratios at birth among migrants from societies that are known for son preference. The initial focus was put on people of Asian origin. Among the first authors, Dubuc and Coleman (2007) document an increase in the sex ratio of births among India-born mothers in England and Wales, especially for higher-order births. The authors hold prenatal sex diagnosis of fetuses and subsequent abortion of female fetuses as main responsible factors for this trend. The authors hence question the adherence of certain India-born immigrants to Western norms of gender equality.

Abrevaya (2009) analyzes immigrant groups in the United States. Using three different sources (Federal birth data, California birth data, and census data) and controlling for a set of observable maternal characteristics, he concludes that over 2,000 Chinese and Indian girls are "missing" in the country between 1991 and 2004, especially due to gender selection at later births. He also sees strong parallels between birth dynamics of Chinese and Indian migrants in the U.S., and the birth statistics found in China and India itself. Almond and Edlund (2008) summarize comparable results for the U.S. Almond, Edlund, and Milligan (2013) confirm a similar outcome for South and East Asian immigrants in Canada. The three papers cited on North America all report substantially elevated sex ratios at birth at higher parities if the previous children were girls. In addition, there is first evidence for a comparable situation among Chinese and Indian migrant families in Australia (SBS Radio, 2015).

We received a number of papers conducting a similar analysis for continental European countries. Verropoulou and Tsimbos (2010) conclude from Greek birth records that Albanian mothers have a five percent higher chance of having a male birth than Greek mothers, and Asian mothers also have a significantly higher sex ratio at birth than native Greeks. The Albanian sex ratio is on average at 1.09 and the ratio of Asians even amounts to 1.29, which compares to 1.05 for natives. However, their sample for analysis comprises only a bit more than 100,000 births in total (out of which 12,000 are Albanian and only 1,200 Asian), which might cause undesired random variation linked to the limited sample size. Gavalas, Rontos, and Nagopoulos (2015) conduct a second study on Greece with a larger sample, and report for Indian and Chinese mothers on average a highly elevated sex ratio at birth of 1.18. They take an additional migrant sample of "East Europeans", which comprises 17 nationalities from Russia to Bulgaria, and find an elevated ratio of 1.08 on average, and of 1.13 for third parity births and above.

Singh, Pripp, Brekke, and Stray-Pedersen (2010) observe a comparable phenomenon in Norway for immigrants from India (1,600 mothers), but not from Pakistan (5,600 mothers), after analyzing 21,000 births. As most of these immigrants stem from the Punjab region, which according to the authors is fairly similar on both sides of the India-Pakistani border except for religious affiliation, they speculate that religion could be responsible for the differences in sex ratios at birth. Italy has also been examined several times (Ambrosetti et al., 2015; Blangiardo & Rimoldi, 2012; Meldolesi, 2012) with results pointing at skewed sex ratio at birth for migrants originating from India, China and Albania, even more so for higher-order births. Finally, González (2014) documents son-biased sex ratios at birth among Asian-born parents in Spain using birth registries from 2007 to 2012. The male bias remains significant also when controlling for a number of family characteristics. However, all in all the observed male bias among Asian-born parents has little impact in absolute terms (less than 100 girls are "missing"), since that population group is very small in Spain.

In summary, all papers have found that immigrants from countries with known son preference continue to display distorted sex ratios at birth also in their new environment. However, nearly all research focuses exclusively on Asian immigrants, and groups from Southeast Europe are neglected despite their large migration waves and their strong son preference that can be traced back in birth statistics for around 100 years (UNFPA, 2012a). Balkan migrants are also particularly interesting as they belong to the same human race/ethnicity as people in Western and Central Europe (Anemone, 2011). In addition they are reasonably proximate in terms of geography, and Jorde and Wooding (2004, p. 30) argue based on human genetic variation analysis that "most individuals from the same geographic region will be more similar to one another than to individuals from a distant region". Hence, potential biological bias that could cause diverging sex ratios at birth because groups belong to different races/ethnicities (Anderson & Ray, 2010; James, 1987) becomes arguably smaller through this approach. The study by Verropoulou and Tsimbos (2010) examines Albanian migrants in Greece, but it relies on birth data from one single year and contrasts two population groups which live directly adjacent and are therefore probably substantially intertwined. The related work on Spain and Italy focuses geographically also on Southern Europe. Hence, to our knowledge this is the first study to examine prenatal sex selection among Southeast European migrants who moved to Central European countries. We also use data from two host countries (Germany and Switzerland), each spanning more than ten years of birth records, to arguably obtain a sufficiently large sample and robust conclusions.

Implementing Gender Selection

Several options during and before pregnancy have been identified in the literature to implement a higher boy-birth likelihood (for a comprehensive overview, see Guilmoto, 2015). Yet it is essential to understand that such mechanisms for putting son preference into practice are only the outcome channel from underlying son preference. In recent decades, the rise of sex-selective abortions has been a key instrument. Two very common methods, sonography and non-invasive prenatal tests, allow for sex identification at early stages of the pregnancy, upon which the parents may act and abort a female embryo. There is no official documentation about this sex-selective practice, and Westoff (2005) notes that its extent is particularly unknown in Eastern Europe.

The level of technology available matters significantly, in particular access to and cost of prenatal diagnosis (Abrevaya, 2009). Kim (2005) develops a theoretical model, where falling cost of gender detection technologies leads to higher sex ratios, and applies it successfully to pregnancy data of Korean women. For India, Arokiasamy (2007) notes how such technologies, once available, have been widely misused. His observed pattern of extremely high sex ratio for both the last live birth and last two live births in selected Indian states is consistent with the higher levels of use of either sonogram/ultrasound or amniocentesis. Zeng et al. (1993) find similar results for China, and Banister (2004) describes that after ultrasound became available, the shortage of girls emerged in areas of the country where it had not existed before. On the other hand, Junhong (2001) argues that the association is not as evident, as prenatal sex determination by ultrasound is illegal in China so that doctors often refuse to inform the parents. Bongaarts (2013) fears that biases in the sex ratio at birth might even further increase as technology becomes more widespread, since in numerous countries the desired sex ratio exceeds the observed ratio, often by a large margin.

An alternative mechanism to eliminate children of undesired sex are neglect and lack of healthcare, which in theory could also be exercised before birth, but which appear unreliable and hence less common. Instead, adjusting the number of children represents a further main mechanism for parents to achieve desired son preference, leading to son-preferred differential fertility-stopping behavior (Filmer, Friedman, & Schady, 2009). The concept denotes that the desired number of children depends on the number of boy births, where typically no more children follow if the most recent child born was male. Zeng et al. (1993), Park and Cho (1995), and Das Gupta (2005), among others, reaffirm this mechanism empirically by documenting highly skewed sex ratios at last birth. Related research has also shown that contraception increases after a boy birth compared to after a girl birth (Arnold, 1997; Retherford & Roy, 2003), while a girl birth increases the likelihood that a mother shortens the birth interval until the next child (Leone, Matthews, & Zuanna, 2003; Milazzo, 2012). Our contribution to this research consists of analyzing the two major mechanisms proposed (sex-selective abortion and fertility-stopping behavior after a son) for different population groups in Germany and Switzerland. Data from abortion as well as birth statistics help us understand better how parents might implement son preference in Central Europe.

Underlying Motives for Gender Selection

A large research body exists both on roots for gender inequality (among others, Doepke & Tertilt, 2009; Doepke, Tertilt, & Voena, 2012; Jayachandran, 2014b), and more specifically on determinants of abnormal sex ratios (among others, see Das Gupta et al., 2003; Kishor, 1993; Miller, 1985; Park, Bowen, & Steinbacher, 2012; Visaria, 1971). Bhaskar and Gupta (2007) caution to interpret a rise in the population's sex ratio as reflecting son preference or discrimination against women, as other factors impact the sex composition of different age groups as well (see also Calvi, 2015). However, with our exclusive focus on the sex ratio at birth we hope to tackle such omitted variable bias. Hesketh and Xing (2006) summarize more than 30 demographic and environmental factors that have been associated with the sex ratio at birth, many of which are naturally interdependent and thus difficult to isolate. We nonetheless attempt to give a brief overview of the most frequently advanced causes.

Declining fertility levels are widely proposed as major determinant for stronger son preference in many parts of the world. Das Gupta and Bhat (1997) and Guilmoto (2009) conclude for Asian and Caucasian countries that such a link intensifies gender discrimination, as the average number of children decreases. Campbell and Campbell (1997) have similar findings for Botswana. Banister (2004) confirms fertility as a key driver behind increasing sex ratios at birth in China, Jayachandran (2014a) does so for India, and Graham (2007) discusses this phenomenon for Singapore. Lin, Liu, and Qian (2014) provide supporting evidence from an extensive study on Taiwan as well. Klasen (2008), however, contests such a cause-and-effect relationship, and proposes to treat fertility decline as an endogenous variable affected by a range of other determinants in a country. He argues that a decrease in fertility might lead to a reduction of gender preference if a third exogenous variable comes into play that causes fertility decline *and* more gender equality.

Next, the role of public policies has been identified as critical. Klasen and Wink (2003) note that state policy can critically influence gender bias in mortality. They cite female inclusion in education and employment as curbing son preference, whereas the one-child policy in China rather led to the opposite. Sudha and Rajan (1999) summarize that policy measures in India that were to address societal female discrimination were insufficient and largely ineffective. For the United States, Abrevaya (2009) discusses how the political decision to not ban gender-selective procedures might be an indirect policy towards fostering son preference. Nandi and Deolalikar (2013) evaluate in a case study a law in India against sex-selective abortion, using a treatment-effect analysis framework. They conclude with strong positive evidence of the power of public policies in constraining son preference.

Similarly the overall political economy is thought to matter, i.e. the political and economic system in place as well as the effects of changes and socioeconomic turmoil. Nicholas and Oxley (1993) argue that there were effects of the industrialization on the level of son preference, and Horrell, Meredith, and Oxley (2009) similarly link socioeconomic development and gender inequality in an analysis of the 19th century British Empire. Das Gupta and Shuzhuo (1999) analyze

China, Korea, and India throughout the 20th century and identify how specific historical events such as famines and war have impacted the level of son preference in these countries. For the same region, Das Gupta (2010) reaffirms that the type of modern state system in place has been critical in influencing son preference.

Moreover, differences in environmental and geographical conditions have been suggested to explain variations in the sex ratio at birth (Arokiasamy, 2004; Attané & Guilmoto, 2007; Guilmoto, 2008, 2012; Rose, 1999). For example, smog, heavy rainfall or floods, earthquakes, or other impactful events around conception have been proposed to affect gender selection (Fukuda et al., 1998; Hansen et al., 1999, Lyster, 1974).

The "cultural" aspect of simply preferring a son has been pointed out as yet another main reason (Bulte et al., 2011), but a number of authors advanced this notion by investigating how such gender discrimination customs originated. This led to discussion on patriarchal traditions, including patrilocality and patrilinearity. Several papers argue that biased sex ratios at birth are rooted in practices of ancient agriculture-based economics (Aldashev & Guirkinger, 2012; Alesina, Giuliano, & Nunn, 2013; Boserup, 1970; Carranza, 2014; Mayer, 1999), while others point at land inheritance customs that benefit sons (Arokiasamy & Goli, 2012; Jain, 2014; Lahiri & Self, 2005; Lahiri & Self, 2007; Sudha, Khanna, Rajan, & Srivastava, 2007). Increased poverty risk from dowries (Das Gupta, 2000), divorce laws (Sun & Zhao, 2011) and relatively less supply of brawn (Rosenblum, 2013) in case of a daughter further contribute to a patriarchal and son preference culture. A further incentive for male offspring lies in the expected care for parents at old age, for which daughters are widely not considered responsible (Bhasin, 1993; Ebenstein, 2014; Geeta, 2007; Larsen, Chung, & Das Gupta, 1998; Sun, 2002).

Recently, access to media has been found to affect patriarchal traditions (Gillard, Howcroft, Mitev, & Richardson, 2008). Jensen and Oster (2009) document the "power of TV" by observing effects from the introduction of cable television in districts in India. They conclude that this new access to media led to significant decreases in the acceptance rates of domestic violence against women, higher female autonomy, and to lower son preference.

Yet, with our geographical focus on population groups in Germany and Switzerland, we believe that the determinants for son preference cited so far have limited applicability, as they unfold rather in developing countries. Four additional socio-economic variables find support in the literature, and they are in our opinion also more plausible in our research context. Income levels in a household are thought to have a robust association with gender selection practices, although the relationship is not linear but tends to follow an inverted u-curve (Bhat & Zavier, 2007; Klasen & Wink, 2003; Sen, 1990). Guilmoto (2015) confirms that the initially positive relationship between household income and prenatal discrimination can be reversed above a certain social level, with son preference falling among the top income earners. Female labor force participation represents an important element in that context, reducing gender inequality and son preference (Klasen & Wink, 2002; Qian, 2008; Sudha et al., 2007). However, Retherford and Roy (2003) note the monotonous trend of increased propensity for sex-selective abortion with higher (female) socioeconomic status. Murthi, Guio, and Drèze (1995), Banister (2004), Chamarbagwala and Ranger (2006), and Echávarri and Husillos (2016) caution similarly that one should not rely on rising income levels to improve the shortage of daughters.

Higher incomes are closely linked to access to better health, the second variable we consider highly relevant. Many studies identify health differences as a key driver behind missing women (Barcellos, Carvalho, & Lleras-Muney, 2014; Basu, 1992; Bhat, 2002; Bose, 2011; Croll, 2000; Jayachandran & Kuziemko, 2011; Timaeus, Harris, & Fairbairn, 1998) although they tend to focus on post-natal mortality. A recent explanation for pre-natal excess female mortality was centered on the prevalence of hepatitis B, but could not be maintained (Das Gupta, 2005; Klasen, 2008; Lin & Luoh, 2008; Oster, 2005; Oster, Chen, Yu, & Lin, 2010). Eguavoen, Odiagbe, and Obetoh (2007) argue that a woman's decision-making power on health predicts the level of son preference. Bharadwaj and Lakdawala (2013) and Agrawal and Unisa (2007) document preferential prenatal treatment of males in Asia, for example through more frequent tetanus shots and visits of antenatal clinics during pregnancy.[67]

Education levels represent our third socio-economic variable, which is usually associated with decreasing female mortality rates (Arokiasamy, 2007; Chung & Das Gupta, 2007; Drèze & Sen, 1995; Foster & Rosenzweig, 2001; Klasen & Wink, 2003). Guilmoto (2015) conceptualizes these empirical findings with the idea that women's increasing access to education and subsequent financial autonomy decreases traditional patriarchal attitudes, and laws against sex discrimination, for example when transmitting property, are more frequently enforced. Alam, van Ginneken, and

[67] Nutrition as an aspect of health is also discussed in the literature, with mixed results, amongst others in Klasen (1996), Sommerfelt and Arnold (1998), Deaton and Drèze (2009). However, that research examines gender-specific effects *after* birth, which is not our direct focus.

Bosch (2007) in return find no effects, results by Abrevaya (2009) also yield mixed conclusions, and Das Gupta (1987) reports that the sex ratio at birth is actually elevated among more educated women for higher order births.

Lastly, religion has been put forward as important determinant for son preference, both in terms of type of religious affiliation, and the extent of religiousness. For South Korea, Kim and Song (2007) find that Christians display prenatal sex selection less than the Buddhist majority. A similar observation can be made for Muslims in India (Guilmoto, 2008). Almond et al. (2013) conduct a detailed study of the son preference for different religious migrant groups in Canada. They find male biased sex ratios at birth driven by immigrants who are neither Christian nor Muslim, with the "worst in class" being Sikh immigrants. The authors argue that the strong protection of the human life in Christianity and Islam extends also to the prenatal phase, i.e. it largely depresses sex-selective abortion. Yet, Guilmoto (2015) notes that for all main religious groups except Judaism a skewed sex ratio at birth can be found somewhere in the world. Hence, it would likely be a hasty judgment to select a single "son preferring" religion. Due to this fact and with the intense public discussion on the impact of (non-traditional) religious groups in Germany and Switzerland in mind, we decide to include it as variable in our analysis as well.

Literature that examined reasons for ongoing son preference among migration groups could so far not resort to wider individual-level data. Thus little knowledge exists in particular on socio-economic determinants which could explain the observed patterns. With regards to our focus region in Southeast Europe, a report by the United Nations Population Fund (UNFPA, 2012a) on Albania points at deeply entrenched attitudes of male superiority manifesting in every aspect of life, and son preference was already visible in birth registries from the 1920s. The communist regime that ruled until 1990 did not manage to overcome this gender inequality despite programs to emancipate Albanian women, and since then even a reverse trend towards again more traditional values can be observed (see also Morris, Herold, Bino, Yili, & Jackson, 2005).

While the history of Albania's neighboring countries have their idiosyncrasies, we believe that other (Western) Balkan countries with biased sex ratios at birth nonetheless share a similar cultural background. Hence it is implausible that current observations of son preference are coincidental or temporary. However, is there "more" than the cultural background that makes gender discrimination at birth persistent until today, even when migrating to a highly gender-equal environment? The third contribution of this research is to explore why this phenomenon continues to exist, in particular among the emigrated Balkan diaspora in Central Europe, with a focus on the four socio-economic variables income, health, education, and religion.

IV. 3. Data Sources and Description

In order to investigate the number of missing women and potential economic determinants, we resort to three micro data sources. These are national abortion data, federal birth records, and household surveys. Abortion and birth records are fully comprehensive per country, while the survey samples are nationally representative. With this range of sources, we aim to balance individual disadvantages of each while leveraging their combined strength.

Abortion Data

We begin our empirical work by investigating whether abortion data are suitable as they represent the arguably most direct source for sex-selective birth behavior. Germany's current legislation that allows legal abortion has been in effect since October 1995 with the "Schwangeren- und Familienhilfeänderungsgesetz" (SFHÄndG), while Switzerland's equivalent law became effective in October 2002 (articles 119-120 of Swiss StGB).

Data and variable structure provided by the national statistical offices differ considerably between countries: Records for Germany (Federal Statistical Office, 2015b) are available for a longer historical period than for Switzerland (BFS, 2015a), namely since 1996[68] (GER) versus 2007 (CHE). We note with regards to the legislative situation that the abortion statistics in both countries cannot be affected by changes in abortion law, as the respective major amendments described above happened earlier. Both countries collect data for the absolute and relative number of abortions per year, as well as the timing of the abortion in terms of weeks of gestation. Germany furthermore provides data on the reason for abortion and the number of previous live births.

Overall, data for Germany and Switzerland allow only for first tentative conjectures since key variables of interest are not included, a fact that was observed also in other Western countries (Abrevaya, 2009).[69] Nonetheless, including these statistics as part of our broad empirical analysis

---

[68] For better legibility, in the analyses we report 2000 as first year for Germany, but the previous years 1996-1999 are qualitatively identical.

[69] Information on the gender of the embryo is not recorded, and likewise existing (previous) live births are not broken down further along gender. We also miss a detailed citizenship split of the parents which would allow to identify our target immigration groups.

will be useful for an understanding of abortion dynamics, as well as for exploring initial hypotheses on gender-selective practices.

Federal Birth Data

Federal birth data for Germany (Federal Statistical Office, 2015c) and Switzerland (BFS, 2015c) document all individual births per country on an annual basis. Such birth registries are also the most widely used source for research on imbalanced sex ratios at birth, respectively "missing women" (Almond et al., 2013; Attané & Guilmoto, 2007; Chung & Das Gupta, 2007; Klasen & Wink, 2003). For Germany, the key variables of interest, i.e. the citizenship of the mother and the sex of the born child, are available as of 2003, whereas the equivalent records for Switzerland can be retrieved theoretically back until 1970. We will, however, limit our data analysis to the time period since 1990, as this roughly coincides with the fall of the iron curtain and the onset of large migration waves from the Balkans to Central Europe. Note that one limitation of the data in both countries is missing direct information on race and immigration background, respectively parentage. That means we are unable to include foreign-born mothers that have received German or Swiss citizenship in our target sample of immigrants, since we have to rely on the technical classification of citizenship as sole migration characteristic. Mothers with Asian race and German citizenship, for instance, would be accounted towards the overall German population group. As a consequence of this strict migrant versus native classification, our estimates of the extent of a "missing women at birth" phenomenon appear rather conservative.

Additional variables collected in the federal birth data vary between the two countries. In the German data, we find information on the citizenship of the newborn, which allows to draw conclusions on the family's migration history. According to current legislation, if the mother is foreign, the newborn may be German because the father is German (*ius sanguinis*), or because at least one of the two foreign-born parents has lived in Germany as legal resident for a minimum of eight years with a permanent right or residence (*ius soli*) (Federal Foreign Office, 2015).[70]

We can then determine through which of these two channels a newborn receives German citizenship given the mother is foreign-born. The father's citizenship is also recorded in the German birth registry for all married couples, which we re-code for our purposes as a dummy

[70] This rule on citizenship has been into force since the year 2000, hence our data sample covering the years 2003-2014 is fully governed by described legislation.

variable of "being German" versus "carrying the same (foreign) citizenship as the mother". Hence, a newborn that is given German citizenship despite both parents being foreigners must be due to the fact that the family (more precisely, at least one parent) has lived in Germany for at least eight years. In contrast, we know from a newborn with foreign citizenship that none of the parents is German, and that both have lived in Germany for less than eight years.

Figure 4.2 summarizes the logic of all potential outcomes. Our inference regarding the time an immigration family already spent in Germany may directly be associated with its degree of socio-cultural assimilation (Robertson, 2001). Put differently, we are able to examine if the sex ratio at birth differs depending on the level of assimilation of the respective immigration group, measured through the years already spent in Germany. This is a particular feature compared to other national birth data statistics, which we are able to exploit for our research.

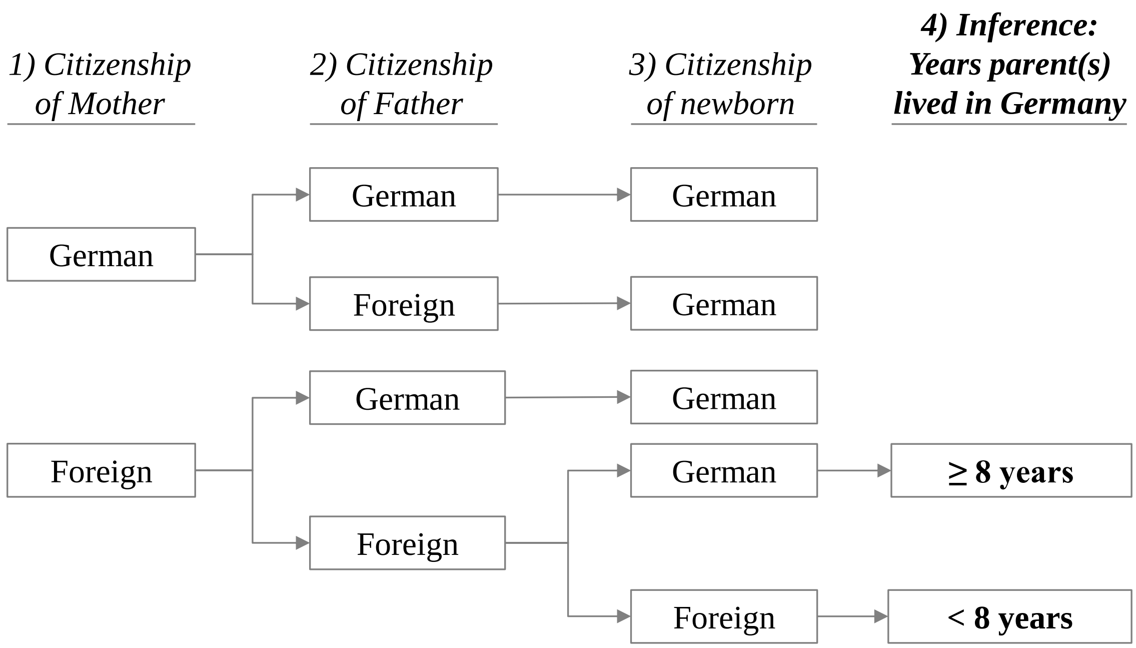


Figure 4.2: Logic of citizenship of newborns and resulting inference on parental time in Germany

The Swiss birth registry includes two other variables that allow additional targeted analyses. It records the age of the mother in seven age brackets, which may serve to examine differential dynamics of gender-selective practices depending on the woman's age and lifecycle situation (Lin et al., 2014; Verropoulou & Tsimbos, 2010). Even more relevant seems the information provided per birth on child parity, i.e. we know how many previous children have been born into the family. While there are unfortunately no data on the sex of these existing children, we can nonetheless investigate the interaction between child parity and sex ratio at birth, which has been suggested as a key relationship (Park & Cho, 1995; Retherford & Roy, 2003; Zeng et al., 1993).

National Household Surveys

As discussed in the literature review, a variety of reasons are put forward to explain sex selective practices for newborns, and determinants why population groups continue to exert such behavior even after they migrated to a new country are not well understood. We attempt to shed new light on this debate towards the end of this chapter, where we move from census to household survey micro data. The German socio economic panel study (SOEP, 2015) is a representative annual study of private households, with nearly 11,000 households, and about 30,000 persons participating in the survey. For our purposes here, it contains useful information on citizenship, migration background, and the number as well as the sex of each child in the household. In addition, it collects a range of socio-economic variables. It is that latter part which we do not find in abortion statistics or birth registries (which are otherwise preferable as they are exhaustive). We base our analysis on version 30 of the SOEP survey, which hence covers the annual rounds from 1984 until 2013. Representative individual data are available for various population groups, including people with migration background.

Given that our statistics so far covered both Germany and Switzerland, it would seem obvious to mirror the micro data analysis also for individuals in Switzerland. To this end, we inspect the suitability of the Swiss Household Panel (SHP), a yearly panel study following a random sample of private households in Switzerland over time. However, the number of households is smaller than the SOEP, and the variables that are of interest to us suffer in particular from poor response rates. For example, the target Balkan and Asian immigrants in the Swiss household survey record only 45 births in total since 1990, which prohibits meaningful analyses, especially since we would want to break those further down along birth parity. Hence, the focus of the micro data analysis will rest on Germany.

Our empirical strategy tests whether a set of demographic and socio-economic variables from the SOEP data are able to explain differences in the children's sex ratio at birth per household, measured as boy-birth likelihood.[71] For this purpose, sufficient data variation and sample size

[71] As we will introduce later, the core regression analysis will consist of a logit model. For this purpose, converting the otherwise preferable outcome variable sex ratio at birth (male births / female births) to a boy-birth likelihood (male births / (male + female births) appears as a more suitable approach.

needs to exist per regression equation. We focus on three specifications: One considers all respondents, i.e. it mirrors the total population in Germany, the second is a sub-sample of foreigners only, i.e. all respondents with migration background, and the third is a sub-sub-sample of only those foreigners with migration background from our target Balkan and Asian countries. As the survey has a set of questions on origin and migration background, we can develop a less technical definition of being a migrant than what we had to resort to in abortion and birth statistics. Together with the questions on sex and parity per child and per household, the survey collects the country of birth per individual through two separate questions.[72] We select all individuals as belonging to our target migration group who indicate as country of birth Albania, Bosnia-Herzegovina, Kosovo, Macedonia, Montenegro, Yugoslavia, China, or India. We then take only the most recent entry per individual to avoid weighing those individuals more who participated several times in the survey, and to have the most up-to-date picture of each household. We observe in several instances that a household in the most recent survey round reports an additional newborn, which is of course critical for our analysis.

In total, these selection criteria yield 208 target migration individuals, which is ca. 0.3 percent of the overall sample of 57,000 cases. This suggests an underrepresentation of this group with respect to the German population, as the German micro census accounts more than two million, or 2.5 percent of the overall population, to these nationalities. Finally, we eliminate all cases with no reported children (either because they explicitly report no children, or because the respective data entries are missing), which leaves us with 108 distinct target migration households. In the SOEP data, this compares to 2,209 foreign households with a general migration background and children, and 16,497 German households with children.

Initial data inspection indicates that the two migration samples (all migrants collectively, and our target migrant sub-sample) may not be large enough to create sufficient data variation, which is needed to explain the comparatively small differences in the sex ratios at birth (see table 4.8 in the appendix for descriptive summary statistics). This is an issue shared with other commonly used sources such as DHS surveys (Guilmoto, 2015). The sex ratios at birth are also surprisingly elevated for the overall sample (which mainly consists of German households) compared to the

[72] These questions are collected as part of the bio information. One entry asks directly for the country of birth; the other asks more generally if the country of birth is outside Germany, and in a second step allows respondents to specify which foreign country they were born in.

migrant sample. The target migration group, however, displays highly male-biased sex ratios at birth up until the third child, which exceed all other samples[73]. Our set of explanatory variables recognizes the general need for "controlling for observable differences in parents' characteristics" (Abrevaya, 2009, p. 15), and takes up the four main determinants of son preference which we discussed in the literature review. Specifically, we look here at underlying *causes* of higher boy-birth likelihood rather than the *mechanisms* to achieve that (stopping rule behavior, sex-selective abortion, neglect of girls, etc.), which literature tends to conflate. We would like to examine a larger set of variables, but the data availability and response rate in the SOEP represents a substantial limitation. In addition to the information on an individual's siblings, our selection thus remains with income levels, religiosity, health, and education.

IV. 4. Empirical Results

Results from Abortion Data

We begin with an overview of abortion dynamics over recent years in Germany and Switzerland (see table 4.1). The total numbers since the 2000s reflect that abortions are indeed sizeable, as on average about one in seven pregnancies is aborted.[74] About half of all pregnant women who decide to abort are single. A number of conceivable scenarios regarding sex-selective abortion may come to mind when inspecting the table. However, none of these can be conclusively verified with the data provided. The summary statistics rather open up an initial informative perspective on the dynamics in both countries, and we summarize in the following first conjectures regarding prenatal gender selection. Yet, we clearly note that the results reflect the need to go beyond abortion data in our "forensic" attempt to find evidence for missing women at birth.

One interpretation of panel B could suggest that availability of early sex-determining technologies may influence abortion dynamics. Both in Switzerland and Germany, elective abortion is only legal up until the twelfth week of gestation, which however, according to some practitioners, already allows to identify the sex given high ultrasound quality (UNFPA, 2012a). Efrat, Akinfenwa, and Nicolaides (1999) find that the accuracy of sex determination increases with

[73] We regard the fourth parity statistics as not reliable given that these comprise only 22 births.
[74] Comparing this figure with Balkan countries proves to be difficult due to the high heterogeneity of officially reported numbers. Albania and Montenegro, for example, report for the 2000s an abortion rate of more than 20%, whereas the rate in Kosovo is only at around 5% (Instat, 2014, p. 51; Johnston 2015).

gestation from 70.3% at 11 weeks, to 98.7% at 12 weeks and 100% at 13 weeks. Parents willing to act on a less than 100% certainty of correct sex determination hence could have a time window to abort within the legal boundaries, and a number of pregnant women time their abortion to just before week 13. In addition, new noninvasive cell-free fetal DNA testing is widely marketed, which allows with a very high level of accuracy sex determination as early as week five after conception (Almond & Edlund, 2008; Mozersky & Mennuti, 2013). While theoretically it is illegal for doctors in Germany and Switzerland to disclose the baby's gender to parents until after week twelve, adherence to this rule is difficult to enforce if results are on hand.

Table 4.1: Summary Statistics of Abortion in Germany and Switzerland

	2000	2007		2014	
	(1)	(2)	(3)	(4)	(5)
	Germany	Germany	Switzerland	Germany	Switzerland
Panel A: Number of legal abortions					
Total	134,609	116,871	10,645	99,715	10,249
Abortions as share of live births	17.5%	17.0%	13.8%	13.8%	11.7%
Share of Single Women	44.5%	52.9%	n/a	57.7%	n/a
Panel B: Weeks of gestation					
8 weeks or less	46%[1]	39%[1]	70%	72%	74%
9-12 weeks	52%[1]	59%[1]	23%	26%	22%
More than 12 weeks	2%	2%	4%	2%	4%
Panel C: Reason for abortion					
Elective abortion	130,945	113,774	n/a	96,080	n/a
Health	3,630	3,072	n/a	3,594	n/a
Rape / Reproductive Coercion	34	25	n/a	41	n/a
Panel D: Previous live births					
Zero	38%	41%	n/a	39%	n/a
One	25%	26%	n/a	25%	n/a
Two or more	36%	33%	n/a	35%	n/a

Footnote 1: Until 2009, Germany applied a different split of the weeks of gestation until week 12 as follows: 7 weeks or less (taken here as proxy for "8 weeks or less"), and 8-12 weeks (taken here as proxy for "9-12 weeks)".
Notes: Detailed data for Switzerland only available as of 2007. Numbers may not add up due to rounding or missing data entries. Switzerland has no data on the reason of abortion and the previous live births.
Source: Federal Statistical Office Germany (Destatis); Swiss Federal Statistical Office (BFS).

After week 12, the two countries report only between two and four percent of abortions. This still translates into 2,780 abortions in Germany for 2014 after gestation week 12, but it remains again speculative how many of those might relate to gender selection. In spite of elective abortion being illegal after week 12, our interest group, i.e. European immigrants from the Balkan, have the realistic alternative to go on a "medical trip" to their home country and receive the desired treatment there. Interviews with Albanian women by the UN note that "there is a law on abortion, but it doesn't get implemented" (UNFPA, 2012a, p. 87). Consequently, cases like the following appear frequently: "My sister-in-law had an abortion when she was at the fourth month of pregnancy. The doctor at the hospital refused to do it, so she paid one of the nurses. The nurse let her in the hospital during the night, wearing a white uniform, and performed the abortion herself" (ibid., p. 85). It seems plausible that pregnant women that are originally from that region will also find a way to abort if their intent and financial background are just strong enough – but this would never appear in the data.[75]

From Panel C we see that at least for Germany elective abortion comprises the overwhelming majority of all cases. This may be interpreted as existence of a societal attitude that widely tolerates a woman's "free" choice to abort – so sex-selective abortion might also be questioned less. Finally, panel D indicates that more than one third of abortions takes place after two children or more. Family planning emerges as key cause for this outcome, in line with declining fertility trends (Banister, 2004; Das Gupta & Bhat, 1997; UNFPA, 2012b; World Bank, 2011). Consequently, there is also a decline as to how often parents might be willing to "try again" for a son. Limiting the number of children has been identified as a major cause of soaring sex ratios at birth (Dyson, 2012; Filmer et al., 2009; Graham, 2007; Guilmoto, 2009; Jayachandran, 2014a). The abortion statistics for Germany, in combination with societal acceptance for abortion, state of the art technology for early sex determination, and the possibility to return relatively easy to the Balkans to abort after week 12, all leave some room for speculations that prenatal gender selection could also exist among migrants in Central Europe. Yet, the findings so far are not conclusive, in particular since we have no information on the gender of aborted embryos.

Evidence from Federal Birth Data

We therefore turn to statistics of the federal birth data as summarized in table 4.2. Aggregate results are broken down by maternal citizenship along the specific birth variables and time periods that are available for Germany and Switzerland as introduced earlier. Based on the insights from figure 4.1, we select three samples of immigrant groups for analysis. First, we look at all immigrants, i.e. all women in Germany who are non-German, respectively all women in Switzerland who are non-Swiss. We then take two sub-samples: On the one hand, we pool

[75] Allahbadia (2002, p. 414) describes a similar pattern for Indian parents in the United States and Canada, who are "courted by American companies" towards highly effective sex selection.

immigrants from the five Western Balkan countries, whose populations have been found to display biased sex ratios at birth (Albania, Kosovo, Montenegro, Macedonia, and Bosnia Herzegovina). On the other hand, we examine Indian and Chinese women to assess their gender selection behavior, and to compare it to the birth dynamics of the highlighted Balkan countries. In addition, we report German, respectively Swiss citizen birth rates as native "baseline" per country. Note that these native ratios are likely biased upwards due to our strict citizenship classification (i.e. all Germans and Swiss with migration background count towards the native ratio), which would lead to rather conservative estimates of population group differences.

The census data indicate that foreign mothers account for a substantial portion of overall births in our timeframe (21% in Germany, and 49% in Switzerland), totaling over two million newborns. In each country, nearly ten percent of these foreign births can be attributed to mothers of the highlighted Balkan countries – more than 170,000 births in total. The largest Balkan immigration group in absolute birth numbers are Kosovars in Germany with 35,000 births, and Macedonians in Switzerland with 26,000 births, over the respective time periods. Chinese and Indian births are little in numbers, reflecting their small population share. We hence note that the distribution weights among immigrant groups is very different from other regions such as North America, which makes our focus on the highlighted Balkan countries all the more relevant.

Our key variable of interest, the sex ratio at birth, yields three interesting findings that hold for both Germany and Switzerland. First, the sex ratio at birth of all foreign mothers shows no sign of gender selection. The ratios in column two are below 1.06 for both countries, which is nearly in line with the ratios of native citizens, and within the range of biologically normal ratios (Anderson & Ray, 2012; Coale, 1991; Klasen & Wink, 2003). However, many of the foreign mothers actually come from neighboring Western European countries. For example, births in Switzerland from Germans alone account for nearly ten percent of foreign births. Hence, potential deviations from normal sex ratios among immigrants are significantly mediated by the large number of fellow Western and Central Europeans who show no gender selection in their countries of origin.

As second insight we note that selected immigrant sub-samples reveal a substantially higher level of sex selection at birth. In Germany, our five Balkan countries combined display a remarkably skewed sex ratio of 1.08, and we find the same ratio for the Chinese and Indian mothers in Switzerland. Given the large number of births from our set of Balkan citizens in Germany, we are able to estimate the sex ratios at birth very precisely. We find that these ratios differ between

Table 4.2: Federal Birth Data in Germany and Switzerland along Citizenship

Mother's Citizenship	(1) Native	(2) All Foreign	(3) 5 Balkan Countries	(4) China and India
Panel A: Germany (2003-2014)				
Total number of births	6,784,294	1,430,242	118,287	22,685
Sex ratio at birth	1.053	1.055	1.077***	1.057
Father same citizenship as mother	n/a	52%	75%	56%
Parents in Germany less than 8 years	n/a	30%	34%	42%
Panel B: Switzerland (1990-2014)				
Total number of births	1,326,729	655,293	58,713	7,362
Sex ratio at birth	1.055	1.059	1.061	1.080
Mother below 30 years	38%	50%	74%	37%
Birth parity	1.79	1.68	1.80	1.46

Notes: The citizenship in columns refers to the mother's citizenship reported at the time of the child's birth. The sex ratio at birth is the ratio of male over female births. "Father same citizenship as mother" indicates the percentage of fathers which has the same citizenship as the mother. As the variable is dummy coded (1 = same citizenship as mother, 0 = German), column 1 has no values since here the two dummy categories coincide. Also, this variable is only collected for married parents, which decreases the sample size by about one third. For Switzerland, "birth parity" denotes the average number of children per mother. Information on this variable is not always collected, so that the sample size for this variable is smaller (there are ca. 250,000 births registered without information on birth parity). Swiss births for 5 Balkan countries include Yugoslavia from 1990-1993. . *, **, *** indicate significance at the 10, 5, and 1 percent level as calculated via a z-test that measures the difference between a given foreign birth and the native birth (German, respectively Swiss)
Source: Federal Statistical Office Germany (Destatis); Swiss Federal Statistical Office (BFS).

Balkan immigrants and Germans on a one percent significance level, i.e. the probability of a newborn being a son is significantly higher if the parents are from the highlighted Balkan region.[76] The ratio for births of Balkan parents in Switzerland is also elevated though not on statistically significant levels.

The findings for both countries are closely aligned with related studies on sex selection practices of immigrants in the Western world (Abrevaya, 2009; Almond et al., 2013; Dubuc & Coleman, 2007; Verropoulou & Tsimbos, 2010). The similarity between Balkan and Asian immigrants may consequently be interpreted as a comparable extent of son preference in South East Europe as has been documented in these Asian countries (Bongaarts & Guilmoto, 2015; Chung & Das Gupta, 2007; Das Gupta, 2010; Zeng et al., 1993).[77] Yet in direct comparison, the

[76] The logic for calculating statistical differences follows Abrevaya (2009), such that we convert the sex ratio at birth into a fraction of boy births out of total births, ranging from 0 to 1, which allows to calculate means, standard errors, and resulting z-tests. We are not able to provide comparable evidence of statistical significance for the Chinese and Indian mothers in Switzerland due to the small sample size and resulting larger standard errors.

[77] While we do not report here the results broken down to individual foreign citizenship due to the comparatively small sample size, qualitative evidence suggests also heterogeneity within the pooled samples. Albanians in Germany, for

migrant ratios observed in the "new" environment (Germany and Switzerland) are below those in the migrants' countries of origin. This may be indirect evidence for migrant assimilation, i.e. migrants tend to adjust their level of sex selection towards natives as they settle down in Central Europe (and we have further detailed analyses in the following to test this hypothesis). Alternatively, it might be due to an attenuating selection effect, meaning that only those migrants come to Central Europe in the first place, who share already a more similar cultural attitude, i.e. lower son preference. Indeed it seems plausible that individuals deciding to migrate are not representative for their original population, but rather more open towards Western traditions and values, including a greater opposition to systematic gender selection at birth.

As third main finding we document that, apart from the within-country variation, there are also between-country differences. Put differently, we see not only that native Germans respectively Swiss and our Balkan immigrants display different sex ratios, but also that the Balkan birth dynamics differ depending on whether the parents live in Germany or in Switzerland. The between-country fluctuations are largest for Chinese and Montenegrin citizens, which we largely attribute to the small sample sizes.[78] The cross-country differences between the average sex ratios at birth for the same population groups are never statistically significant, but they can be observed also in the pooled immigrant groups. Specifically, the sex ratios at birth of the highlighted Balkan countries as well as Chinese and Indian immigrants differ by around 0.02 depending if they are observed in Switzerland or in Germany.

As we previously proposed that a direct comparison between the two countries serves to evaluate consistency of results, it is fair to state that the cross-country findings for biased sex ratios at birth among immigrants are not entirely robust. The fluctuations suggest that prenatal female discrimination is not consistently salient in Germany and Switzerland. Our explorative "forensic" approach is not well suited for predicting such cross-country differences. Perhaps limited sample sizes, and the lack of pronounced gender discrimination at birth contribute to this somewhat non-stationary outcome. We can think of no deeper theoretical reason that would convincingly explain why one of the two country should show a higher value for a particular migration group.

example, have a sex ratio of 1.10 based on 4,294 births, whereas Bosnians have a basically normal ratio of 1.06 based on 27,812 births. The differences closely mirror what can be observed in their countries of origin as well.

[78] See table 4.9 in the appendix for more detailed birth statistics per country.

Inspection of the country-specific variables allows to conjecture on reasons for the varying sex ratios at birth. For the Swiss-specific variables, we observe a striking gap between the Balkan countries and the other groups in terms of the mother's age at birth. Independent of birth parity, only every fourth woman from Albania, Kosovo, Montenegro, Macedonia, or Bosnia-Herzegovina is older than 30 years when giving birth, which contrasts to over 60% of Swiss mothers in that age. Given that early childbearing is more common among women from disadvantaged backgrounds, these findings hint at pronounced socio-economic differences between these groups (Coley & Chase-Lansdale, 1998; Moore et al., 1993; Sonfield, Hasstedt, Kavanaugh, & Anderson, 2013).[79] We also find a negative correlation between elevated sex ratios at birth and the number of children (r= -0.36), which has been observed similarly in the literature (Banister, 2004; Hu & Schlosser, 2015; Jayachandran, 2014a). Mothers from China, for instance, have the highest sex ratio in Switzerland among our groups, while they are on average the oldest and display the lowest fertility levels.

The Germany-specific variables at first glance suggest an association between homogeneity of parental background in terms of common citizenship, and the level of sex selection. Nearly half of the partners of all foreign mothers are Germans, and we observe a balanced sex ratio at birth for this group. For the highlighted Balkan sub-group, however, three quarters of parents share the same Balkan citizenship (only every fourth father is German), and the sex ratio is substantially more skewed. For Asian immigrants, the picture is similar: On aggregate, Chinese and Indian mothers have a German partner in 44% of cases, and biologically normal ratios of sons to daughters. However, a break-down by country reveals that Chinese mothers have German partners in 52% of cases, and a particularly low sex ratio of 1.03. In contrast, only 37% of Indian mothers have a German partner (i.e., 63% have an Indian partner), and they reach a highly elevated sex ratio of 1.09.

While these figures hence suggest an existing correlation, we do not know if the two variables (common citizenship of parents, and sex ratio at birth) are really linked, or if omitted variable bias

[79] We could in principle also assume that the maternal age has a direct effect on the children's sex ratio, for example that a higher age of pregnant women is related to a lower boy-birth likelihood which might explain the sex ratio differences for Swiss versus foreign newborns. However, the medical literature offers no support for such a relationship as overall results are inconclusive. Some studies find a higher likelihood for older mothers to have daughters (Mathews & Hamilton, 2005; Matsuo, Ushioda, and Udoff, 2009), others report the opposite (Takahashi, 1954; Hytten & Leitch, 1971), or conclude that there is no robust association (Rueness, Vatten, & Eskild, 2012; Jacquemyn, Martens, & Martens, 2014). Hence, we do not pursue this hypothesis further.

is present. It could be, for example, that Balkan mothers in general have a strong son preference independent of the citizenship of the father, and that the low number of German partners is unrelated. To tackle such potential bias, we will refine our empirical approach by splitting up the sample further. Similarly, we also interpret the last row in panel A such that the effect from the time spent in Germany requires further analysis, as our basic threshold measure of plus/minus eight years is not conclusive. The impact of both the homogeneity of parental citizenship, and of the length of stay in Germany will hence be examined in further detail in the next section.

Refined Analyses of Birth Statistics per Country

We documented in the previous section that on aggregate levels the sex ratio at birth differs between native citizens, and selected Balkan and Asian immigrant groups. Now we exploit additional country-specific variables from national birth statistics in order to conduct a number of refinement analyses.

German birth statistics provide information on the citizenship of both parents, and indirectly on the time parents with foreign citizenship have been resident in Germany as described earlier. We aim to examine if the time immersed in a new and relatively gender-equal sociocultural environment affects gender-selective practices of immigrants. There is endorsement in the literature for such a relationship (UNFPA, 2012b). Abrevaya (2009) argues that a change of son preference among Asian immigrants could occur in the United States, as second- and third-generation mothers might have a reduced cultural bias, and González (2014) speculates similarly for her country analysis of Spain. However, both studies suffer from insufficient data availability to verify this hypothesis empirically. Almond et al. (2013) find some empirical evidence for a declining son preference in Canada, as second-generation immigrants no longer continued having children in the absence of sons to the same extent as first-generation immigrants.

In table 4.3, we contrast the sex ratios at birth along parental citizenship in panel A, i.e. we look at differences depending if only the mother is immigrant (and the father German), versus both parents being immigrants and sharing the same foreign citizenship (i.e. no parent holds German citizenship). Note that the sample size is about one third smaller than the totals in table 4.2. We are missing information on the citizenship of the father in all those cases, since this is only recorded if the parents are married. Panel B is a deep-dive of the right side of panel A, since it compares the time of permanent residence of all non-German parents, which is our proxy for the level of sociocultural assimilation (see figure 4.2). Given the data structure and legislative situation, we can classify foreign parents in two categories of assimilation, depending on if they have had residence in Germany for longer than eight years or not.

Results in panel A indicate that there are no substantial differences in sex selection at birth depending if both parents share the same foreign citizenship, or if the father is German. This could not be inferred from the aggregate descriptive statistics in table 4.2, which actually rather implied a differential effect depending on parental citizenship homogeneity. However, the highlighted Balkan countries for instance, only move from a sex ratio at birth of 1.073 (father German) to a ratio of 1.079 (father with same foreign citizenship as mother). Note that already the first ratio is considerably above a natural birth rate, which then increases just a bit further if mother and father share the same citizenship.

Table 4.3: Sex Ratio at Birth in Germany along Parental Citizenship and Time of Residence in Germany

Panel A: Parental Citizenship	Father German		Both Parents Foreign	
Mother's Citizenship	Sex Ratio at Birth	Sample Size	Sex Ratio at Birth	Sample Size
German	1.053	4,045,343	n/a	n/a
All Foreign	1.055	519,639	1.054	557,772
5 Balkan Countries	1.073	21,567	1.079	65,513
China and India	1.048	8,978	1.045	11,273
Panel B: Time of residence in Germany	Residence for at least 8 Years		Residence for less than 8 Years	
Mother's Citizenship	Sex Ratio at Birth	Sample Size	Sex Ratio at Birth	Sample Size
German	n/a	n/a	n/a	n/a
All Foreign	1.051	303,585	1.058	254,187
5 Balkan Countries	1.080	40,887	1.078	24,626
China and India	1.030	3,175	1.051	8,098

Notes: Each cell reports the fraction of male over female births along mother's citizenship. In panel A, both parents foreign indicate that the father carries same citizenship as the mother, whose citizenship is given in the very left column. Totals in panel A are about one third smaller than the total of births in table 2, because not all births have information attached regarding the nationality of the father. Panel B is a detailed breakdown of the two columns at the right of panel A, i.e. panel B only contains births from parents who share the same foreign citizenship.
Source: Federal Statistical Office Germany (Destatis).

Our likelihoods in the left column (father German) might also be biased upwards due to our classification which we owe to the data availability. We pointed already earlier at the fact that males with migration background who have obtained German citizenship cannot be differentiated from German males whose ancestors have lived here for generations. This is likely a relevant issue in this context, as foreign females may look for a partner with similar cultural background independent of his actual current citizenship.[80]

Panel B breaks down the all-foreign parents sample by their time of residence in Germany. There is some evidence that the time spent in Germany affects the parents' level of sex selection, although the sex ratio differences are not statistically significant. The Balkan countries on aggregate closely fluctuate around a ratio of 1.08. For Kosovars, who represent the largest both-parents-foreign group with over 20,000 births, the sex ratio at birth moves down from 1.100 with residence less than eight years to 1.088 with residence longer than eight years (see table 4.10 in the appendix). These are all strongly biased ratios, but the sample differences are too closely positioned to be statistically significant.

In summary, intra-generational cultural assimilation measured in length of stay seems to have no significant effect on gender-selective practices of immigrant groups in Germany. This compares to a similar effort by Almond et al. (2013) for Canada, where the authors find that second-generation immigrants still prefer sons at measurable levels. Our findings suggest that several years of residence do not make a substantial difference on how much immigrating parents prefer a son over a girl.

A final analysis specific to Germany examines whether clustering of migrants affects the sex ratio at birth. In other words, has the degree of accumulation of a given population group an additional reinforcing effect on sex selection? The German birth data provide a geographical split along citizenship between all West and East German states (Bundesländer) while excluding Berlin, which we exploit for this purpose. This is because, when disregarding its capital, Germany's migration footprint is remarkably divided between West Germany and the former GDR, today's East Germany. With five percent, the population share with migration background in the East is much lower than in the West with ca. 20 percent (Federal Statistical Office, 2016). This pattern is mirrored in birth statistics: We observe that births by foreign mothers are four times higher in West than in East Germany, relative to births by German mothers. The share of our target group of Balkan migrants relative to the German number of births per region differs even by a factor of ten (2% in the West, 0.2% in the East). Hence, the size and influence of migrant communities differs

[80] While not affecting our interpretations, we found it surprising that for a foreign mother, more births are registered with a foreign father than with a German father, and for our set of Balkan countries both-parents-foreign births are triple as many as only-mother-foreign births.

considerably, which might impact their level of son preference and associated sex selection at birth.

The comparison of the West versus East sex ratios at birth per population group lends some support for such a hypothesis (see table 4.11 in the appendix). Foreigners collectively display no difference between regions, but the target Asian and Balkan immigrants indeed have a higher ratio in West Germany, i.e. where they cluster relatively more. The ratio for these two groups is 0.01 (Balkans), respectively 0.03 (Asians) higher in the West than in the East. However, as birth numbers of these groups in East Germany amount only to a few thousand, these small differences are not statistically significant. Thus, we may only conclude that a trend indication is observable, but a robust pattern cannot be established.[81]

Turning to the Swiss birth registries, additional valuable information provided relates to the birth parity indication, i.e. we know how many live births a mother has previously had. Similar to Abrevaya (2009), we can therefore break down the overall relative sex selection shown per population group along birth parity. Table 4.4 reports statistical results, with panel A pooling all births in Switzerland since 1990, whereas panel B examines births only since the year 2000. In comparison to Swiss birth data in table 4.2 the sample size is reduced since we have not records on birth parity for all births. Moreover, given the naturally decreasing birth sample towards higher parities, we pool all births after first birth up to fourth parity for a combined "higher parity" sex ratio per population group. This ratio hence serves as point of comparison against the respective sex ratio at first birth in order to detect deviating sex selection behavior that can be ascribed to higher parities.[82] In addition, we also present the higher parities individually in the table.

The ratios reveal that the group of Indian and Chinese immigrants in Switzerland seem to engage in substantial sex selection at higher parities. The ratios range from a considerably male-biased value of 1.16 up to 2.60, and the pooled ratio of all their higher parity births is larger than the first parity figure at statistically significant levels. Hence, despite a limited sample we observe clear male bias at higher parities among the Asian target migrants.

---

[81] Note that a comparable analysis for Swiss regions cannot be conducted meaningfully as a geographical split of births results in too small samples and erratic figures. For example, the sex ratio at birth for 2014, even irrespective of parental citizenship, fluctuates between 78.5 and 113.1 on a canton level.

[82] As we focus on the variation of the sex ratio at birth within the same population group but across birth parities, we do not compare the sex ratios to an "unbiased" baseline outside of the given population group. Doing so would conflate the parity effect with the between-population-group effect, whereas here we aim to isolate the parity effect.

For Swiss citizens, higher-parity births are actually less likely to be boys. The large sample allows very precise estimation, and the pattern is nearly linear as parity increases even though the absolute changes are rather small. Abrevaya (2009) explains a very similar outcome for white U.S. citizens with the fact that higher-parity births are "more common among women with lower socioeconomic status and lower-quality prenatal care" (p. 11), which are more prone to harm male fetuses. Hence, we refrain from linking this trend directly to any kind of parental gender preference among Swiss. For all foreign mothers, the sex ratio at birth displays a U-shaped pattern, i.e. it decreases after the first birth, but rises again substantially for the fourth birth.

The sample with five Balkan immigration groups yields very similar results, such that higher parities on aggregate are less likely to be sons than the first birth. The more granular picture along individual parities shows a more male-biased ratio for the first child, followed by rather balanced sex ratios for the second and third child, yet a spiking ratio for the fourth child, which is significantly higher than at first parity in both panels. Considering Macedonian immigrants as example (results are given in table 4.12 in the appendix), we see a particularly strong gender imbalance at birth for fourth parity. Among the 1,100 births registered in Switzerland for that parity and citizenship, there are only 501 girls, which leads to a sex ratio of 1.20. All results are nearly identical for our two time periods reported.

In addition, we are interested whether the full legalization of abortion in Switzerland in 2002 impacted the sex ratio at birth. To this end, we split Swiss birth records in periods 1990-2002 and 2003-2014 and compared the ratios per population group. Again we find no consistent difference in the birth statistics. The total ratio of all births (Swiss and foreigners) slightly increased from 1.053 (1990-2002) to 1.059 (2003-2014), but the ratio for Balkan migrants actually decreased from 1.065 to 1.055 whereas Chinese and Asians rose from 1.066 to 1.087. Hence the liberalization of abortions in Switzerland had no clear impact on the extent of prenatal sex selection. With this in mind, we decide to keep the pooled sample from 1990-2014 as preferred dataset, since the larger number of births allows for more precise and reliable interpretations.

Table 4.4: Sex Ratio at Birth in Switzerland along Mother's Citizenship and Birth Parity

	1st Birth		Higher Parities (2nd - 4th Birth)		2nd Birth		3rd Birth		4th Birth	
	SRB	Sample Size	SRB	Sample Size	SRB	Sample Size	SRB	Sample Size	SRB	Sample Size
Mother's Citizenship	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Full sample (1990-2014)										
Swiss	1.059	501,265	1.053	651,344	1.055	447,379	1.048	154,391	1.049	49,574
All Foreign	1.064	288,897	1.053	293,735	1.050**	211,543	1.050	62,131	1.090**	20,061
5 Balkan Countries	1.080	32,311	1.048	39,374	1.030***	25,440	1.055	10,036	1.148**	3,898
China and India	1.044	4,231	1.157***	2,765	1.145**	2,351	1.148	378	2.600***	36
Panel B: Millenium sample (2000-2014)										
Swiss	1.062	263,407	1.059	338,969	1.058	237,377	1.063	77,777	1.053	23,815
All Foreign	1.062	184,330	1.055	181,316	1.055	133,786	1.049	37,190	1.069	10,340
5 Balkan Countries	1.083	20,629	1.045	23,242	1.033**	16,038	1.044	5,663	1.177**	1,541
China and India	1.032	3,259	1.186***	1,939	1.193**	1,691	1.110	230	1.571	18

Notes: Uneven cells report the share of male over female births along mother's citizenship by birth parity (number of previous births plus 1 per mother). Even columns indicate the total number of births per given birth parity and citizenship. 5 Balkan Countries includes Yugoslavia until 1993. Due to missing entries on birth parity the totals are smaller than the comparable figures in table 4.2. *, **, *** indicate significance at the 10, 5, and 1 percent level as calculated via a z-test that measures the difference between the n^{th} birth and the first birth. Source: Swiss Federal Statistical Office (BFS).

Overall, data for Switzerland confirm what previous literature has concluded on gender preferences of immigrant groups in relation to birth parity (Abrevaya, 2009; Almond et al., 2013; Dubuc & Coleman, 2007; González, 2014). Similar to Germany the detailed perspective for Switzerland provides evidence of sex selection and missing women among selected immigrant groups.[83] However, the phenomenon appears to be confined to few distinct cases and small in absolute numbers. We now estimate how many women seem to be really missing at birth.

Estimates of Missing Women at Birth

In the spirit of the counterfactual exercise by Sen (1990), we aim in this section to quantify how many more women should actually exist in Germany and Switzerland if all parents would behave according to an "unbiased" reference sex ratio at birth. Before discussing the results, two remarks seem appropriate. First, we restrict our estimate of missing women to the number of females "missing" at birth. The literature correctly points at the necessity to observe female death rates throughout a woman's lifetime, since additional factors may also cause an anomalously high number of missing girls during childhood and adolescence, as well as adulthood and old age (Anderson & Ray, 2010, 2015; Milazzo, 2012). However, this is primarily a concern in less developed countries, and we hope to capture nearly all missing women in Germany and Switzerland at the point of birth. We recognize that we might not provide a full picture of excess female death rates, but the existence of additional systematically missing women *after* birth seems unlikely.

Second, the debate on the correct reference sex ratio at birth is far from settled (Bongaarts & Guilmoto, 2015; Klasen & Wink, 2003; Sen, 1992), so that consequently we decide to employ three alternative estimation methods for quantifying missing girls. The first and preferred method is to simply take the observed sex ratio at birth of natives (i.e. Germans in Germany, and Swiss in Switzerland) as reference. We believe this is an intuitive approach, while it also seems conceptually sensible. Both countries are far from displaying systematic female discrimination in terms of gender selection at birth. Therefore, instead of calculating a somewhat arbitrary average of sex ratios at birth in Western countries, we feel more comfortable with taking directly the native

[83] We also employed a second dimension in the data, namely age of the mother, which has been associated with the degree of son preference Verropoulou and Tsimbos (2010); Lin et al. (2014). However, our results for Switzerland are not supportive of any consistent link between those two variables.

ratio of our respective "host" countries. Also, Germans in Germany, and Swiss in Switzerland have each more than one million births recorded in our sample, which provides by itself a highly reliable reference rate.

Nonetheless, we also indicate a second estimate based on the fixed global sex ratio at birth of 1.059 proposed by Coale (1991), and we adopt an alternative approach developed by Klasen and Wink (2003). The latter estimate an individual expected sex ratio at birth per country through a regression in which they link life expectancy to the sex ratio at birth. We apply their regression coefficients and, using life expectancy data from the World Development Indicators, estimate an expected sex ratio at birth per country. The expected sex ratio at birth for a pool of countries (i.e., the highlighted five Balkan countries, as well as China and India combined) is an average of the individual ratios weighted by their number of births recorded in Germany, respectively Switzerland[84].

Results are given in table 4.5, where the observed actual sex ratio at birth is contrasted to each of the three reference rates, which yields the estimated number of missing girls. The number should be read as women who should exist in addition to all born females in these two countries in the time period of 2003 to 2014 in Germany, respectively 1990 to 2014 in Switzerland. Three findings stand out. First of all, the all-foreign group, i.e. immigrants per se, does not lead to women missing at birth in Germany and/or Switzerland. The sex ratio at birth of all foreigners is only marginally higher than the ratio of natives, and any calculation is very sensitive to the choice of reference rate. The counterfactual rates based on the method of Coale or Klasen and Wink indicate no excess female deaths, while the native reference rate does suggest it, but at minimum levels. This outcome is not surprising given the heterogeneity of immigrating foreigners. Most migrated from culturally similar neighboring Western and Central European countries and thus show no son preference, while others come from very different regions including those known for sex-selective practices.

Inspection of immigrant groups from our target countries (which are known for son preference) yields strikingly different results. All our estimates for the highlighted Balkan

[84] See the Appendix for a detailed methodology description. Klasen and Wink (2003) originally use as data source for life expectancy the UN Demographic Yearbook. However, this is unsuitable for our purpose, since our target set of Balkan countries are not included. Also, while an important contribution to the literature, we disregard Anderson and Ray (2010) as alternative estimation method, since they solely focus on China, India, and Sub-Saharan Africa, which play a negligible role as sources of immigration to Germany and Switzerland.

countries, and in five out of six cases also for the China/India sample, indicate an issue of missing women.[85] Independent of the estimation method employed, the numbers are remarkably stable: In Germany, there are systematically at least 1,000 women missing from the five Balkan countries, which is equivalent to two percent of all female births from that population group. For the Asian sample, there are in absolute numbers more missing girls in Switzerland (ca. 80) than in Germany (ca. 40), even though German birth records indicate around three times more births from Chinese and Indian mothers. The estimates are consequently due to the substantially more skewed sex ratio at birth of these Asians in Switzerland, which translates into a higher relative shares of Asian missing women in Switzerland than in Germany.

Finally, comparing results between estimation methods reaffirms that the reference rate of native citizens appears as the most plausible and useful here. This becomes particularly evident when reading the missing girls totals in panel C. The negatively signed results for the “all foreign” group based on the approach by Coale or Klasen and Wink suggest that there are actually too many girls born by immigrants in Germany and Switzerland. This would in return imply a daughter preference in three out of four cases among all foreigners. We think that this sign switch is implausible and due to an imprecise reference rate. The native ratio at birth as reference is the only method which yields a very small number of missing girls from immigrant groups, which seems reasonable given the high level of gender equality in Germany and Switzerland in international comparison. In any case, with these fluctuating estimates depending on the reference rate, it is all the more noteworthy how consistent all estimates are for missing girls from the five Balkan countries as well as from China and India. These two samples are always estimated to cause between 1,100 and 1,600 missing girls in Germany and Switzerland combined over the respective time periods, independent of the reference rate. Thus, while a small number in absolute terms, these selected immigrant groups display an issue of missing girls also in Central Europe, which is unfortunately highly robust. In contrast, all foreigners collectively do not cause a missing girls phenomenon as shown through our range of data analyses.

[85] See also table. 4.13 in the appendix for more detailed results on missing women per population group.

Table 4.5: Estimates of "Missing Women" at Birth in Germany and Switzerland

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Native Ratio			Coale's (1991) method			Klasen and Wink's (2003) method		
Mother's Citizenship	Actual Sex Ratio	Reference Sex Ratio	Missing Women	% missing births	Reference Sex Ratio	Missing Women	% missing births	Reference Sex Ratio	Missing Women	% missing births
Panel A: Germany (2003-2014)										
5 Balkan Countries	1.077	1.053	1,269	2.23%	1.059	958	1.68%	1.055	1,164	2.04%
China and India	1.057	1.053	39	0.35%	1.059	-20	-0.19%	1.053	42	0.39%
All Foreign	1.055	1.053	1,339	0.19%	1.059	-2,381	-0.34%	1.057	-1,375	-0.20%
Panel B: Switzerland (1990-2014)										
5 Balkan Countries	1.061	1.055	205	0.54%	1.059	75	0.20%	1.055	202	0.53%
China and India	1.080	1.055	81	2.30%	1.059	69	1.95%	1.053	89	2.51%
All Foreign	1.059	1.055	1,026	0.32%	1.059	-64	-0.02%	1.057	397	0.12%
Panel C: Totals Germany and Switzerland										
5 Balkan Countries			1,474	1.55%		1,033	1.09%		1,366	1.44%
China and India			120	0.82%		49	0.33%		131	0.90%
All Foreign			2,365	0.23%		-2,445	-0.24%		-977	-0.10%

Notes: Expected sex ratio based on Klasen and Wink's (2003) method are the authors' own calculations based on the regression equation in table 2, column 1 from ibid. Hereby, All Foreign uses the life expectancy from the World Bank for "Europe and Central Asia" as proxy. 5 Balkan Countries in Panel B (Switzerland) includes Yugoslavia until 1993, but Yugoslavia is not included in the weighted expected sex ratio at birth for Klasen and Wink's (2003) method for the 5 Balkan Countries due to missing data on life expectancy at birth in Yugoslavia. % missing births is arrived at by dividing the number of Missing Women by the actual number of female births reported.
Source: Federal Statistical Office Germany (Destatis); Swiss Federal Statistical Office (BFS); World Development Indicators (World Bank).

In this section we employ German household survey data to analyze potential underlying reasons for the diverging sex ratios at birth that we documented based on census data. In other words, we test the explanatory power of a set of socio-economic and demographic variables for the varying sex ratios at birth of different population groups (entire sample, the all-foreign group, and a Balkan and Asian target group). One appealing feature of the German socio-economic panel study (SOEP) for examining this question lies in the fact that the overall sex ratio at birth per population group is reasonably in line with our results from census data: The entire sample representing the population in Germany has a ratio of 1.048, the all-foreign group is slightly higher 1.055, and the target migration group reports a distinctly greater ratio of 1.092. As described earlier, we employ only the SOEP as the sample size for the comparable Swiss household survey is too small. Our first regression analysis consists of the following OLS equation:

(4.1) $\text{Proportion of Male Children}_i = \beta_0 + \beta_1 \text{Migration}_i + \beta_3 X_i + \varepsilon_i$

where *i* designates individual households, and Migration_i is a vector of two dummies whether the individual has a general migration background (all-foreign) and/or originates from our highlighted Balkan and Asian countries. X_i denotes the described additional socio-economic determinants that have been proposed in the literature (Abrevaya, 2009; Almond et al., 2013; Gavalas et al., 2015; Kim & Song, 2007) and have a reasonable response rate in our household survey. These are monthly gross income, level of religiosity, extent of health issues, and education measured as level of schooling; finally ε_i is the heteroscedasticity-robust error term. The dependent variable measures the share of male children in a household, which ranges from zero (no sons, only daughters) to one (only sons, no daughters). The objective of this approach is to examine whether our regressors are able to explain the overall children gender composition in a household, respectively to identify variables that increase the proportion of sons.[86] The advantage of this analysis lies in a larger sample size, as we can include all households with at least one child, and in examining a linear data distribution that complements our second, dichotomous regression analysis.

The second model is based on a categorical outcome variable *P*, which takes the value of 1 if the newborn at indicated parity is a boy, and 0 if it is a girl. Its probabilities are modeled in the following logit equation:

---

[86] We also estimated a fractional logit model where results are qualitatively identical, but with larger coefficients in absolute terms.

(4.2) $P_{i,n}$ (boy = 1) = λ (β_0 + β_1 Migration$_i$ + β_2 Previous Child Female$_{i,n}$ + β_3 X$_i$) ,

where the logit function λ at the right hand side, in addition to the two regressors already introduced for equation (4.1), now employs the dummy Previous Child Female which is coded 1 if any of the previous children in the household are female at given parity, and 0 otherwise[87]; *n* denotes the given birth parity. We hence examine whether our variables are able to explain if the sex of a child per parity is male, and if so, we are interested in the effect indicated by the β coefficients. The equations are estimated using robust standard errors, and we will report coefficients as odds ratios for the dependent variable being male instead of female.[88]

Our procedure for both equations is as follows: We begin by analyzing whether being foreign per se has a significant impact on the children gender composition, respectively the likelihood of having a boy birth. Specifically, in Panel A we take a broad perspective and include a migration background dummy that takes the value of 1 for all-foreign individuals not born in Germany, whereas in Panel B the migration dummy marks only the immigrant sub-sample from our target Balkan and Asian countries. Panel C includes two dummies, i.e. one all-foreign dummy and one target-group dummy, and in Panel D we test whether within all foreigners the target Balkan and Asian migrants have a differential effect.

Table 4.6 presents results for these four main specifications. Inspection of columns one and two for the proportion of male children indicates a clear differentiation of effects depending on the migration group. The all-foreign sample and the target sample carry opposite signs, which are kept throughout all specifications, and even if the overall explanatory power of the model is weak. Being a foreigner per se tends to decrease the share of sons in a household, whereas a migration background from the target Balkan and Asian countries leads to a higher proportion of male children. The significant results in panel C suggest that the all-foreign sample reduces the proportion of male children by two percentages points, whereas migration background from target countries increases it by six percentage points. Also when looking only at foreigners in panel D, the target countries maintain a differential effect towards more sons. None of the additional socio-

[87] Conceptually, it would also be interesting for third and fourth parity births to employ a dummy on whether all of the previous children in the household are female. Such a case, however, is too rare in our sample for meaningful analyses.
[88] In the appendix (tables 4.14 and 4.15) we report the regular coefficients of the logit equations. We also experimented with linear probability models which yielded nearly identical results in terms of sign and significance of coefficients as well as overall model fit. However, due to the dichotomous nature of our dependent variable we prefer to report results from logit models.

economic variables have a meaningful effect, hence we are unable to identify further determinants for a male bias.

Moving on to the logit results in columns three to ten, we find that the likelihood of a boy birth is more affected by the sex of existing children than by a person's migration background, but with changing coefficient signs. The existence of daughters has a significant positive impact on the likelihood of the next child being male, but for second parity only, i.e. when the first child born was a girl. Here, the odds ratios indicate that a boy birth is seven percentage points more likely when the first child was a girl, an effect that remains constant (though no longer significant) if we consider only foreign migrants in panel D. However, the likelihood of a son actually decreases on average for third and fourth births among households in Germany if there have been girls among the older siblings. These higher parity estimates might be somewhat erratic due to finite sample sizes at higher parities, but we also conclude that evidence for sex selection at birth is not consistently found in our household survey.

Next, we have a closer look at the two migration dummies employed, where we focus on second and third parities. We find again that only for the target migration countries the odds ratios are higher than one in nearly all specifications (10 out of 12 in panels B through D), i.e. male-biasing. In contrast, the odds ratio of the all-foreign migration dummy indicates a decreasing boy-birth likelihood throughout all specifications in panels A and C. While the associated standard errors are generally large so that only selected dummies are significant, we nevertheless interpret this sign switch depending on the migration dummy as reaffirming evidence for sex selection among certain immigrant groups only. This is also consistent with the census findings from birth registries, which were not irregular among all foreigners in Germany combined, but substantially biased for our target group. We see a similar link between this sub-sample of Balkan/Asian groups and sex selection displayed here in the SOEP household survey, whereas immigrants overall show an inconspicuous pattern – in our survey data even a higher girl-birth likelihood.

Lastly, our additional controls have in general again no material impact on boy-birth likelihood. Religiosity is the only variable whose odds ratio is consistently below one for our full sample in panels A through C. Estimates suggest that a higher level of religious belief, measured via frequency of church respectively temple visits, decreases the odds for a male child. This likely reflects a general refusal to abort any child independent of gender preference, an attitude which is common among more religious people. In panel D, however, we observe a switch for religiosity from lower to higher boy-birth likelihood after the first birth. This could suggest that the frequency

Table 4.6: Regressions on Proportion of Male Children (OLS), and Boy-Birth Log-likelihood per Parity (Logit, Odds Ratios Shown)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Proportion of Male Children		1st Birth		2nd Birth		3rd Birth		4th Birth	
Panel A: Full sample with migration background dummy										
Migration Background	-0.02	-0.02	0.96	0.97	0.91	0.81	0.96	0.82	0.75	0.37
	(0.01)**	(0.03)	(0.04)	(0.15)	(0.05)*	(0.16)	(0.08)	(0.30)	(0.10)**	(0.22)*
Previous Child Female	n/a	n/a	n/a	n/a	1.07	1.07	0.82	0.93	0.75	1.45
					(0.04)*	(0.07)	(0.07)***	(0.14)	(0.11)**	(0.58)
Additional Controls	no	yes	no	yes	no	yes	no	yes	no	yes
Sample Size	18,706	4,776	18,706	4,776	13,035	3,205	4,672	900	1,490	214
(Pseudo) R-Squared	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03
Wald Chi-Square (p-value)			0.34	0.68	0.05	0.08	0.03	0.70	0.01	0.21
Panel B: Full sample with dummy for migration background from target countries										
Migration Background from Target Countries	0.04	0.12	0.91	0.71	1.04	4.69	1.04	0.99	0.53	0.78
	(0.03)	(0.15)	(0.17)	(0.48)	(0.22)	(5.11)	(0.28)	(1.45)	(0.22)	(0.84)
Previous Child Female	n/a	n/a	n/a	n/a	1.07	1.07	0.82	0.93	0.75	1.45
					(0.04)*	(0.07)	(0.05)***	(0.14)	(0.11)**	(0.58)
Additional Controls	no	yes	no	yes	no	yes	no	yes	no	yes
Sample Size	18,706	4,776	18,706	4,776	13,035	3,205	4,672	900	1,490	214
(Pseudo) R-Squared	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
Wald Chi-Square (p-value)			0.63	0.65	0.18	0.05	0.01	0.75	0.05	0.37
Panel C: Full sample with two migration background dummies										
Migration Background	-0.02	-0.03	0.96	0.98	0.91	0.74	0.96	0.81	0.77	0.41
	(0.01)***	(0.03)	(0.04)	(0.16)	(0.05)*	(0.15)	(0.08)	(0.31)	(0.11)*	(0.24)
Migration Background from Target Countries	0.06	0.15	0.94	0.72	1.13	6.24	1.08	1.21	0.64	0.90
	(0.03)*	(0.15)	(0.18)	(0.50)	(0.24)	(6.92)*	(0.31)	(1.83)	(0.28)	(0.65)
Previous Child Female	n/a	n/a	n/a	n/a	1.07	1.07	0.82	0.93	0.74	1.40
					(0.04)*	(0.08)	(0.05)***	(0.14)	(0.11)**	(0.57)
Additional Controls	no	yes	no	yes	no	yes	no	yes	no	yes
Sample Size	18,706	4,776	18,706	4,776	13,035	3,205	4,672	900	1,490	214
(Pseudo) R-Squared	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03
Wald Chi-Square (p-value)			0.61	0.76	0.08	0.04	0.03	0.81	0.02	0.30
Panel D: All Foreign sub-sample										
Migration Background from Target Countries	0.06	0.20	0.96	0.95	1.17	6.68	0.96	2.02	0.67	
	(0.03)*	(0.15)	(0.19)	(0.67)	(0.25)	(7.27)*	(0.28)	(2.53)	(0.32)	*insufficient sample size*
Previous Child Female	n/a	n/a	n/a	n/a	1.06	1.19	0.69	4.78	0.69	
					(0.11)	(0.49)	(0.12)**	(6.53)	(0.22)	
Additional Controls	no	yes	no	yes	no	yes	no	yes	no	
Sample Size	2,209	172	2,209	172	1,640	107	693	31	275	
(Pseudo) R-Squared	0.01	0.02	0.01	0.01	0.01	0.04	0.01	0.07	0.01	
Wald Chi-Square (p-value)			0.85	0.79	0.64	0.39	0.12	0.84	0.43	

Notes: In columns 1 and 2 the OLS coefficients are reported, where the dependent variable is the share of male children in a household, ranging from 0 (no children are male, all female) to 1 (all children are male, no females). In columns 3 to 10 a logit model estimates the boy-birth likelihood per parity indicated. The coefficient from the logit model is reported for the indicated independent variable along with the robust standard error in parentheses. Migration Background denotes whether the individual is born outside of Germany; Migration Background from Target Countries denotes whether the individual is born in Yugoslavia, Albania, Croatia, Bosnia-Herzegovina, Macedonia, Kosovo, India, or China. Previous Child Female denotes as dummy whether any of the previous children in the household are female. The dependent variable is a dummy whether the child at given birth parity in a household is male (=1) or female (=0). Even columns include the following controls: Monthly gross household income in EUR in logs; the level of religiosity as measured by frequency of church/mosque/temple visits; health issues, which measures the subjective extent of general sickness and lack of health; and the level of schooling. See the appendix for detailed variable description. The sample comprises singular data from 1984-2013; in case of several data entries by the same household individual over different years only the most recent data are used. The number of births deviates from the sample size at higher parities since less and less parents report a second, third, or fourth child. *, **, *** indicate significance at the 10, 5, and 1 percent level. Source: German Socio-Economic Panel Study (SOEP).

of religious activities has a different effect among migration groups than among the general German population, probably because migrants tend to follow a different religion. A differential effect of religious affiliation on son preference has previously been documented in a study on Korea (Kim & Song, 2007). Unfortunately, the data here do not allow a further break-down along religious affiliation due to poor response rates, but among the all-foreign group the level of religiosity rather increases the odds of a male child after first birth. In any case, this variable is never estimated precisely enough in our equations to reach significance levels, so that conclusions remain somewhat speculative.

Overall the explanatory power of the model is quite limited, which is somewhat in line with our expectations. It mirrors the earlier finding that variations of the sex ratio at birth between groups are small. In particular for the first child estimates are likely to be obtained by chance. As we expect parents with gender preferences to display a corresponding deliberate behavior, if any, towards later births (for the first birth many might just "give it a try" for a son), this outcome is plausible. Increasing evidence for son preference towards higher parities has previously been documented among immigrant groups (Almond et al., 2013; Dubuc & Coleman, 2007). A further statistical reason for insignificant findings is simply due to the limited sample size, especially for second and third births: due to poor response rates for the control variables, the theoretically available number of birth observations decreases by around 80 percent, and for panel D, by even more than 90 percent. The sample for fourth parity is small in any case, as only few households actually have four children. This also forces us to omit estimates for the full specification at fourth birth in panel D, as there would be only twelve observations.

Finally, we examine the explanatory power of our socio-economic regressors for our target migrants from the Balkans and Asia only. The objective is to identify potential reasons why these groups maintain sex selection practices in Germany, in spite of living in a very different socio-economic environment than in their home countries. Is that phenomenon due to "cultural heritage" only, or can we identify additional variables that affect the odds of having a male versus a female child? We again estimate an OLS model for the proportion of male children per household in panel A, as well as a logit model for the boy-birth likelihood per given parity in panels B through D. The empirical work is impeded due to a strongly reduced sample, as we only look at survey respondents from our highlighted countries, i.e. a further subset of panel D in table 4.6. For this reason, we cannot include all regressors simultaneously, and we cannot conduct estimates for fourth parity births.

Table 4.7 summarizes results for these target immigration groups. In line with previous findings, we cannot identify significant determinants for the odds of having a male child. This suggests again that observable household characteristics cannot explain the variations of the sex of the child. If anything, religiosity appears to be most suitable for explaining boy-birth likelihoods. The regressor is significant in two out of three cases, and we are able to reaffirm in panels B through D the switch in the odds ratio from less to more than one after first parity, which we found in table 4.6 for the all-foreign group as well.

For parents who migrated from the Balkans, China, or India, religiosity favors the likelihood for a girl at first birth parity. However, for the second child, each additional level of religiosity increases the odds for a male child by 2.4 times. Furthermore, the dummy indicating whether older female siblings exist carries the expected sign, but has a significant effect only at second parity. This is in line with what Abrevaya (2009) reports for Chinese and Indian mothers in California. In our case, given that the first child was a girl, the odds for the second child being a boy increase by about two times. Our remaining explanatory variables employed are erratic. Income levels seem to increase boy birth likelihood but never significantly, and the level of health as well as the level of schooling of an individual yield no consistent picture either. All OLS results in panel A similarly yield no statistical evidence. In summary, the micro level analyses based on the German socio-economic panel study showed the difficulty to establish robust patterns between socio-economic respectively demographic variables, and the sex of a child. We interpret this mostly as a lack of systematic prenatal female discrimination that would be salient enough for statistical significance. In other words, we find no evidence for sex selection in households in Germany, i.e. a newborn being a boy or a girl results overall from a biological and random process, even for migrant sub-groups. In addition, the limited number of observations represent a considerable empirical challenge.

Nonetheless, three further findings stand out from the OLS and logit estimates. First, existing female siblings tend to matter. If the first child is a girl, the odds for a male birth at second parity always increase, though for higher parities the effects from existing female siblings become smaller or even turn negative. Second, if there is any sex selection, it is associated only with selected migrant groups, which is consistent with our earlier findings from birth registries.

Table 4.7: Regressions on Proportion of Male Children (OLS), and Boy-birth Log-likelihood per Parity (Logit, Odds Ratios Shown) for Target Migration Groups

Dependent Variable: Boy-birth likelihood	(1)	(2)	(3)	(4)	(5)
Panel A: Proportion of Male Children					
Monthly Income	0.07				
	(0.04)				
Religiosity		-0.03			
		(0.05)			
Health Issues			-0.02		
			(0.03)		
Level of Schooling				0.03	
				(0.10)	
Sample Size	60	60	108	20	
R-Squared	0.04	0.01	0.01	0.01	
Panel B: First birth					
Monthly Income	1.19				
	(0.30)				
Religiosity		0.42			
		(0.20)*			
Health Issues			1.21		
			(0.20)		
Level of Schooling				1.41	
				(0.64)	
Sample Size	60	60	108	20	
(Pseudo) R-Squared	0.01	0.10	0.01	0.03	
Wald Chi-Square (p-value`	0.51	0.07	0.24	0.46	
Panel C: Second birth					
Monthly Income	1.17				
	(0.36)				
Religiosity		2.42			
		(0.90)**			
Health Issues			0.97		
			(0.17)		
Level of Schooling				2.11	
				(0.97)*	
Previous Child Female					2.04
					(0.87)*
Sample Size	51	51	92	15	92
(Pseudo) R-Squared	0.01	0.09	0.01	0.09	0.02
Wald Chi-Square (p-value`	0.59	0.02	0.88	0.10	0.09
Panel D: Third birth					
Monthly Income	1.32				
	(0.54)				
Religiosity		1.84			
		(0.99)			
Health Issues			1.06		
			(0.25)		
Level of Schooling				0.33	
				(0.54)	
Previous Child Female					1.35
					(0.87)
Sample Size	25	25	54	8	54
(Pseudo) R-Squared	0.02	0.05	0.01	0.05	0.01
Wald Chi-Square (p-value`	0.49	0.25	0.80	0.50	0.64

Notes: In panel A OLS coefficients are reported, where the dependent variable is the share of male children in a household, ranging from 0 (no children are male, all female) to 1 (all children are male, no females). In panels B through D a logit model estimates the boy-birth likelihood per parity indicated. Reported is the coefficient from a logit model for the indicated independent variable along with the robust standard error in parentheses. The dependent variable is a dummy whether the child at given birth parity in a household of the migration target countries (Yugoslavia, Albania, Croatia, Bosnia-Herzegovina, Macedonia, Kosovo, India, China) is male (=1) or female (=0). Previous Child Female denotes as dummy whether any of the previous children in the household are female. The regressors are: Monthly gross household income in EUR in logs; the level of religiosity as measured by frequency of church/mosque/temple visits; health issues, which measures the subjective extent of general sickness and lack of health; and the level of schooling. See the appendix for detailed variable description. The sample comprises singular data from 1984-2013; in case of several data entries by the same household individual over different years only the most recent data are used. Due to the limited sample size not all regressors can be included simultaneously. Sign Switch for Girl-birth Likelihood denotes that when employing the girl-birth likelihood instead of the boy-birth likelihood as outcome variable, the respective regressor switches signs. *, **, *** indicate significance at the 10, 5, and 1 percent level. Source: German Socio-Economic Panel Study (SOEP).

Specifically, the target migration sample of Balkan and Asian origin is more likely to have male children, but migrants altogether in Germany (the all-foreign group) do not share this pattern. Third, the level of religiosity appears to be the only socio-economic factor with some explanatory power, which is closely related to findings by Almond et al. (2013). However, the specific effect of that variable effect depends on the sample. For the total sample, more religiosity leads on average to a lower boy-birth likelihood, which we explain with the tendency among more religious people to oppose abortion. Religiosity among migrants, however, leads to higher odds of having a son, both for all foreigners collectively and for our target migration sub-sample. We believe this is due to the different meaning of religiosity for the average German versus the migrant population, which leads to divergent behavior regarding sex selection practices and probably overall son preference.

IV. 5. CONCLUDING REMARKS

Even in the 21st century discrimination of women remains a central issue (Barcellos et al., 2014; Duflo, 2005), which is not only confined to the developing world. For Europe, prenatal excess female mortality continues to be a reality in certain Balkan countries, and large migration waves from that region could have the potential to disseminate gender selection practices further across the continent. The objective of this chapter has been to examine if migrants, who moved to Germany and Switzerland and who originate in particular from Balkan and Asian countries that are known for son preference, display biased sex ratios at birth also in the new environment. In a "forensic" approach and using different micro data sources, we systematically compared those target migrants to all migrants collectively as well as to the native population. As second step, we attempted to identify underlying motives for gender selection with a focus on socio-economic determinants.

We screened abortion statistics and birth registry data in Germany and Switzerland to examine sex ratios at birth along maternal citizenship. Empirical results indicate that mothers from the target Balkan and Asian countries tend to maintain an overly high sex ratio in their new environment, as compared to native Germans, respectively Swiss. In contrast, the sex ratio at birth of immigrants collectively is only marginally higher than the native population. Hence, Central Europe is not broadly importing the missing women phenomenon, but there is evidence that foreigners from selected countries and regions continue to have biased sex ratios at birth.

Birth data for Germany allowed us to further examine effects from cultural adaptation. However, neither the citizenship of the father, i.e. the distinction if he is German versus from the same country as the mother, nor the parental length of stay in Germany at the point of birth materially affect the level of sex selection. The Swiss data provided additional information on birth parity, which revealed that Indian and Chinese migrants appear to engage in significant sex selection at higher parities, whereas Balkan migrants displayed a strongly elevated sex ratio only at fourth parity. Considering three different reference rates of the sex ratio at birth, we consistently estimated around 1,500 missing girls from the Balkan and Asian immigrant groups in Germany (2003-2014) and Switzerland (1990-2014) combined. Hence while we identified prenatal gender selection among distinct immigrant groups, the impact in absolute perspective is rather small.

We also attempted to identify potential underlying reasons for sex selection at birth, by employing the Germany SOEP household survey. It contains information on sex and parity of each newborn as well as demographic and socio-economic indicators, though the sample of foreign respondents is limited. Overall findings showed the difficulty to establish robust patterns between socio-economic respectively demographic variables, and the sex of a child. This confirms census data results which have not yielded evidence for broad systematic gender selection either. The gender of children of households in Germany is above all determined by nature, i.e. through a random outcome, even for migrant sub-groups. Nevertheless, our regressions suggest that composition of previous parities tends to matter, as the boy-birth likelihood always increases significantly if the first child was a girl. We also observe that any prenatal female discrimination is associated with the selected Balkan and Asian households in Germany, and not with foreigners collectively. This mirrors the aggregate findings from national birth registries. Finally, out of our socio-economic determinants only religiosity comes close to being significant, but the direction of the effect depends on the sample examined. Among all households, more intense religious beliefs decrease sex selection, which might be due to a general opposition to abortion among more religious people. For foreigners, however, more religiosity leads to a higher male bias at birth.

We propose future research to refine the explanatory analyses for gender preferences, which for selected groups seems to persist independent of the geographical environment. Likewise, a more longitudinal perspective could help in the understanding of the level of son preference over time. Policy makers would benefit from enhanced knowledge on the pace of adjustment of the sex ratio at birth towards "normal" levels. Given the global crises together with unfavorable fertility rates and demographic trends in Central Europe, migration will most likely continue to be a key socio-economic variable in the future. We aimed to contribute to an informed debate on its consequences for gender inequality at birth by analyzing the relatively rich set of existing migration data from previous decades. These might also be indicative for gauging the future migration impact on son preference, too. While unexpected trends may change the picture, we currently find effects of pre-natal sex selection to remain small and thus the need for action to be limited.

IV. 6. Appendix

Appendix Table 4.8: Descriptives for the German socio economic panel study (SOEP)

	(1)	(2)	(3)
Averages (unless otherwise indicated)	Total Sample	All Foreign	Target Migration Sample[1]
Sample Size (at least One Child)	18,706	2,209	108
Sample Size (at least Two Children)	13,035	1,640	87
Sample Size (at least Three Children)	4,672	693	49
Sample Size (at least Four Children)	1,490	275	22
Number of Children	1.71	1.80	1.94
Sex Ratio at Birth: First Child	1.08	1.04	1.08
Sex Ratio at Birth: Second Child	1.00	0.93	1.18
Sex Ratio at Birth: Third Child	1.04	1.02	1.33
Sex Ratio at Birth: Fourth Child	1.11	0.90	0.83
Log Average Monthly Income	7.39	7.33	6.95
Religiosity	2.11	2.15	2.08
Health Issues	2.65	2.58	2.79
Level of Schooling	2.06	2.09	1.90

Notes: 1 Target Migration Sample includes all respondents who indicate as country of birth Albania, Bosnia-Herzegovina, Kosovo, Macedonia, Montenegro, Yugoslavia, China, or India. The sex ratio at birth is calculated as total male births over total female births per parity given. Log Average Monthly Income is the gross monthly income in EUR reported per household. Religiosity refers to the frequency of church/temple visits on a scale from 1 (never) to 5 (daily). Health issues is the subjective individual's health situation on a scale from 1 (very good) to 5 (bad). Level of schooling is the highest school education achieved on a scale from 1 (lower secondary school) to 5 (high school graduation with general qualification for university entrance). Source: German socio economic panel study (SOEP).

Appendix Table 4.9: Detailed Statistics for Federal Birth Data in Germany and Switzerland along Citizenship

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Mother's Citizenship	Native	All Foreign	5 Balkan Countries[1]	China and India	Kosovar	Montenegrin	Macedonian	Albanian	Bosnian	Chinese	Indian
Panel A: Germany (2003-2014)											
Total number of births	6,784,294	1,430,242	118,287	22,685	34,552	31,425	20,204	4,294	27,812	11,735	10,950
Sex ratio at birth	1.053	1.055	1.077***	1.057	1.087***	1.075**	1.078*	1.095	1.062	1.032	1.085
Father same citizenship as mother	n/a	52%	75%	56%	72%	79%	83%	52%	72%	48%	63%
Parents in Germany less than 8 years	n/a	30%	34%	42%	31%	41%	31%	37%	31%	39%	44%
Panel B: Switzerland (1990-2014)											
Total number of births	1,326,729	655,293	58,713	7,362	14,011	287	25,966	962	14,144	4,019	3,343
Sex ratio at birth	1.055	1.059	1.061	1.080	1.039	1.126	1.073	1.060	1.048	1.091	1.066
Mother below 30 years	38%	50%	74%	37%	72%	58%	80%	60%	67%	32%	43%
Birth parity	1.79	1.68	1.80	1.46	1.78	1.68	1.75	1.58	1.72	1.43	1.49

1: 5 Balkan Countries includes Yugoslavia from 1990-1993 (38,620 births) with an overall sex ratio at birth of 1.069.
Notes: The citizenship in columns refers to the mother's citizenship reported at the time of the child's birth. The sex ratio at birth is the ratio of male over female births. *, **, *** indicate significance at the 10, 5, and 1 percent level as calculated via a z-test that measures the difference between a given foreign birth and the native birth (German, respectively Swiss). "Father same citizenship as mother" indicates the ratio of fathers in percentage that have the same citizenship as the mother. As the variable is dummy coded (1 = same citizenship as mother, 0 = German), column 7 has no values since here the two dummy categories coincide. Also, this variable is only collected for married parents, which decreases the sample size by about 42,000 births. For Switzerland, "birth parity" is the average number of children per mother. This variable is not collected, so that the sample size for this variable is smaller (there are ca. 200,000 births registered without birth parity allocation) Source: Federal Statistical Office Germany (Destatis); Swiss Federal Statistical Office (BFS).

Appendix Table 4.10: Detailed Sex Ratio at Birth in Germany along Parental Citizenship and Time of Residence in Germany

Panel A: Parental Citizenship	Father German		Both Parents Foreign	
Mother's Citizenship	Sex Ratio at Birth	Sample Size	Sex Ratio at Birth	Sample Size
German	1.053	4,045,343	n/a	n/a
All Foreign	1.055	519,639	1.054	557,772
5 Balkan Countries	1.073	21,567	1.079	65,513
China and India	1.048	8,978	1.045	11,273
Kosovar	1.082	7,722	1.092	20,181
Montenegrin	1.061	4,690	1.080	17,930
Macedonian	1.035	2,656	1.084	12,611
Albanian	1.206	1,352	1.034	1,479
Bosnian	1.058	5,147	1.060	13,312
Chinese	0.987	5,216	1.052	4,897
Indian	1.140	3,762	1.040	6,376
Panel B: Time of residence in Germany	Residence for at least 8 Years		Residence for less than 8 Years	
Mother's Citizenship	Sex Ratio at Birth	Sample Size	Sex Ratio at Birth	Sample Size
German	n/a	n/a	n/a	n/a
All Foreign	1.051	303,585	1.058	254,187
5 Balkan Countries	1.080	40,887	1.078	24,626
China and India	1.030	3,175	1.051	8,098
Kosovar	1.088	12,895	1.100	7,286
Montenegrin	1.082	9,959	1.076	7,971
Macedonian	1.094	8,598	1.062	4,013
Albanian	0.949	651	1.107	828
Bosnian	1.064	8,784	1.053	4,528
Chinese	0.993	1,152	1.070	3,745
Indian	1.052	2,023	1.034	4,353

Notes: Each cell reports the fraction of male over female births along mother's citizenship. In panel A, both parents foreign indicate that the father carries same citizenship as the mother, whose citizenship is given in the very left column. Panel B is a detailed breakdown of the two columns at the right of panel A, i.e. panel B only contains births from parents who share the same foreign citizenship. Source: Federal Statistical Office Germany (Destatis).

Appendix Table 4.11: Federal Birth Data in West and East Germany along Citizenship

	(1)	(2)	(3)	(4)
Mother's Citizenship	Native	All Foreign	5 Balkan Countries	China and India
Panel A: West Germany (2003-2014)				
Total number of births	5,355,309	1,279,692	110,039	18,882
Sex ratio at birth	1.054	1.056	1.077	1.058
Panel B: East Germany (2003-2014)				
Total number of births	1,135,770	58,227	2,397	1,693
Sex ratio at birth	1.050	1.056	1.066	1.025
Difference	0.004	—	0.011	0.033

Notes: The citizenship in columns refers to the mother's citizenship reported at the time of the child's birth. The sex ratio at birth is the ratio of male over female births. West Germany comprises all West German states (Bundesländer), and East Germany all East German states. Berlin is excluded from both samples, which yields a reduction of the total number of births of East and West together, compared to Germany totals.

Appendix Table 4.12: Detailed Sex Ratio at Birth in Switzerland along Mother's Citizenship and Birth Parity

	1st Birth		Higher Parities (2nd - 4th Birth)		2nd Birth		3rd Birth		4th Birth	
	SRB	Sample Size	SRB	Sample Size	SRB	Sample Size	SRB	Sample Size	SRB	Sample Size
Mother's Citizenship	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Full sample (1990-2014)										
Swiss	1.059	501,265	1.053	651,344	1.055	447,379	1.048	154,391	1.049	49,574
All Foreign	1.064	288,897	1.053	293,735	1.050**	211,543	1.050	62,131	1.090**	20,061
5 Balkan Countries	1.080	32,311	1.048	39,374	1.030***	25,440	1.055	10,036	1.148**	3,898
China and India	1.044	4,231	1.157***	2,765	1.145**	2,351	1.148	378	2.600***	36
Kosovar	1.045	6,108	1.036	7,378	1.005	4,823	1.102	1,972	1.090	583
Montenegrin	1.577	134	0.778***	128	0.698***	90	1.364	26	0.500**	12
Macedonian	1.099	11,895	1.054	13,502	1.055	9,084	1.011**	3,318	1.196	1,100
Albanian	1.016	494	1.090	370	1.094	266	1.079	79	1.083	25
Bosnian	1.076	6,057	1.032	7,158	1.018	5,146	1.078	1,687	1.019	325
Chinese	1.067	2,338	1.162***	1,377	1.155	1,166	1.148	189	1.750	22
Indian	1.061	1,893	1.152***	1,388	1.135	1,185	1.148	189	6.000***	14
Panel B: Millenium sample (2000-2014)										
Swiss	1.062	263,407	1.059	338,969	1.058	237,377	1.063	77,777	1.053	23,815
All Foreign	1.062	184,330	1.055	181,316	1.055	133,786	1.049	37,190	1.069	10,340
5 Balkan Countries	1.083	20,629	1.045	23,242	1.033**	16,038	1.044	5,663	1.177**	1,541
China and India	1.032	3,259	1.186***	1,939	1.193**	1,691	1.110	230	1.571	18
Kosovar	1.045	6,108	1.036	7,378	1.005	4,823	1.102	1,972	1.090	583
Montenegrin	1.577	134	0.778***	128	0.698***	90	1.364	26	0.500**	12
Macedonian	1.090	9,751	1.051	10,781	1.052	7,499	0.987**	2,540	1.297**	742
Albanian	1.143	330	1.061	235	0.967	177	1.136	47	4.500**	11
Bosnian	1.107	4,306	1.050	4,720	1.047	3,449	1.069	1,078	1.010	193
Chinese	1.059	1,839	1.180***	1,027	1.194	895	1.140	122	0.667	10
Indian	0.997	1,420	1.192***	912	1.193	796	1.077	108	7.000***	8

Notes: Uneven cells report the share of male over female births along mother's citizenship by birth parity (number of previous births plus 1 per mother). Even columns indicate the total number of births per given birth parity and citizenship. 5 Balkan Countries includes Yugoslavia until 1993. Due to missing entries on birth parity the totals are smaller than the comparable figures in table 4.9. *, **, *** indicate significance at the 10, 5, and 1 percent level as calculated via a z-test that measures the difference between the n^{th} birth and the first birth. Source: Swiss Federal Statistical Office (BFS).

Background Information on the Reference Sex Ratio at Birth based on Klasen and Wink (2003):

Klasen and Wink (2003), building on Klasen (1994), propose a "variable" sex ratio at birth, arguing that a secular upward trend in the sex ratio at birth in rich countries requires rejection of a "stable" sex ratio at birth. "Whenever better health and nutrition lower the rates of spontaneous abortions and miscarriages and reduce the incidence of stillbirths, the sex ratio at birth increases" (ibid., p. 269). As people in more developed countries also enjoy higher life expectancies, they conduct a regression of the observed sex ratio at birth on the life expectancy in a country. Hereby the sample with observed sex ratios at birth consists of countries which have complete birth registration data, at least 5,000 births per year, and no evidence of sex-selective abortion.

Then, using the estimated regression coefficients, they calculate an expected "unbiased" sex ratio at birth per country that is now associated with the respective average life expectancy. Hence, in addition to biological factors, the authors also consider the individual state of development per country, proxied by the life expectancy, which influences the sex ratio at birth. As data source for life expectancy and the observed sex ratio at birth they use the UN Demographic Yearbook.

Generally, the expected "unbiased" sex ratio at birth by Klasen and Wink (2003) for countries with excess female mortality (e.g., China, India) are below the ratio for Coale (1991), leading to a higher estimated number of "missing women".

For our purposes, we adopt the regression coefficients estimated by Klasen and Wink (2003), consisting of the constant = 0.991, and the beta for life expectancy = 0.00087. By taking life expectancy data from the World Development Indicators, we the estimate an expected sex ratio at birth for each country in our sample. For combined migration groups, e.g. the group of Indian and Chinese mothers together, we take an average of the individual country results weighted by the actual number of births in Germany and Switzerland. For the all-foreign group, we proxy the average life expectancy of all migrants by employing the aggregate World Bank life expectancy value for "Europe and Central Asia", yielding an expected sex ratio at birth of 1.057.

Appendix Table 4.13: Detailed Estimates of "Missing Women" at Birth in Germany and Switzerland

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Native Ratio			Coale's (1991) method			Klasen and Wink's		
Mother's Citizenship	Actual Sex Ratio	Reference Sex Ratio	Missing Women	% missing births	Reference Sex Ratio	Missing Women	% missing births	Reference Sex Ratio	Missing Women	% missing births
Panel A: Germany (2003-2014)										
5 Balkan Countries	1.077	1.053	1,269	2.23%	1.059	958	1.68%	1.055	1,164	2.04%
China and India	1.057	1.053	39	0.35%	1.059	-20	-0.19%	1.053	42	0.39%
Kosovar	1.087	1.053	532	3.22%	1.059	441	2.67%	1.052	551	3.33%
Montenegrin	1.075	1.053	316	2.09%	1.059	234	1.54%	1.056	279	1.84%
Macedonian	1.078	1.053	225	2.32%	1.059	172	1.77%	1.056	198	2.04%
Albanian	1.095	1.053	80	3.92%	1.059	69	3.36%	1.058	71	3.44%
Bosnian	1.062	1.053	114	0.85%	1.059	42	0.31%	1.057	65	0.49%
Chinese	1.032	1.053	-117	-2.02%	1.059	-147	-2.55%	1.056	-133	-2.30%
Indian	1.085	1.053	155	2.96%	1.059	127	2.41%	1.049	176	3.35%
All Foreign	1.055	1.053	1,339	0.19%	1.059	-2,381	-0.34%	1.057	-1,375	-0.20%
Panel B: Switzerland (1990-2014)										
5 Balkan Countries	1.061	1.055	205	0.54%	1.059	75	0.20%	1.055	202	0.53%
China and India	1.080	1.055	81	2.30%	1.059	69	1.95%	1.053	89	2.51%
Kosovar	1.039	1.055	-106	-1.54%	1.059	-129	-1.87%	1.052	-85	-1.24%
Montenegrin	1.126	1.055	9	6.68%	1.059	9	6.32%	1.056	9	6.63%
Macedonian	1.073	1.055	209	1.67%	1.059	165	1.32%	1.056	199	1.59%
Albanian	1.060	1.055	2	0.43%	1.059	0	0.09%	1.058	1	0.16%
Bosnian	1.048	1.055	-48	-0.69%	1.059	-71	-1.03%	1.057	-59	-0.86%
Chinese	1.091	1.055	65	3.38%	1.059	58	3.03%	1.056	63	3.28%
Indian	1.066	1.055	16	0.99%	1.059	11	0.67%	1.049	26	1.60%
All Foreign	1.059	1.055	1,026	0.32%	1.059	-64	-0.02%	1.057	397	0.12%
Panel C: Totals Germany and Switzerland										
5 Balkan Countries			1,474	1.55%		1,033	1.09%		1,366	1.44%
China and India			120	0.82%		49	0.33%		131	0.90%
Kosovar			427	1.40%		312	1.02%		466	1.52%
Montenegrin			325	2.11%		242	1.57%		288	1.87%
Macedonian			434	1.95%		337	1.52%		397	1.78%
Albanian			82	2.73%		69	2.30%		71	2.37%
Bosnian			67	0.74%		-29	-0.33%		6	0.07%
Chinese			-52	-0.67%		-89	-1.15%		-70	-0.91%
Indian			171	2.50%		137	2.00%		202	2.94%
All Foreign			2,365	0.23%		-2,445	-0.24%		-977	-0.10%

Notes: Expected sex ratio based on Klasen and Wink's (2003) method are the authors' own calculations based on the regression equation in table 2, column 1 from ibid. Hereby, All Foreign uses the life expectancy from the World Bank for "Europe and Central Asia" as proxy. 5 Balkan Countries in Panel B (Switzerland) includes Yugoslavia until 1993, but Yugoslavia is not included in the weighted expected sex ratio at birth for Klasen and Wink's (2003) method for the 5 Balkan Countries due to missing data on life expectancy at birth in Yugoslavia. % missing births is arrived at by dividing the number of Missing Women by the actual number of female births reported.
Source: Federal Statistical Office Germany (Destatis); Swiss Federal Statistical Office (BFS); World Development Indicators (World Bank).

Appendix Table 4.14: Boy-Birth Log-likelihood Regressions (Logit, Regular Coefficients Shown)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1st Birth		2nd Birth		3rd Birth		4th Birth	
Panel A: Full sample with migration background dummy								
Migration Background	-0.04	-0.03	-0.09	-0.21	-0.04	-0.20	-0.29	-0.98
	(0.04)	(0.16)	(0.05)*	(0.20)	(0.08)	(0.37)	(0.13)**	(0.57)*
Previous Child Female	n/a	n/a	0.06	0.06	-0.20	-0.07	-0.30	0.37
			(0.04)*	(0.07)	(0.07)***	(0.15)	(0.14)**	(0.40)
Additional Controls	no	yes	no	yes	no	yes	no	yes
Sample Size	18,706	4,776	13,035	3,205	4,672	900	1,490	214
(Pseudo) R-Squared	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03
Wald Chi-Square (p-value)	0.34	0.68	0.05	0.08	0.03	0.70	0.01	0.27
Panel B: Full sample with dummy for migration background from target countries								
Migration Background from Target Countri	-0.09	-0.34	0.04	1.54	0.04	-0.01	-0.64	-1.08
	(0.19)	(0.67)	(0.21)	(1.09)	(0.28)	(1.47)	(0.42)	(0.84)
Previous Child Female	n/a	n/a	0.06	0.07	-0.20	-0.08	-0.30	0.30
			(0.04)*	(0.07)	(0.07)***	(0.15)	(0.14)**	(0.41)
Additional Controls	no	yes	no	yes	no	yes	no	yes
Sample Size	18,706	4,776	13,035	3,205	4,672	900	1,490	214
(Pseudo) R-Squared	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
Wald Chi-Square (p-value)	0.63	0.65	0.18	0.05	0.01	0.75	0.05	0.37
Panel C: Full sample with two migration background dummies								
Migration Background	-0.04	-0.02	-0.10	-0.30	-0.04	-0.21	-0.26	-0.90
	(0.05)	(0.16)	(0.05)*	(0.21)	(0.08)	(0.38)	(0.14)*	(0.59)
Migration Background from Target Countri	-0.06	-0.33	0.13	1.83	0.08	0.19	-0.45	-0.71
	(0.19)	(0.69)	(0.21)	(1.11)*	(0.28)	(1.51)	(0.43)	(0.65)
Previous Child Female	n/a	n/a	0.06	0.07	-0.20	-0.07	-0.30	0.34
			(0.04)*	(0.07)	(0.07)***	(0.15)	(0.14)**	(0.40)
Additional Controls	no	yes	no	yes	no	yes	no	yes
Sample Size	18,706	4,776	13,035	3,205	4,672	900	1,490	214
(Pseudo) R-Squared	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03
Wald Chi-Square (p-value)	0.61	0.76	0.08	0.04	0.03	0.81	0.02	0.30
Panel D: All Foreign sub-sample								
Migration Background from Target Countri	-0.04	-0.05	0.16	1.90	-0.04	0.71	-0.40	
	(0.20)	(0.70)	(0.22)	(1.09)*	(0.30)	(1.25)	(0.47)	
Previous Child Female	n/a	n/a	0.06	0.18	-0.37	1.57	-0.36	*insuffi-*
			(0.10)	(0.41)	(0.18)**	(1.36)	(0.32)	*cient*
Additional Controls	no	yes	no	yes	no	yes	no	*sample*
Sample Size	2,209	172	1,640	107	693	31	275	*size*
(Pseudo) R-Squared	0.01	0.01	0.01	0.04	0.01	0.07	0.01	
Wald Chi-Square (p-value)	0.85	0.79	0.64	0.39	0.12	0.84	0.43	

Notes: Reported is the coefficient from a logit model for the indicated independent variable along with the robust standard error in parentheses. Migration Background denotes whether the individual is born outside of Germany; Migration Background from Target Countries denotes whether the individual is born in Yugoslavia, Albania, Croatia, Bosnia-Herzegovina, Macedonia, Kosovo, India, or China. Previous Child Female denotes as dummy whether any of the previous children in the household are female. The dependent variable is a dummy whether the child at given birth parity in a household is male (=1) or female (=0). Even columns include the following contrds: Monthly gross household income in EUR in logs; the level of religiosity as measured by frequency of church/mosque/temple visits; health issues, which measures the subjective extent of general sickness and lack of health; and the level of schooling. See the appendix for detailed variable description. The sample comprises singular data from 1984-2013; in case of several data entries by the same household individual over different years only the most recent data are used. The number of births deviates from the sample size at higher parities since less and less parents report a second, third, or fourth child. *, **, *** indicate significance at the 10, 5, and 1 percent level. Source: German Socio-Economic Panel Study (SOEP) .

Appendix Table 4.15: Boy-birth Log-likelihood Regressions for Target Migration Groups (Logit, Regular Coefficients Shown)

Dependent Variable: Boy-birth likelihood	(1)	(2)	(3)	(4)	(5)
Panel A: First birth					
Monthly Income		0.18			
		(0.25)			
Religiosity			-0.86		
			(0.47)*		
Health Issues				0.19	
				(0.17)	
Level of Schooling					0.34
					(0.46)
Sample Size		60	60	108	20
Wald Chi-Square (p-value)		0.51	0.07	0.24	0.46
Panel B: Second birth					
Previous Child Female	0.71				
	(0.43)*				
Monthly Income		0.16			
		(0.30)			
Religiosity			0.88		
			(0.37)**		
Health Issues				-0.03	
				(0.18)	
Level of Schooling					0.75
					(0.46)*
Sample Size	92	51	51	92	15
Wald Chi-Square (p-value)	0.09	0.59	0.02	0.88	0.10
Panel C: Third birth					
Previous Child Female	0.30				
	(0.64)				
Monthly Income		0.28			
		(0.41)			
Religiosity			0.61		
			(0.54)		
Health Issues				0.06	
				(0.24)	
Level of Schooling					-1.10
					(1.63)
Sample Size	54	25	25	49	54
Wald Chi-Square (p-value)	0.64	0.49	0.25	0.80	0.50

Notes: Reported is the coefficient from a logit model for the indicated independent variable along with the robust standard error in parentheses. The dependent variable is a dummy whether the child at given birth parity in a household of the migration target countries (Yugoslavia, Albania, Croatia, Bosnia-Herzegovina, Macedonia, Kosovo, India, China) is male (=1) or female (=0). Previous Child Female denotes as dummy whether any of the previous children in the household are female. The regressors are: Monthly gross household income in EUR in logs; the level of religiosity as measured by frequency of church/mosque/temple visits; health issues, which measures the subjective extent of general sickness and lack of health; and the level of schooling. See the appendix for detailed variable description. The sample comprises singular data from 1984-2013; in case of several data entries by the same household individual over different years only the most recent data are used. Due to the limited sample size not all regressors can be included simultaneously. Sign Switch for Girl-birth Likelihood denotes that when employing the girl-birth likelihood instead of the boy-birth likelihood as outcome variable, the respective regressor switches signs. *, **, *** indicate significance at the 10, 5, and 1 percent level. Source: German Socio-Economic Panel Study (SOEP).

V. General Conclusion and Research Outlook

V. 1. New Perspectives on Inequality

Inequality is a central theme in economics, and it continues to have outstanding importance in public discussions around the globe. This dissertation was set out to explore distinct aspects of inequality, namely of inequality in income and gender, and two factors were salient for motivating this research focus.

On the one hand, in spite of declining poverty rates in absolute levels, there is evidence that income inequality in the world is growing (Milanovic 2009; 2011). In a recent study, Oxfam (2016) claimed that the richest one percent now have more wealth than the rest of the world combined, and 62 people own the same as half the world. Are variables which are typically associated with development and believed to affect economic growth also responsible for this trend? In other words, do determinants for development cause at the same time larger income inequalities? Consequences for national income distribution from changes in variables associated with development, such as trade integration or institutional quality, are subject to ongoing controversies. While effects for the overall inequality in a country have been examined, more granular analyses of the changes in the underlying income distribution are rare. It is essential, however, to know if a certain development determinant affects only, say the richest ten percent in a country, or the entire population.

On the other hand, broad consensus arose that inequalities between individuals also have an impact on economic development when the inequalizing feature is not personal income, but other individual characteristics. Discriminating behavior towards women, for example, may not only be morally unjustified, but also causes direct negative economic effects. Hence, working towards equality of opportunities between genders, and removing obstacles that prevent women from reaching their full economic potential will be positive for women's individual freedom, and for the economic development of a nation. But many sides of gender inequality are still poorly understood, especially when considering the issue from an economic angle. Is there an underlying economic rationale why in many parts of the world girls receive less education than boys? What might be economic reasons for families to desire a son rather than a daughter; and if families are immersed in a new socio-economic environment, does this affect their sex selection behavior? Examining such questions in order to learn more about gender inequality comprise the second major motivation for our work.

The scale of both debates, inequality in income and in gender, is extensive and multifaceted. In the introductory chapter, we documented the voluminous income inequality research and discussed its importance for economics throughout time. We also pointed at the strong renewed interest it receives at present. Gender inequality is a younger theme in economics, but has also gained immense traction towards its high relevance today in academia as well as policy-making. Hence, our contribution aims to advance the economic development literature by providing new perspectives not only on income inequality, but also on two aspects of gender inequality.

V. 2. Synthesis of Dissertation Findings

For synthesizing our main findings, interpreting them, and linking them to the literature, we revisit in this conclusion each of the three research questions of this book as addressed in detail in the respective chapter. Chapter 2 covered an aspect of income inequality by seeking to answer: What are the effects of fundamental development factors on different income groups? We advanced the related literature with this empirical approach by going beyond analyses of average income levels or overall inequality in a country, so that observed aggregate inequality changes can be pinpointed to specific income groups. This work was also the first to examine all key growth regressors more systematically, and regressions were estimated for several time periods and in various specifications for additional robustness.

We examined effects in an explorative manner through a set of cross-country growth regressions. The dataset was newly constructed for 138 countries over 30 years by combining income distribution data from the World Income Inequality Database with average income levels from the Penn World Tables. Consequently, we were able to analyze effects on different income groups and systematically examined the lowest quintile, the median, the average, the top quintile, and the top decile of the income distribution. Focusing on the development factors geography, trade integration, and institutional quality, and applying the established instruments for the latter two in order to ensure causal interpretation, the empirical estimates of the five income groups over six five-year periods yielded a number of interesting findings.

The point estimates generally confirm the related literature results for average income levels, but there are indeed substantial differential effects of the variables on low versus high income groups. The impact of geography (distance from the equator) clearly discriminates between income groups, as its coefficients strongly decrease from bottom to top incomes. The variable even switches signs in many cases once we passed the mean income, suggesting that the rich – in contrast to the poor – actually benefit from equator proximity. In fact, we showed that for geography the point estimates for the top incomes lie generally outside the 95 percent confidence interval of the poorest income. While the variable tends to lose significance at conventional levels when controlling for trade and institutional quality, the coefficient dynamics are very robust. In an alternative specification where we measured geographical conditions by the prevalence of malaria, the point estimates turned statistically significant and again keep the observed pattern.

Trade integration, measured as the share of exports and imports combined of national GDP and instrumented with Frankel and Romer's (1999) constructed trade shares, displays negative effects for all income levels. Yet, the granular perspective revealed that the coefficient values for rich incomes turn out to be higher in (absolute) size and significance levels as compared to the poor incomes. This suggests an equalizing effect of trade for different income levels within a country, as the rich are comparatively worse affected than the poor. We also noted that trade adds nearly no explanatory power to the model, so that the overall association between trade volumes and income levels is rather weak.

Institutional quality, the third fundamental factor for development we examined, is instrumented in our main specification following the approach of Hall and Jones (1999) using language data. The variable displays positive and highly significant values for all incomes, i.e. better institutions are beneficial for all income groups. It also increases the explanatory power of the model substantially. When comparing income groups, we noted an increasing coefficient from bottom to top, such that the effect on top income groups is on average 20 percent higher than for the bottom income groups. We confirmed these results when using settler mortality data based on the work by Acemoglu, Johnson and Robinson (2001) as alternative instrument.

Overall, coefficient trends move evenly across income groups so that results for median and average income groups are close to a linear interpolation of top and bottom incomes. Our results are consistent over time but the explanatory power of the empirical analysis increases for lower incomes. A range of econometric tests confirmed the model's overall validity and we controlled our results for a number of additional variables. These indicate that world regions have a sizeable effect on the results. In addition, the variables health and human capital, the latter using lagged values as instrument, enter significantly but do not alter the described patterns of the fundamental development factors. In particular our estimates provided no direct support for the hypothesis that human capital is a more basic source for growth than institutions.

In chapter 2 we analyzed for the first time simultaneously the set of variables that are widely thought to matter most for development. For average income levels, the findings closely confirm results by Rodrik (2004). More importantly, our study provides an alternative perspective compared to the conclusions suggested for different income groups by Dollar et al. (2002; 2013; 2014). These papers generally do not find robust association between growth variables and the income differences of rich versus poor groups. They have taken a narrower approach by examining rather detailed variables, while not employing all fundamental development variables simultaneously. We hope to offer a more systematic picture of the impact of variables that matter most for development. Indeed we learned that geographic conditions are an important variable preventing higher income equality, and institutions also rather harm a more equal distribution of incomes. Larger trade volumes, in contrast, have the potential to support income convergence. What we do have in common with related studies is the consistency of empirical findings over time. This suggests that the estimates' stability over different time periods gives no occasion to assume that effects substantially change depending on, say, the decade under investigation. As far as data availability allows researchers to go back in time, no time period warrants special attention.

Higher trade integration is often perceived skeptically, as especially the poor are thought to be the "globalization losers" suffering from foreign competition as discussed in chapter 1. Our conclusions from this chapter, which are strictly based on empirics i.e. on what the data speak, propose rather the opposite. Based on the numerous estimations conducted, a plausible interpretation of coefficients suggests that in a more open economy the poor are less marginalized than before. There are some indications that the rich, however, may face perhaps for the first time real competition in so far monopoly-like structures from which they had previously benefitted. Nonetheless, we would not go that far as to interpret the trade estimates as unconditional advocacy of "tearing down the borders". Our study is limited to econometric modeling, hence the cross-country regressions generally disregard more granular, country-specific circumstances that need to be taken into consideration when discussing trade policies.

The discriminating effects of geographic conditions entail even stronger implications. Addressing the inequalizing consequences of this exogenously given variable arguably comes with no conceivable downside. We documented that geographic conditions seem to be related to disease exposure: Rich income groups are not really affected by hot, tropical climate conditions close to the equator, while effects for the poor are dramatic. Higher mortality rates and lower productivity due to poor health likely in such an environment contribute to a widening income gap of the poor vis-à-vis the rich. Thus, policies targeted at creating a more level playing field in terms of accessible and improved public health and disease control should have a sizeable impact in curbing the discriminating effects of geography. International health organizations have increased their efforts in order to help developing countries, but many government leaders still have considerably more options to implement better health conditions. Ameliorations in sanitary conditions and higher vaccination coverage in geographically disadvantaged regions, for instance, might be a less "visible" success for policy-makers, but our results lend support to the conclusion that these would in fact benefit the poor.

Chapter 3 then shifted the research focus to gender inequality and sought to answer: Does marriage age affect educational gender inequality? This question contributed to the strand of literature which relates cultural customs and traditions to gender gaps in education, but which has not provided international evidence on effects of age of marriage on educational achievements. We also incorporated in our empirical strategy the notion that causality goes from marriage age to education, and offered several ways for mitigating endogeneity issues.

The economic conceptualization originated in the observation that women get married at young age in many countries, and on average also at a younger age than men, i.e. we observe a spousal age gap around the world. Simultaneously marriage age shows robust correlation patterns with education levels. We developed a simple economic framework on how a woman's timing of marriage affects her education. Societal expectations of marriage age signal the timing for child-rearing, as marriage is widely regarded as the primary and still the socially most accepted institution for conceiving children. Anticipated family offspring affects future female labor force participation, since due to deeply rooted societal conventions, wives tend to be responsible for raising children. Men are generally not affected by such family planning. This split of gender roles lowers the economic incentive only for female education, because the educational pay-off for women in the labor market is negatively affected by child-rearing. The timing of marriage itself is related to exogenous socio-cultural customs, which crucially influence individual decisions on marriage. The earlier a woman is expected to get married, the shorter her anticipated pay-off to educational investments such that educational investments are lower for younger marriage age than for older marriage age. We thus hypothesized that for countries where women get married younger, their achieved level of education is likely to be lower.

The theoretical considerations were found to hold in our empirical analyses. To this end, we built a global panel data set from 1980-2010 and examined the effects of marriage age on different measures of educational achievement. The results yielded that the female age at marriage has a theory-consistent and highly significant effect on female education: Each year of marriage postponement for women is associated with a three percentage points higher female completion rate in secondary schooling, and to about three weeks, or 13 percent longer female tertiary education. Recognizing the substantial likelihood to have biased results due to endogeneity issues, we employ fixed effects and different instrumentation strategies: For our core specification we instrument the culture-induced domestic female marriage age with a weighted average of the marriage age in adjacent countries. Alternative instruments established in the literature also confirm the findings though not on consistently significant levels.

In a second step, we investigated the effects of spousal age gaps, i.e. the female *relative* to the male marriage age, and obtained even stronger results. Each additional year between wife and husband, i.e. an increase of the spousal age gap by one year, reduces the female secondary schooling completion rate by 10 percentage points and the time women spend at university by one month. Finally, we employed a quasi difference-in-difference strategy to isolate the differences between women and men regarding marriage age and educational achievement, i.e. we specifically examine spousal age gap effects on educational gaps. The methodological rationale was to eliminate potential confounding factors that affect the level of educational achievements jointly for women and men as we focus only on gender differences. The results were again reaffirming. In summary, there is clear evidence that spousal age gaps affect female education significantly more negatively than male education.

For a number of refinement analyses, we moved from a panel analysis to the most recent cross-section for better data availability. There we found that also gender parity in literacy is significantly negatively affected by younger female marriage age, respectively a larger spousal age gap, whereas a measure of primary schooling attainment and the quality of education are not. We documented inter-generational effects of marriage age as well, i.e. the marriage age pattern of the parental generation also influences the children's educational gender inequality. Furthermore, marriage age impacts even more aspects of female discrimination, namely the rate of teenage pregnancies and female participation in politics. We regard this as corroborating the hypothesis that the causal link from marriage age to educational gender inequality is no accidental empirical finding, but part of a broader robust pattern. Finally we provided strong empirical evidence that marriage age and conventional measures of gender discrimination do not act as substitutes, i.e. the effects cannot be closely replicated by employing commonly used variables of gender discrimination instead of marriage age.

These findings of chapter 3 added to the gender inequality research an international macro perspective on the relationship between marriage age and gender gaps in education. The methodology was also among the few in the related literature that offers a set of approaches to address endogeneity issues. Finally, this chapter to our knowledge was the first to examine the impact of other "regular" gender discrimination variables, i.e. we tackle omitted variable bias and document that marriage age is no substitute to conventional measures of gender discrimination. Our core findings confirm the literature consensus of a positive relationship between marriage age and levels of education. It also reaffirms the causal link between female marriage age and female education that could already be documented on a micro level (Field & Ambrus, 2008; Maertens, 2013). Finally, results from the quasi diff-in-diff specifications employed provided rare evidence that female education is significantly more negatively affected by marriage age than male education.

The marriage age in a country is highly dependent on cultural customs and traditions, but it can also be altered by imposing minimum marriage age legislation. Child and teenage marriage is nowadays outlawed or highly restricted in most countries of the world, which recognizes the fact that for the development of young people, in particular of females, such marriage timing is harmful. Our research lends further support to that notion by providing a new economic explanation on why early female marriage has a negative impact on female education. Informing societies about this economic link needs to be made a priority in order to address educational gender inequality successfully, since powerful forces against marriage age regulations represent an ongoing challenge. Recently, a ruling by the Turkish Constitutional Court stirred a heated debate, as an existing law punishing sexual acts with children under 15 years old was annulled (The Guardian, 2016). Human rights activists now fear this may make child marriage more acceptable again in Turkey, a country that is simultaneously negotiating European Union membership. This reflects that female marriage age is an issue not only for the least developed nations, but for basically all societies around the globe.

Finally, chapter 4 examined gender inequality at birth, a second aspect of gender inequality, by asking: Does Central Europe import the missing women phenomenon? For answering this research question, we focused on sex selection at birth and we selected Germany and Switzerland as two cases studies, which appear suitable due to their large migration influx over recent years. Thus we set out to analyze whether different migrant groups, especially from Balkan and Asian countries that are known for strong son preference, also engage in sex selection at birth in their new environment. This analysis extends the literature on missing women at birth among migrant groups. Specifically, it examined for the first time effects in Central Europe and conducted a detailed analysis of Balkan immigrants, who are not only known to engage in sex-selective practices in their countries of origin, but also belong to the same major race as people in Central Europe such that biological reasons for different sex ratios at birth can arguably be well mitigated. The analysis of potential socio-economic reasons for sex selection at birth based on household surveys has also not been performed similarly in the literature before.

To answer this complex question empirically, we resorted to three different micro data sources. We began by investigating national abortion statistics. These alone could not provide conclusive evidence, but left room for conjectures on how sex selective abortion may exist in Germany and Switzerland based on three major observations: the large absolute number of abortions in each country reflects a general societal acceptance that could facilitate (discrete) sex selection, modern medical sex determination techniques allow for a time window to abort within legal boundaries, and a significant share of abortions happens only after two children or more, suggesting deliberate family (and therefore potentially gender) planning.

For more direct and conclusive evidence, we looked in a second step into several years of national birth registries, going back until 1990 for Switzerland, respectively until 2003 for Germany. Empirical results on the share of newborn males versus females indicated that mothers from the target Balkan and Asian countries, in contrast to native Germans respectively Swiss, tend to maintain a high sex selection at birth in their new environment. The sex ratios at birth of these target groups reach up to 1.08 and above, which cannot be reconciled with biological explanations. However, the same target foreign population groups display between-country fluctuations of their average sex ratio at birth, which somewhat challenges the consistency of results. In any case, the sex ratio at birth of foreigners collectively is in contrast only marginally higher than the native population. Hence, we found that the missing women phenomenon is partly imported to Central Europe, since foreigners from selected countries continue to engage in sex selection at birth, but at modest levels in absolute terms.

We then inspected further country-specific variables provided in the birth data in order to better understand some of the associated sociodemographic dynamics. The German data indirectly provide information on the time parents with (both) foreign citizenship have been resident in the country, i.e. we know for how many years migrants have been exposed to a new and arguably more gender-equal sociocultural environment. Yet we found no significant effects from such cultural assimilation on gender-selective practices of immigrant groups in Germany. In addition, there are no substantial differences in sex selection behavior depending if both parents share the same foreign citizenship, or if the father is German. The Swiss birth registries allowed for an examination of sex selection along birth parities. These indicate a skewed ratio at statistically significant levels for all higher parity births among the target Asian parents, whereas for Balkan migrants the sex ratio at higher parities is overall not elevated, except spiking ratios at fourth birth.

The final analysis based on birth data estimated the number of missing women at birth, i.e. how many more women should actually exist in Germany and Switzerland if no parents would practice sex selection at birth. Bearing in mind the ongoing debate over a correct “unbiased” reference sex ratio at birth, we quantified the missing women at birth using three alternative counterfactuals. The results indicated on the one hand that migrants overall do not lead to missing women, since the sex ratio at birth of all foreigners is nearly in line with that of natives. On the other hand, gender selection at birth of the target Balkan and Asian countries translates into missing women that amount to 1,100-1,600 girls in Germany and Switzerland combined over the respective time periods, independent of the counterfactual applied.

In the last section of chapter 4, we explored underlying reasons for the gender inequality at birth. Focusing on socio-economic determinants, we tested the explanatory power of a set of variables for the boy-birth likelihood in a household. As data source we employed the German socio-economic panel study (SOEP) alone, as the sample size for the comparable Swiss household survey turned out to be too small. Overall findings provided no evidence for systematic patterns between socio-economic determinants and sex selection. This outcome mirrors results from national birth data, which have not yielded evidence for collective sex selection practices either. The gender of children in households in Germany is above all determined by nature, i.e. through a random outcome. Nevertheless, our regressions suggested that the sex of older siblings tends to matter, as the boy-birth likelihood raises significantly if the first child was a girl. Balkan, Chinese and Indian immigrants also increase the odds for a boy, whereas immigrants collectively do not. This is also consistent with the findings from national birth registries. Finally, out of our socio-economic determinants only religiosity came close to being significant, but with diverging effects depending on the group under examination. Among all households in Germany, higher religious beliefs decrease son likelihood, which might be due to a general opposition vis-à-vis abortion among more religious people. For foreigners only, however, more religiosity leads to a higher boy-birth likelihood.

The documented distortions of sex ratios at birth among selected migrant groups and the resulting quantification of missing women at birth is closely in line with related findings in other Western countries (Abrevaya, 2009; Almond & Edlund, 2013; Dubuc & Coleman, 2007; González, 2014). In addition to Asian migrants, we were able to identify similar effects for migrants from the Balkans as well. Still, with the absolute level of female discrimination at birth turning out to be small we consequently lack robust associations between any socio-economic variables and the boy-birth likelihood. In short, we show that there is no systematic prenatal gender inequality in Central Europe.

While the record migration influx of 2015 seems to remain a singular peak, there are good reasons to believe that Central Europe will continue to attract migrants in the coming years and decades. Apart from humanitarian crises due to war and conflict that cause people to migrate, the economic power of countries like Germany and Switzerland together with their unfavorable demographic structure and resulting need for external migration represent further and persistent factors. With regards to migrants, their countries of origin tend to be more and more unpredictable. While five years ago few people would have seen Syrians as the main migration group coming to Germany, recent reports note a sharp migration increase of Chechens and other Caucasian population groups (Bidder & Reimann, 2016). If this trend stabilizes such that a larger Caucasian community develops in Central Europe, it would be highly interesting to examine also their levels of son preference and sex selection behavior. This is because the Caucasus region is yet another part of the world whose societies have been found to display substantially male-biased sex ratios at birth (Guilmoto & Duthé, 2013). For now, however, our conclusion suggests that female discrimination in the form of "imported" prenatal gender inequality does not represent a significant issue in Central Europe.

V. 3. Research Outlook

This dissertation offers a number of potential links for future research endeavors. Our granular analysis in chapter 2 on income inequality was characterized by a strong empirical focus, i.e. we "let the data speak". Unanswered questions remain regarding an economic theory that might be able to explain what we see in the data. Such a theoretical explanation would ideally encapsulate all empirical aspects observed in order to consider the interdependencies between the fundamental development variables we employed. Yet, the construction of a model that is able to link the data findings with economic theory for certain sub-aspects would also be welcome, and perhaps more feasible as starting point. For instance, while the economic literature on trade is truly rich, future contributions on why trade affects different income groups the way we observed it would be a step forward in the inequality literature.

Moreover, our results identified the important role of the state of public health for elevating poor income levels. For building on these findings, an investigation on which public health measures are most beneficial, respectively which ones have a more or less differentiating impact for different income groups would be valuable. Such empirical work would potentially generate better insights when moving to a regional level with similarly developed countries, where one may then compare different health strategies. An alternative would be micro-studies using individual data. These could be also analyzed within a country by comparing for instance the effect of health policies on different income levels for different states or districts. Our work in chapter 2 operates rather on the aggregate level, offering a set of new directions for the "inclusive growth" literature level. We hope future research may take up individual elements for further analysis.

Our modest contribution in chapter 3 on educational gender inequality also leaves a number of questions open for additional research. While it provides support for having and strictly enforcing a minimum marriage age requirement as well as compulsory schooling in order to mitigate the most adverse consequences, further policy recommendations based on our findings are difficult to draw. The cross-country perspective naturally limits the degree of specificity we can have at hand when discussing our findings in a national policy framework. As a further point and applicable to all instrumental variable estimates, the methodological approach rests on a set of assumptions that cannot be conclusively proven. The reader also needs to "believe" to a certain degree in the explanation we propose for isolating the causal effects. While we aimed to provide a careful argument for our instruments chosen, we cannot disprove all potential doubts.

Ideally, follow-up studies would access micro level data from various regions to track married couples with different socio-economic and cultural backgrounds, and thus obtain a more granular understanding of individual education dynamics. These may also apply alternative methods for resolving endogeneity concerns, as successfully shown among others by Field & Ambrus (2008). One valuable source might be the Demographic and Health Surveys (DHS), which collect a number of these variables and have been exploited for related purposes (Garenne, 2014;

Westoff, 2003). In addition, these data contain unmarried individuals which could also allow more refined analyses. In this dissertation, we decided to follow a macro data approach in order to complement the number of already existing micro studies, and more importantly, in order to exploit a data source containing a substantially higher number of countries (the DHS survey is limited to ca. 60 countries and excludes e.g. China, whereas we analyze up to 130 countries). Moreover, we have not found an identification strategy based on micro data that would solve the causality issues between the marriage age and education per se more convincingly than our suggested approach. Hence we believe that also a macro data analysis as presented here merits consideration, but more individual and geographically focused data sets may provide more specific policy recommendations.

Another point of departure is based on the conjecture that the relationship between marriage age and education does not follow a strictly linear pattern, but includes some sort of life-cycle aspect as well. In other words, extrapolating the positive association we found much further would lead to an implausible outcome for very old marriage ages, namely extraordinary high education levels. Hence, there needs to be a point in life after which the marginal benefit from delaying marriage gradually decreases. This is a complex issue which we did not examine in depth with the methodological framework at hand, and which requires more explicit integration of literature strands on life-cycle theories. Lastly, the analyses always considered only the first marriage, yet with divorce rates globally increasing the significance of the first marriage (including the associated age at which groom and bride marry) may alter.

One of the questions remaining open from chapter 4, which is also recurrently raised in the literature, refers to the sustainability of son preference and sex selection over the long term. Put differently, does the extent of sex selection at birth among "affected" population groups decrease over time, and if so when? Our proxy variable of cultural assimilation in Germany yielded no conclusive evidence, which mirrors findings of related studies. Researchers have difficulties to investigate this issue further as the data are in general unsuitable, providing an intra-generational perspective only. But it would be most interesting to learn more about the next generation(s) of, say Asian or Balkan immigrants. Do we still find measurable differences in the sex ratio at birth of their children? Can we perhaps even identify variables that are associated with a more rapid decrease in son preference? This would also help understand if the phenomenon is transient when brought by immigrants and consequently, how much policy makers and other involved groups should be concerned with it. Every girl that is not born only because the parents think it has the "wrong" gender is an unacceptable case, but knowing whether this form of gender inequality lasts over generations would be a key insight to better comprehend and tackle the issue.

Furthermore, our household survey analysis suffered from poor response rates and resulting low sample size. We would welcome a similar effort using a different data set that could be able to resolve these issues and thus provide higher statistical reliability. Identifying what affects the level of son preference among migrant groups, other than culture, remains a key element of the research agenda in this field, and we could only offer some initial evidence.

The different chapters in summary demonstrate that inequality may take quite different forms. Yet we find many points of contact, and consequently the interconnectedness of different types of inequality should be kept in mind. Findings by the International Monetary Fund, for instance, indicate that countries with higher income inequality also tend to have larger gender gaps in health, education, labor market participation, etc. (Gonzales, Jain-Chandra, Kochhar, Newiak, & Zeinullayev, 2015). Singular economic perspectives may help disentangle this multifaceted topic, and a scientific approach may also provide a more factual view in this often emotionally led discourse.

Nonetheless, when striving to better manage global inequalities holistically, the interdependencies of its singular dimensions need to be better understood. We recognize that, with our approach, potential inequality "spillovers" to other parts of economy and society cannot be captured adequately. We treated the different inequality aspects rather in isolation, which was essential for methodology and overall feasibility purposes, and which is a scientific simplification and abstraction shared with many other contributions. Therefore we hope that this dissertation, whose findings could be all based on empirical evidence, offered valuable new perspectives on inequality in income and gender for its readers.

Bibliography

Abrevaya, J. (2009). Are There Missing Girls in the United States? Evidence from Birth Data. *American Economic Journal: Applied Economics*, 1(2), 1–34.

Acemoglu, D. (2003). Cross-country Inequality Trends. *The Economic Journal, 113*(485), 121-149.

Acemoglu, D. (2010). Theory, General Equilibrium, and Political Economy in Development Economics. *Journal of Economic Perspectives, 24*(3), 17–32.

Acemoglu, D., & Robinson, J. A. (2000). Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *Quarterly Journal of Economics, 115*(4), 1167–1199.

Acemoglu, D., Johnson, S., & Robinson, J.A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review, 91*(5), 1369-1401.

Acemoglu, D., Johnson, S., & Robinson, J.A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics, 117*(4), 1231-1294.

Adelman, I., & Robinson, S. (1989). Income distribution and development. In Chenery, H., & Srinivasan, T. N. (Eds.), *Handbook of development economics 2* (949-1003). Berkeley: University of Berkeley Press.

Agarwala, A. N., & Singh, S. P. (1957). *The Economics of Underdevelopment.* New York: Oxford University Press.

Aghion, P., Caroli, E., & Garcia-Penalos, C. (1999). Inequality and economic growth: the perspective of the new growth theories. *Journal of Economic Literature, 37*(4), 1615–1660.

Agnihotri, S. (1996). Juvenile sex ratios in India: a disaggregated analysis. *Economic and Political Weekly, 31*(52), 3369–3382.

Agrawal, S., & Unisa, S. (2007). Discrimination from Conception to Childhood: A Study of Girl Children in Rural Haryana, India. In I. Attané & C. Z. Guilmoto (Eds.), *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia* (247–267). Paris: Committee for International Cooperation in National Research in Demography.

Ahluwalia, M. S., Carter, N. G., & Chenery, H. B. (1979). Growth and Poverty in Developing Countries. *Journal of Development Economics, 6*, 299–341.

Alam, N., van Ginneken, J., & Bosch, A. (2007). Decreases in Male and Female Mortality and Missing Women in Bangladesh. In I. Attané & C. Z. Guilmoto (Eds.), *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia* (161–182). Paris: Committee for International Cooperation in National Research in Demography.

Albouy, D. Y. (2012). The Colonial Origins of Comparative Development: An Empirical Investigation: Comment. *American Economic Review, 102*(6): 3059-3076.

Alcalá, F., & Ciccone, A. (2004). Trade and Productivity. *Quarterly Journal of Economics, 119*(2), 612-645.

Aldashev, G., & Guirkinger, C. (2012). Deadly anchor: Gender bias under Russian colonization of Kazakhstan. *Explorations in Economic History, 49*(4), 399–422.

Alderson, A. S., & Nielsen, F. (2002). Globalization and the Great U-turn: Income Inequality Trends in 16 OECD Countries. *American Journal of Sociology, 107*, 1244–99.

Alesina, A., & Perotti, R. (1996). Income Distribution, Political Instability, and Investment. *European Economic Review, 40*(6), 1203–1228.

Alesina, A., & Rodrik, D. (1994). Distributive Politics and Economic Growth. *Quarterly Journal of Economics, 109*(2), 465–490.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth, 8*(2), 155–194.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *Quarterly Journal of Economics, 128*(2), 469–530.

Alexander, K. L., & Eckland, B. K. (1974). Sex differences in the educational attainment process. *American Sociological Review*, 39(5), 668-682.

Allahbadia, G. N. (2002). The 50 Million Missing Women. *Journal of Assisted Reproduction and Genetics, 19*(9), 411–416.

Almond, D., & Edlund, L. (2008). Son biased sex ratios in the U.S. 2000 census. *Proceedings of the National Academy of Sciences of the United States of America, 105*(15), 5681–5682.

Almond, D., Edlund, L., & Milligan, K. (2013). Son preference and the persistence of culture: evidence from South and East Asian Immigrants to Canada. *Population and Development Review, 39*(1), 75–95.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy, 113*(1), 151-184.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2008). Using selection on observed variables to assess bias from unobservables when evaluating swan-ganz catheterization. *American Economic Review, 98*(2), 345-350.

Alvaredo, F, Atkinson, A. B., Piketty, T., & Saez, E. (2015). *The World Top Incomes Database*. Retrieved from http://topincomes.g-mond.parisschoolofeconomics.eu/.

Ambrosetti, E., Ortensi, L. E., Castagnaro, C., & Attili, M. (2015). Sex Imbalances at Birth in Migratory Context: Evidence from Italy. *Genus, 71*(2-3), 29-51.

Anderberg, D., Hener, T., & Wilson, T. (2014). *Assortative Mating–The Role of Education in Marital Age Gaps.* Retrieved from http://www.eea-esem.com/files/papers/EEA-ESEM/2014/2533/MASTER_agegap_small.pdf

Anderson, S., & Ray, D. (2010). Missing Women: Age and Disease. *Review of Economic Studies, 77*(4), 1262–1300.

Anderson, S., & Ray, D. (2012). The Age Distribution of Missing Women in India. *Economic and Political Weekly, 47*(47-48), 87–95.

Anderson, S., & Ray, D. (2015). *Missing Unmarried Women*. NBER Working Paper No. 21511. Retrieved from http://www.econ.nyu.edu/user/debraj/Papers/AndersonRayUnmarried.pdf

Anemone, R. L. (2011). *Race and human diversity: A biocultural approach.* Upper Saddle River, NJ: Prentice Hall/Pearson.

Angrist, J. (2002). How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation. *Quarterly Journal of Economics, 117*(3), 997–1038.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Angrist, J. D., & Pischke, J.-S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con Out of Econometrics. *Journal of Economic Perspectives, 24*(2), 3–30.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies, 58*(2), 277-297.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*(1), 29-51.

Arnold, F. (1997). Gender preferences for children. *Demographic and Health Surveys Comparative Studies 23.*

Arokiasamy, P. (2004). Regional Patterns of Sex Bias and Excess Female Child Mortality in India. *Population, 59*(6), 833–863.

Arokiasamy, P. (2007). Sex ratio at birth and excess female child mortality in India: trends, differentials and regional patterns. In I. Attané & C. Z. Guilmoto (Eds.), *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia* (49–72). Paris: Committee for International Cooperation in National Research in Demography.

Arokiasamy, P., & Goli, S. (2012). Explaining the Skewed Child Sex Ratio in Rural India: Revisiting the Landholding-Patriarchy Hypothesis. *Economic and Political Weekly, 47*(42), 85–94.

Arrow, K. J. (1973). The Theory of Discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in Labor Markets.* Princeton: Princeton University Press.

Atkinson, A. B. (1997). Bringing Income Distribution in From the Cold. *The Economic Journal, 107*(441), 297–321.

Atkinson, A. B. (2011). The Restoration of Welfare Economics. *American Economic Review, 101*(3), 157–161.

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature, 49*(1), 3–71.

Atkinson, A.B., & Brandolini, A. (2001). Promise and Pitfalls in the Use of 'Secondary' Data-Sets: Income Inequality in OECD Countries as a Case Study. *Journal of Economic Literature*, *39*, 771-800.

Atkinson, A.B., & Brandolini, A. (2009). On data: a case study of the evolution of income inequality across time and across countries. *Cambridge Journal of Economics, 33,* 381-404.

Atkinson, A.B., & Piketty, T. (2007). *Top Incomes over the Twentieth Century.* Oxford and New York: Oxford University Press.

Atkinson, A.B., Piketty, T., & Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature, 49*(1), 3-71.

Attané, I., & Guilmoto, C. Z. (Eds.). (2007). *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia*: Paris: Committee for International Cooperation in National Research in Demography.

Balakrishnan, R., Steinberg, C., & Syed, M. (2013). *The Elusive Quest for Inclusive Growth: Growth, Poverty, and Inequality.* IMF Working Paper WP/13/152. Retrieved from https://www.researchgate.net/profile/Ravi_Balakrishnan/publication/259911776_The_Elusive_Quest_for_Inclusive_Growth_Growth_Poverty_and_Inequality_in_Asia/links/0f31752e820fa869e0000000.pdf

BAMF. (2016). *476.649 Asylanträge im Jahr 2015.* Retrieved from https://www.bamf.de/SharedDocs/Meldungen/DE/2016/201610106-asylgeschaeftsstatistik-dezember.html

Bandiera, O., & Natraj, A. (2013). *Does Gender Inequality Hinder Development and Economic Growth. Evidence and Policy Implications.* World Bank. Retrieved from http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2013/02/21/000158349_20130221100323/Rendered/PDF/wps6369.pdf

Banerjee, A. V., & Duflo, E. (2009). The Experimental Approach to Development Economics. *Annual Review of Economics, 1,* 151-178.

Banerjee, A., Duflo, E., Ghatak, M., & Lafortune, J. (2013). Marry for what: caste and mate selection in modern India. *American Economic Journal: Microeconomics, 5*(2), 33-72.

Banerjee, B., & Knight, J. B. (1985). Caste Discrimination in the Indian Urban Labour Market. *Journal of Development Economics, 17*(1), 277–307.

Banister, J. (2004). Shortage of Girls in China Today. *Journal of Population Research, 21*(1), 19–45.

Banks, C. A., & Arnold, P. (2001). Opinions towards sexual partners with a large age difference. *Marriage & Family Review, 33*(4), 5-18.

Barcellos, S. H., Carvalho, L. S., & Lleras-Muney, A. (2014). Child Gender and Parental Investments in India: Are Boys and Girls Treated Differently?. *American Economic Journal: Applied Economics, 6*(1), 157-189.

Bardhan, P. (1996). *The Nature of Institutional Impediments to Economic Development.* Center for International and Development Economics Research. Retrieved from http://escholarship.org/uc/item/8w12j7q1

Barnard, H. C. (1969). *Education and the French revolution*. Cambridge: Cambridge University Press.

Barro, R. J., & Sala-i-Martin, X. (1995) *Economic Growth.* New York: McGraw-Hill.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics, 106*(2), 407–443.

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth, 5*(1), 5–32.

Barro, R. J. (2001). Human Capital and Growth. *American Economic Review, (91)*2, 12-17.

Barro, R. J. (2008). *Inequality and Growth Re-Visited.* ADB Working Paper on Regional Economic Integration No. 11. Retrieved from https://adb-test.atmire.com/bitstream/handle/11540/1762/wp11-inequality-growth-revisited.pdf?sequence=1

Barro, R. J., & Lee, J. W. (1994). Sources of economic growth. *In Carnegie-Rochester conference series on public policy, 40*, 1-46.

Barro, R. J., & Lee, J. W. (2001). International Data on Educational Attainment: Updates and Implications. *Oxford Economic Papers, 53*(3), 541-563.

Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics, 104*, 184-198.

Basu, A. M. (1992). *Culture, the Status of Women, and Demographic Behavior.* Oxford: Oxford University Press.

Basu, A. M., & Das Gupta, M. (2001). Family Systems and the Preferred Sex of Children. In J. Hoem (Ed.), *International Encyclopedia of Social and Behavioral Sciences.*

Basu, A.M. (2002). Why does education lead to lower fertility? A critical review of some of the possibilities. *World Development 30*(10), 1779–1790.

Baum, C., Schaffer, M. E., & Stillman, S. (2003). Instrumental variables and GMM: Estimation and testing. *Stata Journal, 3*(1): 1-31.

Baum, C., Schaffer, M. E., & Stillman, S. (2007). Enhanced Routines for Instrumental Variables/GMM Estimation and Testing. *Stata Journal, 7*(4): 465–506.

Bazzi, S., & Clemens, M.A. (2013) Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth. *American Economic Journal: Macroeconomics, 5*(2), 152-186.

Becker, G. S. (1971). *The economics of discrimination.* Chicago: University of Chicago Press.

Becker, G. S. (1973). A theory of marriage: Part I. *The Journal of Political Economy, 81*, 813-846.

Becker, G. S. (1974). A theory of marriage: Part II. *The Journal of Political Economy*, *82*, 11-26.

Becker, S. O., Cinnirella, F., & Woessmann, L. (2010). The trade-off between fertility and education: evidence from before the demographic transition. *Journal of Economic Growth, 15*(3), 177-204.

Benabou, R. (1996). *Inequality and Growth.* NBER Macroeconomics Annual, 11–73.

Benería, L. (1995). Toward a greater integration of gender in economics. *World Development, 23*(11), 1839–1850.

Benería, L., & Sen, G. (1981). Accumulation, Reproduction, and "Women's Role in Economic Development": Boserup Revisited. *Signs, 7*(2), 279–298.

Benería, L., Berik, G., & Floro, M. (2015). *Gender, Development and Globalization: Economics as If All People Mattered.* New York: Taylor & Francis.

Benhabib, J., & Spiegel, MM. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics, 34*(2), 143-173.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy, 75*(4, Part 1), 352–365.

Berman, E., Bound, J., & Griliches, Z. (1994). Changes in the Demand for Skilled Labor within U. S. Manufacturing: Evidence from the Annual Survey of Manufactures. *Quarterly Journal of Economics, 109*(2), 367–397.

Bertola, G. (1993). Factor shares and savings in endogenous growth. *American Economic Review, 83*, 1184-1198.

BFS. (2015a). *Anzahl Schwangerschaftsabbrüche. Informationszentrum, Sektion Demografie und Migration.* Retrieved from http://www.bfs.admin.ch/

BFS. (2015b). B*ilanz der ständigen Wohnbevölkerung nach Staatsangehörigkeit, 2014.* Informationszentrum, Sektion Demografie und Migration. Retrieved from http://www.bfs.admin.ch/

BFS. (2015c). *Lebendgeburten nach detaillierter Staatsangehörigkeit des Kindes, 1970-2014.* Informationszentrum, Sektion Demografie und Migration. Retrieved from http://www.bfs.admin.ch/

Bharadwaj, P., & Lakdawala, L. K. (2013). Discrimination Begins In The Womb: Evidence Of Sex-Selective Prenatal Investments. *Journal of Human Resources, 48*(1), 71-113.

Bhasin, K. (1993). *What is Patriarchy?* New Delhi: Kali for Women.

Bhaskar, V., & Gupta, B. (2007). India's missing girls: biology, customs, and economic development. *Oxford Review of Economic Policy, 23*(2), 221–238.

Bhat, P. M. (2002). On the Trail of 'Missing' Indian Females: II: Illusion and Reality. *Economic and Political Weekly, 37*(52), 5244–5263.

Bhat, P. M., & Zavier, A. J. F. (2007). Factors Influencing the Use of Prenatal Diagnostic Techniques and the Sex Ratio at Birth in India. *Economic and Political Weekly, 42*(24), 2292–2303.

Bhrolcháin, M. N. (1992). Age difference asymmetry and a two-sex perspective. *European Journal of Population/Revue européenne de Démographie, 8*(1), 23-45.

Bianchi, S. M., Milkie, M. A., Sayer, L. C., & Robinson, J. P. (2000). Is anyone doing the housework? Trends in the gender division of household labor. *Social Forces, 79,* 191–234.

Bidder, B., & Reimann, A. (2016). *Neue Fluchtbewegung: Darum kommen so viele Tschetschenen nach Deutschland.* Retrieved from http://www.spiegel.de/politik/ausland/fluechtlinge-darum-kommen-viele-tschetschenen-nach-deutschland-a-1099180.html

Bigsten, A. (1983). *Income Distribution and Development: Theory, Evidence, and Policy.* London: Heinemann.

Blanchard, O., & Fischer, S. (1989). *Lectures on macroeconomics.* Cambridge, London: MIT Press.

Blanden, J., Gregg, P., & Machin, S. (2005). *Intergenerational Mobility in Europe and North America*. London: Centre for Economic Performance.

Blangiardo, G. C., & Rimoldi, S. (2012*). 100 milioni di bambine mancano all'appello nel mondo. E in Italia?*. Neodemos 24/10/2012. Retrieved from http://www.neodemos.info/100-milioni-di-bambine-mancano-allappello-nel-mondo-e-in-italia/

Blood, R. O., & Wolfe, D. M. (1960) *Husbands and Wives.* New York: The Free Press.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87,* 115-143.

Bongaarts, J. (2013). The implementation of preferences for male offspring. *Population and Development Review, 39*(2), 185–208.

Bongaarts, J., & Guilmoto, C. Z. (2015). How Many More Missing Women? Excess Female Mortality and Prenatal Sex Selection, 1970–2050. *Population and Development Review, 41*(2), 241–269.

Borooah, V. K., Diwakar, D., Mishra, V. K., Naik, A. K., & Sabharwal, N. S. (2014). Caste, inequality, and poverty in India: A re-assessment. *Development Studies Research, 1*(1), 279–294.

Bose, S. (2011). The Effect of Women's Status and Community on the Gender Differential in Children's Nutrition in India. *Journal of Biosocial Science*, *43*(5), 513–533.

Boserup, E. (1970). *Woman's Role in Economic Development*. London: George Allen and Unwin Ltd.

Bourguignon, F. (1981). Pareto Superiority of Unegalitarian Equilibria in Stiglitz' Model of Wealth Distribution with Convex Saving Function. *Econometrica, 49*(6), 1469–1475.

Bourguignon, F., & Morrisson, C. (2002). Inequality among World Citizens: 1820 – 2002. *American Economic Review, 92*(4), 727-744.

Bourguignon, F., & Verdier, T. (2000). Oligarchy, Democracy, Inequality, and Growth. *Journal of Development Economics, 62*(2), 285–313.

Brandth, B., & Kvande, E. (2016). Fathers and flexible parental leave. *Work, Employment & Society*, *30*(2), 275-290.

Bruno, M., Ravallion, M., & Squire, L. (1998). Equity and growth in developing countries: old and new perspectives on the policy issues. In Tanzi, V. & Chu, K. (Eds.), *Income Distribution and High-Quality Growth*. Cambridge: MIT Press.

Bulte, E., Heerink, N., & Zhang, X. (2011). China's One-Child Policy and 'the Mystery of Missing Women': Ethnic Minorities and Male-Biased Sex Ratios. *Oxford Bulletin of Economics and Statistics, 73*(1), 21–39.

Cagatay, N., & Ertürk, K. (2004). *Gender and Globalization: a Macroeconomic Perspective*. Retrieved from http://www.berlin-divercity.de/diwiki/images/1/17/Gender_and_globalisation.pdf

Cagatay, N., Elson, D., & Grown, C. (1995). Introduction. *World Development Special Issue on Gender, Adjustment and Macroeconomics, 23*(11).

Cain, G. C. (1986). The economic analysis of labor market discrimination: A survey. In Ashenfelter, O. C., & Layard, R. (Eds.), *Handbook of Labor Economics, vol. 1* (709-730). North Holland.

Caldwell, J. C., Reddy, P. H., & Caldwell, C. (1983). The causes of marriage change in South India. *Population Studies, 37*(3), 343-361.

Calvi, R. (2015). *Why Are Older Women Missing in India? The Age Profile of Bargaining Power and Poverty*. Retrieved from https://www.bc.edu/content/dam/files/schools/cas_sites/economics/pdf/Seminars/SemF2015/CALVI_JMP_Nov28.pdf

Campbell, E. K., & Campbell, P. G. (1997). Family Size and Sex Preferences and Eventual Fertility in Botswana. *Journal of Biosocial Science, 29*(2), 191–204.

Canberra Group (2001). *Final Report and Recommendations.* Ottawa: Expert Group on Household Income Statistics, The Canberra group. Retrieved from http://www.lisdatacenter.org/wp-content/uploads/canberra_report.pdf

Carmichael, S. (2011). Marriage and power: Age at first marriage and spousal age gap in lesser developed countries. *The History of the Family, 16*(4), 416-436.

Carmichael, S., De Moor, T., & van Zanden, J. L. (2011). *"When the heart is baked, don't try to knead it." Marriage age and spousal age gap as a measure of female 'agency'.* CGEH Working Paper 19, Centre for Global Economic History. Retrieved from http://www.cgeh.nl/sites/default/files/WorkingPapers/CGEH.WP_.No19.Carmichaeletal.pdf

Carranza, E. (2014). Soil Endowments, Female Labor Force Participation, and the Demographic Deficit of Women in India. *American Economic Journal: Applied Economics, 6*(4), 197–225.

Castelló-Climent, A., & Hidalgo-Cabrillana, A. (2012). The Role of Educational Quality and Quantity in the Process of Economic Development. *Economics of Education Review, 31*, 391-409.

Casterline, J. B., Williams, L., & McDonald, P. (1986). The age difference between spouses: variations among developing countries. *Population Studies, 40*(3), 353-374.

Cervellati, M., & Sunde, U. (2005). Human capital formation, life expectancy, and the process of development. *American Economic Review, 95*(5), 1653-1672.

Cervellati, M., & Sunde, U. (2007). *Human Capital, Mortality and Fertility: A Unified Theory of the Economic and Demographic Transition*. IZA Discussion Paper, 2905. Retrieved from ftp://repec.iza.org/RePEc/Discussionpaper/dp2905.pdf

Chahnazarian, A. (1988). Determinants of the sex ratio at birth: Review of recent literature. *Social biology, 35*(3-4), 214–235.

Chamarbagwala, R., & Ranger, M. (2006). *India's Missing Women: Disentangling Cultural, Political and Economic Variables* (Working Paper). Indiana University, Kelley School of Business, Department of Business. Retrieved from https://kelley.iu.edu/riharbau/repec/iuk/wpaper/bepp2006-05-chamarbagwala-ranger.pdf

Checchi, D., & García-Peñalosa, C. (2008). Labour market institutions and income inequality. *Economic Policy, 23*(56), 602-649.

Chiappori, P. A., Iyigun, M., & Weiss, Y. (2009). Investment in Schooling and the Marriage Market. *American Economic Review, 99*(5), 1689-1714.

Chung, W., & Das Gupta, M. (2007). The Decline of Son Preference in South Korea: The Roles of Development and Public Policy. *Population and Development Review, 33*(4), 757–783.

Ciccone, A., & Jarociński, M. (2010). Determinants of Economic Growth: Will Data Tell? *American Economic Journal: Macroeconomics, 2*(4), 222-246.

Clark, J. B. (1891). Distribution as Determined by a Law of Rent. *Quarterly Journal of Economics, 5*(3), 289–318.

Coale, A. J. (1991). Excess Female Mortality and the Balance of the Sexes in the Population: An Estimate of the Number of "Missing Females". *Population and Development Review, 17*(3), 517–523.

Cohen, D., & Soto, M. (2007). Growth and human capital: good data, good results. *Journal of Economic Growth, 12*(1), 51-76.

Coley, R. L., & Chase-Lansdale, P. L. (1998). Adolescent pregnancy and parenthood: Recent evidence and future directions. *American Psychologist, 53*(2), 152–166.

Cordasco, F. (1976). *A brief history of education: a handbook of information on Greek, Roman, medieval, Renaissance, and modern educational practice.* Totowa, NJ: Rowman & Littlefield.

Correa, E. A., Jetter, M., & Agudelo, A. M. (2016). Corruption: Transcending borders. *Kyklos, 69*(2), 183-207.

Cragg, J. G., & Donald, S. G. (1993). Testing Identifiability and Specification in Instrumental Variable Models. *Econometric Theory, 9*(2), 222–240.

Croll, E. J. (2000). *Endangered Daughters*. London: Routledge.

Croll, E. J. (2001). Amartya Sen's 100 Million Missing Women. *Oxford Development Studies, 29*(3), 225–244.

Dahl, G. B., & Moretti, E. (2008). The demand for sons. *The Review of Economic Studies, 75*(4), 1085–1120.

Danziger, L., & Neuman, S. (1999). On the age at marriage: theory and evidence from Jews and Moslems in Israel. *Journal of Economic Behavior & Organization, 40*(2), 179-193.

Das Gupta, M. (1987). Selective Discrimination against Female Children in Rural Punjab. *Population and Development Review, 13*(1), 77–100.

Das Gupta, M. (2000). Population and Resources: An Exploration of Reproductive and Environmental Externalities. *Population and Development Review, 26*(4), 643–689.

Das Gupta, M. (2005). Explaining Asia's "Missing Women": A New Look at the Data. *Population and Development Review, 31*(3), 529–535.

Das Gupta, M. (2010). Family Systems, Political Systems and Asia's Missing Girls: The Construction of Son Preference and Its Unravelling. *Asian Population Studies, 6*(2), 123–152.

Das Gupta, M., & Bhat, P. M. (1997). Fertility Decline and Increased Manifestation of Sex Bias in India. *Population Studies, 51*(3), 307–315.

Das Gupta, M., & Shuzhuo, L. (1999). Gender Bias in China, South Korea and India 1920-1990: Effects of War, Famine and Fertility Decline. *Development and Change, 30*(3), 619–652.

Das Gupta, M., Chung, W., & Shuzhuo, L. (2009). Evidence for an Incipient Decline in Numbers of Missing Girls in China and India. *Population and Development Review, 35*(2), 401–416.

Das Gupta, M., Zhenghua, J., Bohua, L., Zhenming, X., Chung, W., & Hwa-Ok, B. (2003). Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea. *Journal of Development Studies, 40*(2), 153–187.

David, H. P. (1970). *Family Planning and Abortion in the Socialist Countries of Central and Eastern Europe*. New York: Population Council.

Davies, J.B., Shorrocks, A., Sandstrom, S., & Wolff, E. (2007). *The World Distribution of Household Wealth. Center for Global, International and Regional Studies.* Retrieved from http://escholarship.org/uc/item/3jv048hx

De Gregorio, J., & Lee, J. W. (2002). Education and income inequality: new evidence from cross-country data. *Review of income and wealth, 48*(3), 395-416.

Dearden, N. (2016, May 20). *TTIP symbolises the worst of global capitalism. Cameron pushes it at his peril. The Guardian.* Retrieved from http://www.theguardian.com/commentisfree/2016/may/20/david-cameron-support-ttip

Deaton, A. S. (2010). Instruments, Randomization, and Learning about Development. *Journal of Economic Literature, 48*(2), 424–455.

Deaton, A., & Drèze, J. (2009). Food and Nutrition in India: Facts and Interpretations. *Economic and Political Weekly, 44*(7), 42–65.

Deininger, K., & Squire, L. (1996). A New Data Set Measuring Income Inequality. *World Bank Economic Review, 10*(3), 565-591.

Deininger, K., & Squire, L. (1998). New Way of Looking at Old Issues: Inequality and Growth. *Journal of Development Economics, 57*(2), 259–287.

Deshpande, A. (2011). *The grammar of caste: Economic discrimination in contemporary India.* Oxford: Oxford University Press.

Devroye, D., & Freeman, R. B. (2001). *Does inequality in skills explain inequality in earnings across advanced countries?* NBER Working Paper No. 8140. Retrieved from http://www.nber.org/papers/w8140.pdf

Diamond, J. (1997). *Guns, Germs, and Steel: the fates of human societies*. W.W. Norton, New York, NY.

Díaz-Giménez, J., & Giolito, E. (2013). Accounting for the Timing of First Marriage. *International Economic Review, 54*(1), 135-158.

DiMaggio, P. (1994). Culture and economy. In N. J. Smelser & R. Swedberg (Eds.) *The Handbook Of Economic Sociology* (27-57). New York and Princeton, NJ: Princeton University Press.

Doepke, M., & Tertilt, M. (2009). Women's Liberation: What's in It for Men? *Quarterly Journal of Economics, 124*(4), 1541–1591.

Doepke, M., Tertilt, M., & Voena, A. (2012). The Economics and Politics of Women's Rights. *Annual Review of Economics, 4*, 339–372.

Dollar, D., & Gatti, R. (1999) *Gender Inequality, Income and Growth: Are Good Times Good for Women?.* Washington DC: World Bank.

Dollar, D., & Kraay, A. (2002). Growth is Good for the Poor. *Journal of Economic Growth, 7*, 195-225.

Dollar, D., & Kraay, A. (2003), Institutions, Trade and Growth. *Journal of Monetary Economics, 50*(1), 133-162.

Dollar, D., & Kraay, A. (2004). Trade, growth, and poverty. *The Economic Journal, 114*(493), 22-49.

Dollar, D., Fisman, R., & Gatti, R. (2001). Are women really the "fairer" sex? Corruption and women in government. *Journal of Economic Behavior & Organization, 46*(4), 423-429.

Dollar, D., Kleineberg, T., & Kraay, A. (2014). *Growth, Inequality, and Social Welfare. Cross-Country Evidence*. World Bank Policy Research Working Paper No. 6842.

Dollar, D., Kleineberg, T., & Kraay, A. (2016). Growth Still is Good for the Poor. *European Economic Review, 81*, 68-85.

Dougherty, C. (2005). Why are the returns to schooling higher for women than for men?. *Journal of Human Resources, 40*(4), 969-988.

Drèze, J., & Sen, A. (1995). *India: Economic Development and Social Opportunity*. Oxford: Oxford University Press.

Dubuc, S., & Coleman, D. (2007). An Increase in the Sex Ratio of Births to India-born Mothers in England and Wales: Evidence for Sex-Selective Abortion. *Population and Development Review, 33*(2), 383–400.

Duflo, E. (2005). *Gender Equality in Development*. BREAD Policy Paper No. 011, Duke University.

Duflo, E. (2012). Women Empowerment and Economic Development. *Journal of Economic Literature, 50*(4), 1051-79.

Durham, M. E. (1909). *High Albania*. London: Edward Arnold.

Dykstra, S., Dykstra, B., & Sandefur, J. (2014). *We Just Ran Twenty-Three Million Queries of the World Bank's Web Site*. CGD Working Paper 362.

Dyson, T. (2012). Causes and Consequences of Skewed Sex Ratios. *Annual Review of Sociology, 38*, 443–461.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics, 84*(2), 755-776.

Easterly, W., & Levine, R. (2003). Tropics, germs, and crops: how endowments influence economic development. *Journal of monetary economics, 50*(1), 3-39.

Easterly, W., & Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics*, *112*(4), 1203–1250.

Ebenstein, A. (2014). *Patrilocality and Missing Women* (Working Paper). Retrieved from https://scholars.huji.ac.il/avrahamebenstein/patrilocality-and-missing-women

Eberhardt, M., & Teal, F. (2011). Econometrics for grumblers: A new look at the literature on cross-country growth empirics. *Journal of Economic Surveys, 25*, 109–155.

Echávarri, R., & Husillos, J. (2016). The Missing Link Between Parents' Preferences and Daughters' Survival: The Moderator Effect of Societal Discrimination. *World Development, 78,* 372–385.

Edlund, L. (1999). Son preference, sex ratios, and marriage patterns. *Journal of Political Economy*, *107*(6), 1275-1304.

Edlund, L., & Pande, R. (2002). Why have women become left-wing? The political gender gap and the decline in marriage. *Quarterly Journal of Economics, 117*, 917-961.

Efrat, Z., Akinfenwa, O. O., & Nicolaides, K. H. (1999). First-trimester determination of fetal gender by ultrasound. *Ultrasound Obstet Gynecol, 13,* 305–307.

Eguavoen, A. N., Odiagbe, S. O., & Obetoh, G. I. (2007). The status of women, sex preference, decision making and fertility control in Ekpoma community of Nigeria. *Journal of Social Science*, *15*(1), 43–49.

Elson, D. (1991). Male Bias in the Development: An Overview. In D. Elson (Ed.), *Male Bias in the Development Process: An Overview*. Manchester: Manchester University Press.

Elson, D. (1999). Labor Markets as Gendered Institutions: Equality, Efficiency, and Empowerment Issues. *World Development*, *27*(3), 611.

Elson, D., & Cagatay, N. (2000). The Social Content of Macroeconomic Policies. *World Development, 28*, 1347–1364.

Engermann, S., & Sokoloff, K. (1997). Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States. In S. Haber (Ed). *How Latin America Fell Behind.* Stanford CA: Stanford University Press.

Esteve-Volvart, B. (2000). *Sex Discrimination and Growth*. IMF Working Paper No. 84, International Monetary Fund.

Esteve-Volvart, B. (2004). Gend*er Discrimination and Growth: Theory and Evidence from India*. DEDPS Working Paper 42, London: London School of Economics and Political Science.

Eswaran, M. (2014). *Why Gender Matters in Economics*. Princeton: Princeton University Press. Retrieved from https://books.google.co.uk/books?id=ZD6OAwAAQBAJ

Federal Foreign Office. (2015). *New provisions regulating choice of citizenship.* Retrieved from http://www.auswaertiges-amt.de/EN/EinreiseUndAufenthalt/Optionsregelung.html?searchArchive=0&searchEngineQueryString=citizenship&searchIssued=0&searchIssuedAfter=27.11.2013

Federal Statistical Office. (2015a). *Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund – Ergebnisse des Mikrozensus – 2014, Wiesbaden.* Retrieved from https://www.destatis.de/DE/Startseite.html

Federal Statistical Office. (2015b). *Gesundheit. Schwangerschaftsabbrüche, Wiesbaden.* Retrieved from https://www.destatis.de/DE/Startseite.html

Federal Statistical Office. (2015c). *Lebendgeborene 2003-2014 nach der Staatsangehörigkeit und dem Geschlecht des Kindes sowie nach der Staatsangehörigkeit der Eltern, Wiesbaden*. Retrieved from https://www.destatis.de/DE/Startseite.html

Federal Statistical Office. (2016). *7 - Sozialstruktur und soziale Lagen*. Retrieved from https://www.destatis.de/DE/Publikationen/Datenreport/Downloads/Datenreport2016Kap7.pdf;jsessionid=5B569D87A27F8EA5D0E72163D5601708.cae3?__blob=publicationFile

Feenstra, R. C., & Hanson, G. H. (1996). Foreign Investment, Outsourcing, and Relative Wages. In R.C. Feenstra, G.M. Grossman & D.A. Irwin (Eds.), *The Political Economy of Trade Policy: Papers in Honor of Jagdish Bhagwati* (89-127). Cambridge, MA: MIT Press.

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review, 105*(10), 3150-3182. Data available for download at www.ggdc.net/pwt

Fernández, R., Guner, N., & Knowles, J. (2005). Love and money: A theoretical and empirical analysis of household sorting and inequality. *Quarterly Journal of Economics, 120*(1), 273-344 .

Feyrer, James. 2009. *Trade and Income – Exploiting Time Series in Geography.* NBER Working Paper No. 14910.

Field, E., & Ambrus, A. (2008). Early marriage, age of menarche, and female schooling attainment in Bangladesh. *Journal of Political Economy, 116*(5), 881-930.

Filmer, D., Friedman, J., & Schady, N. (2009). Development, Modernization, and Childbearing: The Role of Family Sex Composition. *World Bank Economic Review, 23*(3), 371–398.

Fischer, R.D. (2001). The evolution of inequality after trade liberalization. *Journal of Development Economics, 66*, 555-579.

Foreman-Peck, J. S. (2011). The Western European marriage pattern and economic development. *Explorations in Economic History*, *48*(2), 292-309.

Foster, A., & Rosenzweig, M. (2001*). Missing women, the marriage market and economic growth.* Unpublished manuscript. Retrieved from http://www.cid.harvard.edu/archive/events/cidneudc/papers/sex.pdf

Frankel, J.A., & Rose, A. (2002). An Estimate of the Effect of Common Currencies on Trade and Income. *Quarterly Journal of Economics, 117*(2), 437-466.

Frankel, J.A., & Romer, D. (1999). Does Trade Cause Growth? *American Economic Review, 89*(3), 379-399.

Freud, S. (1962). *Three Essays on the Theory of Sexuality*. New York: Avon Books.

Fryer, R.G., & Levitt, S.D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics, 2*(2), 210–240.

Fukuda, M., Fukuda, K., Shimizu, T., & Moller, H. (1998). Decline in sex ratio at birth after Kobe earthquake. *Human reproduction*, *13*(8), 2321–2322.

Gallo, C. (2002). *Economic Growth and Income Inequality: Theoretical Background and Empirical Evidence.* Working Paper No. 119. University College London. Retrieved from https://www.bartlett.ucl.ac.uk/dpu/publications/latest/publications/dpu-working-papers/WP119.pdf

Gallup, J.L., Sachs, J.D. (2001). The economic burden of malaria. *The American journal of tropical medicine and hygiene*, *64*(1), 85-96.

Gallup, J.L., Sachs, J.D., Mellinger, A. (1999). Geography and economic development. *International Regional Science Review, 22*(2), 179-232.

Galor, O. (2009). *Inequality and Economic Development: An Overview*. Brown University. Retrieved from https://core.ac.uk/download/files/153/7067349.pdf

Galor, O., & Moav, O. (2004). From Physical to Human Capital Accumulation: Inequality and the Process of Development. *Review of Economic Studies*, *71*(7), 1001–1026.

Galor, O., & Tsiddon, D. (1997). Technological Progress, Mobility, and Economic Growth. *American Economic Review, 87*, 363–382.

Galor, O., & Weil, D. N. (1996). The Gender Gap, Fertility, and Growth. *American Economic Review*, *86*(2), 374–387.

Garenne, M. (2004). Age at marriage and modernisation in sub-Saharan Africa. *Southern African Journal of Demography, 9*(2), 59-79.

Garenne, M. (2014). *Trends in marriage and contraception in sub-Saharan Africa: A longitudinal perspective on factors of fertility decline.* DHS Analytical Studies No. 42. ICF Internation, Rockville, Maryland, USA. Retrieved from http://dhsprogram.com/pubs/pdf/AS42/AS42.pdf.

Gavalas, V., Rontos, K., & Nagopoulos, N. (2015). Sex ratio at birth in twenty-first century Greece: the role of ethnic and social groups. *Journal of Biosocial Science*, *47*(3), 363-375.

Geeta, V. (2007). *Patriarchy*. Kolkata: Stree.

George, H. (1960 [1879]). *Progress and Poverty: An Inquiry Into the Cause of Industrial Depressions, and of Increase of Want with Increase of Wealth…the Remedy*. New York: Robert Schalkenbach Foundation.

Gillard, H., Howcroft, D., Mitev, N., & Richardson, H. (2008). “Missing women”: Gender, ICTs, and the shaping of the global economy. *Information Technology for Development, 14*(4), 262–279.

Gimbutas, M. (2007). *The Goddesses and Gods of Old Europe: Myths and Cult Images.* Berkeley: University of California Press.

Gjonca, A. (2004). The Crossroads of Europe: A first look at regional adult mortality differences in the Balkans. *Space population societies, 3,* 665–679.

Gjonca, A., Aasve, A., & Mencarini, L. (2008). Albania: Trends and patterns, proximate determinants and policies of fertility change. *Demographic Research, 19*(11), 261–292.

Gjonca, A., Aasve, A., & Mencarini, L. (2009). The Highest Fertility in Europe - for how long? Determinants of Fertility Change in Albania. *Demográfia*, *52*(5), 76–96.

Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2004). Do institutions cause growth?. *Journal of Economic Growth*, *9*(3), 271-303.

Glick, P., & Sahn, D. E. (2000). Schooling of girls and boys in a West African country: the effects of parental education, income, and household structure. *Economics of Education Review*, *19*(1), 63-87.

Glick, P., Handy, C., & Sahn, D. E. (2015). Schooling, Marriage and Age of First Birth in Madagascar. *Population Studies*, *69,* 219-236.

Goldberg, P. K., & Pavcnik, N. (2007). Distributional Effects of Globalization in Developing Countries. *Journal of Economic Literature, 45,* 39–82.

Goldin C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review, 96*(2), 1-21.

Goldin C., & Katz, L. F. (2002). The power of the pill: oral contraceptives and women's career and marriage decisions. *Journal of Political Economy, 110*(4), 730-770.

Goldin C., Katz L.F., & Kuziemko, I. (2006). The homecoming of American college women: the reversal of the college gender gap. *Journal of Economic Perspectives*, *20*(4), 133-156.

Gonzales, C., Jain-Chandra, S., Kochhar, K., Newiak, M., & Zeinullayev, T. (2015). *Catalyst for Change: Empowering Women and Tackling Income Inequality*. IMF. Retrieved from http://www.imf.org/external/pubs/ft/sdn/2015/sdn1520.pdf

González, L. (2014). *Missing Girls in Spain* (Working Paper). Barcelona Graduate School of Economics. Retrieved from http://www.unav.edu/matrimonioyfamilia/observatorio/uploads/33214_BarcelonaGse-Gonzalez_Missing-girls-2014.pdf

Goodman, A., Johnson, P., & Webb, S. (1997). *Inequality in the UK*. New York: Oxford University Press.

Gradstein, M. (2007). Inequality, Democracy, and the Protection of Property Rights. *Economic Journal, 117*(1), 252–269.

Graham, E. (2007). Son Preference, Female Deficit and Singapore's Fertility Transition. In I. Attané & C. Z. Guilmoto (Eds.), *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia* (89–106). Paris: Committee for International Cooperation in National Research in Demography.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica*, *45*, 1-22.

Grossbard-Shechtman, S. (1993). *On the Economics of Marriage: A Theory of Marriage, Labor, and Divorce*. Boulder, CO: Westview Press.

Grossman, V., & Stadelmann, D. (2013). Wage Effects of High-skilled Migration: International Evidence, *World Bank Economic Review, 27*(2), 297-319.

Guilmoto, C. Z. (2008). Economic, Social and Spatial Dimensions of India's Excess Child Masculinity. *Population, 63*(1), 91–117.

Guilmoto, C. Z. (2009). The sex ratio transition in Asia. *Population and Development Review, 35*(3), 519–549.

Guilmoto, C. Z. (2010). *High Sex Ratios at Birth in Southeast Europe*. Retrieved from http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers15-05/010063779.pdf

Guilmoto, C. Z. (2012). Son Preference, Sex Selection, and Kinship in Vietnam. *Population and Development Review, 38*(1), 31–54.

Guilmoto, C. Z. (2015). The Masculinization of Births. Overview and Current Knowledge. *Population, 70*, 201–264.

Guilmoto, C. Z., & Duthé, G. (2013). Masculinization of births in Eastern Europe. *Population & Societies, 506*, 1–4.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science, 320*, 1164-1165.

Gustafson, P., & Fransson, U. (2015). Age Differences Between Spouses: Sociodemographic Variation and Selection. *Marriage & Family Review, 51*(7), 610-632.

Hajnal, J. (1965). European Marriage Patterns in Perspective. In D. V. Glass & D.E.C. Eversley (Eds.), *Population in History* (101-143). Aldine, Chicago.

Hall, D. (1994). *Albania and the Albanians*. London: Pinter.

Hall, R., & Jones, C.I. (1999). Why Do Some Countries Produce So Much More per Worker than Others? *Quarterly Journal of Economics, 114*, 83-116.

Hansen, D., Moller, H., & Olsen, J. (1999). Severe periconceptional life events and the sex ratio in offspring: follow up study based on five national registers. *BMJ, 319*(7209), 548–549.

Hanushek E.A., & Woessmann, L. (2009). *Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation*. NBER Working Paper No. 14633.

Hanushek, E. A. (2015). Why Standard Measures of Human Capital are Misleading. *Journal of Economic Policy, 37*(2), 22-39.

Hanushek, E. A., & Woessmann, L. (2011). The economics of international differences in educational achievement. In E. A. Hanushek, S. Machin, L. Woessmann, L. (Eds.), *Handbook of the Economics of Education* (89-200). Amsterdam, London: North Holland.

Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth, 17*(4), 267-321.

Harrison, B., & Bluestone, B. (1988). *The Great U-Turn*. New York: Basic Books.

Hausmann, R., Rodrik, D., & Velasco, A. (2006). Growth Diagnostics. In J. Stiglitz & N. Serra (Eds.), *The washington consensus reconsidered: Towards a new global governance* (324-355). New York: Oxford University Press.

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In E. A. Hanushek, S. Machin, L. Woessmann, L. (Eds.), *Handbook of the Economics of Education* (307–458). Amsterdam, London: North-Holland.

Helpman, E. (1997). *General Purpose Technologies and Economic Growth.* Cambridge, MA: MIT Press.

Heshmati, A. (2006). The world distribution of income and income inequality: a review of the economics literature. *Journal of World-Systems Research, 12*(1), 60-107.

Hesketh, T., & Xing, Z. W. (2006). Abnormal sex ratios in human populations: causes and consequences. *Proceedings of the National Academy of Sciences of the United States of America, 103*(36), 13271–13275.

Hill, M. A., & King, E. (1995). Women's education and economic well-being. *Feminist Economics, 1*(2), 21-46.

Hodder, I. (2005). Women and men at Catalhöyük. *Scientific American, 15,* 34-41.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica, 56*(6), 1371-1395.

Horrell, S., Meredith, D., & Oxley, D. (2009). Measuring misery: Body mass, ageing and gender inequality in Victorian London. *Explorations in Economic History, 46*(1), 93–119.

Hu, L., & Schlosser, A. (2015). Prenatal Sex Selection and Girls' Well-Being: Evidence from India. *The Economic Journal, 125*(587), 1227–1261.

Huntington, S. P. (1996). Th*e Clash of Civilizations and the Remaking of World Order.* New York: Simon and Schuster.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: a meta-analysis. *Psychological bulletin, 107*(2), 139-155.

Hytten, F. E., & Leitch, I. (1971). *The Physiology of Human Pregnancy (2nd ed.).* Oxford: Blackwell Scientific Publications Ltd.

ILO (2013). *Key Indicators on the Labour Market: 7th edition*. Geneva: ILO. Retrieved from http://www.ilo.org/empelm/what/lang--en/WCMS_114240.

Inglehart, R., & Welzel, C. (2005). *Modernization, cultural change, and democracy: The human development sequence*. New York and Cambridge: Cambridge University Press.

Instat. (2012). *Population and Housing Census 2011, Tirana*. Retrieved from http://www.instat.gov.al/en/census/census-2011/publications.aspx

Instat. (2014). *Albania. Population and Population Dynamics. New Demographic Horizons?,* Tirana. Retrieved from http://www.instat.gov.al/media/242042/population_and_population_dynamics_-_new_demographic_horizons.pdf

Irwin, D., & Terviö, M. (2002). Does Trade raise Income? Evidence from the Twentieth Century. *Journal of International Economics, 58*(1), 1-18.

Iyigun, M., & Walsh, R. P. (2007). Endogenous gender power, household labor supply and the demographic transition. *Journal of Development Economics, 82*(1), 138-155.

Jacquemyn, Y., Martens, E., & Martens, G. (2014). Pregnancy at late premenopausal age: outcome of pregnancies at 45 years and older in Flanders, Belgium. *Journal of obstetrics and gynaecology, 34*(6), 479–481.

Jain, T. (2014). Where There Is a Will: Fertility Behavior and Sex Bias in Large Families. *The Journal of Human Resources*, *49*(2), 393–423.

James, W. H. (1987). The human sex ratio. Part 1: A review of the literature. *Human Biology, 59*(5), 721–752.

Jaumotte, F., Lall, S., & Papageorgiou, C. (2013). Rising Income Inequality: Technology, or Trade and Financial Globalization. *IMF Economic Review, 61*(2), 271–309.

Jayachandran, S. (2014a). *Fertility Decline and Missing Women*. NBER Working Paper No. 20272. Retrieved from http://faculty.wcas.northwestern.edu/~sjv340/fertility_decline.pdf

Jayachandran, S. (2014b). *The roots of gender inequality in developing countries*. NBER Working Paper No. 20380. Retrieved from http://faculty.wcas.northwestern.edu/~sjv340/roots_of_gender_inequality.pdf

Jayachandran, S., & Kuziemko, I. (2011). Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India. *Quarterly Journal of Economics, 126,* 1485–1538.

Jeffrey, C. (2002). Caste, Class, and Clientelism: A Political Economy of Everyday Corruption in Rural North India. *Economic Geography, 78*(1), 21–42.

Jelnov, P. (2014). *The Marriage Age U-shape*. Retrieved from http://econ.biu.ac.il/files/economics/seminars/the_marriage_age_u-shape.pdf

Jensen, R., & Oster, E. (2009). The power of TV: Cable television and women's status in India. *Quarterly Journal of Economics, 124*(3), 1057–1094.

Jensen, R., & Thornton, R. (2003). Early female marriage in the developing world. *Gender & Development*, *11*(2), 9-19.

Johnston, W. R. (2015). *Historical Abortion Statistics, Kosovo*. Johnston Archive. Retrieved from http://www.johnstonsarchive.net/policy/abortion/ab-kosovo.html

Jorde, L. B., & Wooding, S. P. (2004). Genetic variation, classification and 'race'. *Nature genetics*, *36*(11), 28–33.

Kabeer, N., & Natali, L. (2013). Gender Equality and Economic Growth: Is there a Win-Win?. *IDS Working Papers, 2013*(417), 1-58.

Kaldor, N. (1961). *Capital accumulation and economic growth*. London: Palgrave Macmillan.

Kalemli-Ozcan, S., Ryder, H. E., & Weil, D. N. (2000). Mortality decline, human capital investment, and economic growth. *Journal of Development Economics, 62*(1), 1-23.

Kaufmann, D., Kraay, A., & Zoido-Lobatón, P. (2002). *Governance Matters II-Updated Indicators for 2000/01.* World Bank Policy Research Department Working Paper No. 2772, Washington DC.

Kijima, Y. (2006). Caste and Tribe Inequality: Evidence from India, 1983–1999. *Economic Development and Cultural Change, 54*(2), 369–404.

Kim, D.-S., & Song, Y.-J. (2007). Does Religion Matter? A Study of Regional Variations in Sex Ratio at Birth in Korea. In I. Attané & C. Z. Guilmoto (Eds.), *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia* (183–206). Paris: Committee for International Cooperation in National Research in Demography.

Kim, J. (2005). Sex Selection and Fertility in a Dynamic Model of Conception and Abortion. *Journal of Population Economics, 18*, 41–67.

Kim, Y. S. (2010). *The Impact of Rainfall on Early Child Health*. University of Maryland, Maryland. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.189.2943&rep=rep1&type=pdf

King, E. M., & Hill, M. A. (Eds.). (1997). *Women's education in developing countries: Barriers, benefits, and policies.* World Bank, Baltimore, MD: John Hopkins University Press.

Kishor, S. (1993). "May God Give Sons to All": Gender and Child Mortality in India. *American Sociological Review, 58*(2), 247–265.

Klasen, S. (1994). Missing Women Reconsidered. *World Development, 22*(7), 1061–1071.

Klasen, S. (1996). Nutrition, Health, and Mortality in Sub Saharan Africa: Is There a Gender Bias? *Journal of Development Studies, 32*, 913–933.

Klasen, S. (1999). *Does gender inequality reduce growth and development: evidence from cross-country regressions*. Policy research report on gender and development working paper series; no. 7. Washington, DC: The World Bank.

Klasen, S. (2002). Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development. *World Bank Economic Review, 16*(3), 345-373.

Klasen, S. (2008). Missing Women: Some Recent Controversies on Levels and trends in Gender Bias in Mortality. In K. Basu & R. Kanbur (Eds.), *Arguments for a Better World: Essays in Honor of Amartya Sen*. Oxford: Oxford University Press.

Klasen, S., & Lamanna, F. (2009). The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries. *Feminist Economics, 15*(3), 91-132.

Klasen, S., & Wink, C. (2002). A Turning Point in Gender Bias in Mortality? An Update on the Number of Missing Women. *Population and Development Review, 28*(2), 285–312.

Klasen, S., & Wink, C. (2003). "Missing Women": Revisiting the Debate. *Feminist Economics, 9*(2-3), 263–299.

Kleibergen, F., & Paap, R. (2006). Generalized Reduced Rank Tests Using the Singular Value Decomposition. *Journal of Econometrics, 133*(1), 97–126.

Knack, S., & Keefer, P. (1995). Institutions and economic performance: cross-country tests using alternative institutional measures. *Economics & Politics, 7*(3), 207-227.

Knowles, S., Lorgelly, P. K., & Owen, P. D. (2002). Are Educational Gender Gaps a Brake on Economic Development? Some Cross-Country Empirical Evidence. *Oxford Economic Papers, 54*(1), 118–49.

Korenman, S., & Neumark, D. (1990). *Marriage, motherhood, and wages*. NBER Working Paper No. 3473. Retrieved from http://www.nber.org/papers/w3473.pdf

Korenman, S., & Neumark, D. (1991). Does marriage really make men more productive?. *Journal of Human Resources, 26*, 282-307.

Kosovo Agency of Statistics. (2014). *Census 2011. askdata*. Retrieved from http://ask.rks-gov.net/en/population/category/22-statistics-of-births

Kountouris, Y., & Remoundou, K. (2016). Cultural Influence on Preferences and Attitudes for Environmental Quality. *Kyklos, 69*(2), 369–397.

Kraay, A. (2006). When is growth pro-poor? Evidence from a panel of countries. *Journal of Development Economics, 80*(1), 198–227.

Krueger A.B., & Lindhal, M. (2001). Education for Growth: Why and For Whom? *Journal of Economic Literature, 39*, 1101–1136.

Krugman, P. (1995). Growing World Trade: Causes and Consequences. *Brookings Papers on Economic Activity, 1*, 327–377.

Kurz, H. D. (1994). Growth and distribution. *Review of Political Economy, 6*(4), 393–420.

Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review, 45(*1), 1–28.

Lagerlöf, N. P. (2003). Gender equality and long-run growth. *Journal of Economic Growth, 8*(4), 403-426.

Lahiri, S., & Self, S. (2005). *Exogamy and bias against daughters in health-care provision: A theory and evidence from two Northern states in India (mimeo).* Southern Illinois University. Retrieved from http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1033&context=econ_dp

Lahiri, S., & Self, S. (2007). Gender bias in education: the role of inter-household externality, dowry and other social institutions. *Review of Development Economics, 11*(4), 591–606.

Lanjouw, P., & Stern, N. (1991). Poverty in Pulanpur. *The World Bank Economic Review, 5*(1), 23–55.

Larsen, U., Chung, W., & Das Gupta, M. (1998). Fertility and Son Preference in Korea. *Population Studies, 52*(3), 317–325.

Lehmiller, J. J., & Agnew, C. R. (2008). Commitment in age-gap heterosexual romantic relationships: A test of evolutionary and socio-cultural predictions. *Psychology of Women Quarterly, 32*, 74-82.

Leigh, A. (2007). How Closely Do Top Income Shares Track Other Measures of Inequality?. *Economic Journal, Royal Economic Society, 117*(524), 619-633.

Leone, T., Matthews, Z., & Zuanna, G. D. (2003). Impact and Determinants of Sex Preference in Nepal. *International family planning, 29*(2), 69–75.

Levant, R. F., Slattery, S. C., & Loiselle, J. E. (1987). Fathers' involvement in housework and child care with school-aged daughters. *Family Relations, 36*(2), 152-157.

Levine, R., & Renelt, D. (1992). A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review, 82*(4), 942-963.

Li, H., Squire, L., & Zou, H. (1998). Explaining International and Intertemporal Variations in Income Inequality. *Economic Journal, 108*, 26–43.

Li, T. (2008). *Spousal age gap and family-produced elderly care*. Working Paper. Retrieved from http://www.aeaweb.org/annual_mtg_papers/2009/retrieve.php?pdfid=420

Lin, M.-J., & Luoh, M.-C. (2008). Can Hepatitis B Mothers Account for the Number of Missing Women? Evidence from Three Million Newborns in Taiwan. *American Economic Review, 98*(5), 2259–2273.

Lin, M.-J., Liu, J.-T., & Qian, N. (2014). More Missing Women, Fewer Dying Girls: The Impact of Sex-Selective Abortion on Sex at Birth and Relative Female Mortality in Taiwan. *Journal of the European Economic Association*, *12*(4), 899–926.

Lipset, S. M. (1960). *Political Man: The Social Basis of Modern Politics*. New York: Doubleday.

Lise, J., & Seitz, S. (2011). Consumption inequality and intra-household allocations. *The Review of Economic Studies*, *78*(1), 328-355.

Lockheed, M. E., & Verspoor, A. M. (1991). *Improving primary education in developing countries.* Oxford: Oxford University Press.

Lopez, J. H. (2004). *Pro-Poor Growth. A Review of What We Know (and of What We Don't)*. Mimeo, World Bank. Retrieved from http://www.eldis.org/vfile/upload/1/document/0708/DOC17880.pdf

Lopez, J. H. (2005). *Growth and Inequality: Are they Connected?* Mimeo, World Bank. Retrieved from http://siteresources.worldbank.org/INTPGI/Resources/ 342674-1115051862644/GROWTHINEQJHLvf.pdf

Lorentzen, P., McMillan, J., & Wacziarg, R. (2008). Death and development. *Journal of Economic Growth*, *13*(2), 81-124.

Loury, G. (1977). A Dynamic Theory of Racial Income Differences. In P. Wallace & A. LeMund (Eds.), *Women, Minorities, and Employment Discrimination* (86-153). Lexington, MA: Lexington Books.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3-42.

Lucas, R. E. (2015). Human Capital and Growth. *American Economic Review: Papers & Proceedings*, *105*(5): 85–88.

Lundberg, M., & Squire, L. (2003). The simultaneous evolution of growth and inequality. *The Economic Journal*, *113*(487), 326–344.

Lyster, W. R. (1974). Altered Sex-Ratio after London Smog of 1952 and Brisbane Flood of 1965. *Journal of Obstretics and Gynaecology of the British Commonwealth*, *81*(8), 626–631.

Maertens, A. (2013). Social norms and aspirations: Age of marriage and education in rural India. *World Development*, *47*, 1-15.

Malthus, T. R. (1798). *An essay on the principle of population, as it affects the future improvement of society*. London: J. Johnson.

Malthus, T. R. (1830). *A Summary View of the Principle of Population*. London: John Murray.

Mankiw, G. N., Romer, D., & Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, *107*(2), 407–437.

Mankiw, N.G., Phelps, E.S., & Romer, P.M. (1995). The Growth of Nations. *Brookings Papers on Economic Activity, 1*, 275–326.

Mansour, H., & McKinnish, T. (2014). Who Marries Differently Aged Spouses? Ability, Education, Occupation, Earnings, and Appearance. *Review of Economics and Statistics, 96*(3), 577-580.

Mariani, F. (2012). The Economic Value of Virtue. *Journal of Economic Growth, 17*(4), 323-356.

Marini, M. M. (1978). The Transition to Adulthood: Sex Differences in Educational Attainment and Age at Marriage. *American Sociological Review, 43*(4), 483-507.

Marx, K. ([1867] 2014). *Das Kapital: Kritik der politischen Ökonomie (ungekürzte Ausgabe).* Hamburg: Nikol.

Mason, K. O. (1986). The Status of Women: Conceptual and Methodological Issues in Demographic Studies. *Sociological Forum, 1*(2), 284-300.

Mason, K. O., & Smith, H. (2003). *Women's Empowerment and Social Context: Results from Five Asian Countries.* Washington, DC: Gender and Development Group, World Bank. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.200.6209&rep=rep1&type=pdf

Mathews, T. J., & Hamilton, B. E. (2005). Trend analysis of the sex ratio at birth in the United States. *National Vital Statistics Reports, 53*(20), 1–17.

Matsuo, K., Ushioda, N., & Udoff, L. C. (2009). Parental aging synergistically decreases offspring sex ratio. *The journal of obstetrics and gynaecology research, 35*(1), 164–168.

Matz, J. A. (2013). *Ethnicity, Marriage and Family Income.* ZEF Discussion Papers on Development Policy No. 180.

Mauldin, W. P., Berelson, B., & Sykes, Z. (1978). Conditions of fertility decline in developing countries, 1965-75. *Studies in Family Planning, 9*(5), 89-147.

Mauro, P. (1995). Corruption and growth. *Quarterly Journal of Economics, 110*(3), 681-712.

Mayer, P. (1999). India's falling sex ratios. *Population and Development Review, 25*(2), 323–343.

Maynes, M. J. (1985). *Schooling in Western Europe: A social history*. Suny Press.

McCleary, R. M., & Barro, R. J. (2006). Religion and Economy. *Journal of Economic Perspectives*, *20*(2): 49-72.

McCrate, E. (1999). Class. In J. Peterson & M. Lewis (Eds.), *The Elgar Companion to Feminist Economics.* Cheltenham, UK: Edward Elgar Publishing.

McKinsey & Company (2015). *The Power of Parity: How Advancing Women's Equality Can Add $12 Trillion to Global Growth*. McKinsey Global Institute. Retrieved from http://www.mckinsey.com/insights/growth/how_advancing_womens_equality_can_add_12_trillion_to_global_growth

Meldolesi, A. (2012), *Mai nate. Perché il mondo ha perso 100 milioni di donne*. Milano: Mondadori Università.

Mensch, B. S., Singh, S., & Casterline, J. B. (2005). Trends in the timing of first marriage among men and women in the developing world. In C. B. Lloyd (Ed.), *The changing transitions to adulthood in developing countries: Selected Studies*. Washington, DC: National Academies Press.

Meslé, F., Vallin, J., & Badurashvili, I. (2007). A sharp increase in sex ratio at birth in the Caucasus. Why? How? In I. Attané & C. Z. Guilmoto (Eds.), *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia* (73–88). Paris: Committee for International Cooperation in National Research in Demography.

Milanovic, B. (2000). Determinants of cross-country income inequality: An augmented Kuznets' hypothesis. In M. Uvalic & V. Franicevic (Eds.), *Equality, participation, transition — Essays in the honor of Branko Horvat* (48–79). London: St. Martin's.

Milanovic, B. (2009). *Global Inequality and the Global Inequality Extraction Ratio. The Story of the Past Two Centuries*. World Bank. Retrieved from https://openknowledge.worldbank.com/handle/10986/4237

Milanovic, B. (2011). *Worlds apart: measuring international and global inequality*. Princeton: Princeton University Press.

Milazzo, A. (2012). Wh*y Are Adult Women Missing? Son Preference and Maternal Survival in India.* Bocconi University, Department of Economics, University of Chicago.

Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics, 24*(3), 1071-1100.

Miller, B. D. (1985). The Unwanted Girls: A Study of Infant Mortality Rates. *Manushi, 29*, 18–20.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy, 66*(4), 281–302.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. Human Behavior & Social Institutions No. 2.

Mincer, J., & Polachek, S. (1974). Family Investments in Human Capital: Earnings of Women. *Journal of Political Economy, 82*(2), 76-108.

Monna, B., & Gauthier, A. H. (2008). A review of the literature on the social and economic determinants of parental time. *Journal of Family and Economic Issues, 29*(4), 634-653.

Montesquieu, C.L., Baron de. ([1748]1989). *The Spirit of the Laws*. Cambridge: Cambridge University Press.

Moore, K. A., Myers, D. E., Morrison, D. R., Nord, C. W., Brown, B., & Edmonston, B. (1993). Age at First Childbirth and Later Poverty. *Journal of Research on Adolescence, 3*(4), 393–422.

Morris, L., Herold, J., Bino, S., Yili, A., & Jackson, D. (2005). *Albania reproductive health survey, 2002. Final report*. Tirana: Albania Ministry of Health and Institute of Statistics.

Morrison, A. R., Raju, D., & Sinha, N. (2007). *Gender equality, poverty and economic growth*. World Bank Policy Research Working Paper No. 4349.

Mozersky, J., & Mennuti, M. T. (2013). Cell-free fetal DNA testing: who is driving implementation? *Genetics in Medicine*, *15*(6), 433–434.

Mulhern, J. (1959). *A history of education: a social interpretation*. New York: Ronald Press.

Murthi, M., Guio, A.-C., & Drèze, J. (1995). Mortality, Fertility, and Gender Bias in India: A District-Level Analysis. *Population and Development Review, 21*(4), 745–782.

Murtin, F., & Wacziarg, R. (2014). The democratic transition. *Journal of Economic Growth, 19*(2), 141-181.

Nandi, A., & Deolalikar, A. B. (2013). Does a Legal Ban on Sex-Selective Abortions Improve Child Sex Ratios? Evidence from a Policy Change in India. *Journal of Development Economics, 103*, 216–228.

Naqvi, S. N. H. (1996). The Significance of Development Economics. *World Development, 24*(6), 975–987.

Neckerman, K. M. (2004). *Social Inequality*. New York: Russell Sage Foundation.

Neckerman, K. M., & Torche, F. (2007). Inequality: Causes and Consequences. *Annual Review of Sociology, 33*(1), 335–357.

Nicholas, S., & Oxley, D. (1993). The living standards of women during the industrial revolution. *The economic history review, 46*(4), 723–749.

Noguer, M., & Siscart, M. (2005). Trade raises Income: a precise and robust result. *Journal of International Economics, 65*(2), 447-460.

Nolan, P., & Lenski, G. (2004). *Human societies: an introduction to macrosociology*. Boulder, Colorado: Paradigm Publishers.

Odedokun, M. O., & Round, J. I. (2004). Determinants of income inequality and its effects on economic growth: Evidence from African countries. *African Development Review, 16(*2), 287–327.

OECD. (2011). *Doing Better for Families*. Retrieved from http://www.oecd.org/els/soc/47701118.pdf

Ondrich, J., Richardson, J. D., & Zhang, S. (2006). A Further Investigation of the Link between Trade and Income. *International Economic Journal, 20*(1), 19-36.

Orton, M., & Rowlingson, K. (2007). *Public attitudes to economic inequality*. York: Joseph Rowntree Foundation.

Oster, E. (2005). Hepatitis B and the Case of the Missing Women. *Journal of Political Economy, 113*(6), 1163–1216.

Oster, E. (2014). *Unobservable selection and coefficient stability: Theory and evidence*. Brown University and NBER. Retrieved from https://www.brown.edu/research/projects/oster/sites/brown.edu.research.projects.oster/files/uploads/Unobservable_Selection_and_Coefficient_Stability.pdf

Oster, E., Chen, G., Yu, X., & Lin, W. (2010). Hepatitis B Does Not Explain Male-Biased Sex Ratios in China. *Economics Letters, 107*(2), 142-144.

Oxfam. (2016). *An Economy For the 1%*. Retrieved from http://oxf.am/Znhx

Papanek, G. S., & Kyn, O. (1986). The Effect of Income Distribution on Development, the Growth Rate, and Economic Strategy. *Journal of Development Economics, 23*, 55–66.

Park, C. B., & Cho, N.-H. (1995). Consequences of Son Preference in a Low-Fertility Society: Imbalance of the Sex Ratio at Birth in Korea. *Population and Development Review, 21*(1), 59–84.

Park, S., Bowen, W. M., & Steinbacher, R. (2012). Social and demographic dimensions of sex selection technologies: Review and analysis of the research literature. *Biodemography and Social Biology, 58*(1), 62–74.

Paukert, F. (1973). Income Distribution at Different Levels of Development: A Survey of Evidence. *International Labour Review, 108*, 97–125.

Pekkarinen, T. (2012). Gender differences in education. *Nordic Economic Policy Review, 1*, 165-194.

Piketty, T. & Saez, E. (2003). Income Inequality in the United States, 1913-1998. *Quarterly Journal of Economics, 118*(1), 1-39.

Piketty, T. (1995). Social Mobility and Redistributive Politics. *Quarterly Journal of Economics, 110*(3), 551–584.

Piketty, T. (2001). *Les hauts revenus en France au XXe siècle: Inegalités et redistributions 1901–1998.* Paris: Grasset.

Piketty, T. (2003). Income Inequality in France, 1901–1998. *Journal of Political Economy, 111*(5), 1004–1042.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge, MA: Harvard University Press.

Plan International. (2008). *Paying the Price: The Economic Cost of Failing to Educate Girls.* Woking, England. Retrieved from https://www.planusa.org/docs/PayingthePrice.pdf

Polachek, S. W., & Xiang, J. (2009). *The gender pay gap across countries: A human capital approach.* Retrieved from http://www.sole-jole.org/9203.pdf

Polasek, O., Kolcic, I., Kolaric, B., & Rudan, I. (2005). Sex ratio at birth and war in Croatia (1991–1995). *Human reproduction, 20*(9), 2489–2491.

Porter, M., Stern, S., & Green, M. (2014). *Social Progress Index 2014*. Retrieved from http://www.socialprogressimperative.org/publications

Potrafke, N., & Ursprung, H. W. (2012). Globalization and gender equality in the course of development. *European Journal of Political Economy, 28*(4), 399-413.

Preston, S.H., Heuveline, P., & Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes.* Malden, MA: Blackwell Publishing.

Pritchett, L. (1997). Divergence, Big Time. *Journal of Economic Perspectives, 11*(3), 3–17.

Pritchett, L. (2001). Where has all the education gone?. *The World Bank Economic Review, 15*(3), 367-391.

Psacharopoulos, G. (1988). Education and development: a review. *The World Bank Research Observer, 3*(1), 99-116.

Psacharopoulos, G., & Woodhall, M. (1985). *Education for development: An analysis of investment choices.* Oxford: Oxford University Press.

Pyke, K., & Adams, M. (2010). What's Age Got to Do With It? A Case Study Analysis of Power and Gender in Husband-Older Marriages. *Journal of Family Issues, 31*(6), 748-777.

Qian, N. (2008). Missing Women and the Price of Tea in China. The Effect of Sex-Specific Earnings on Sex Imbalance. *Quarterly Journal of Economics, 123*(3), 1251–1285.

Quah D. (2001). *Some simple arithmetic on how income inequality and economic growth matter*. Paper presented at WIDER conference on Growth and Poverty, Helsinki. Retrieved from http://s3.amazonaws.com/academia.edu.documents/29138/a5s3su4onpfkn410009.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1477582685&Signature=mbFIlBcGrS%2FiOsh5jzuD8yO3Ktc%3D&response-content-disposition=inline%3B%20filename%3DPDF.pdf

Quah, D. (2002). *One Third of the World's Growth and Inequality*. Retrieved from http://eprints.lse.ac.uk/2019/1/One_Third_of_the_World's_Growth_and_Inequality.pdf

Ram, R. (1984). Population Increase, Economic Growth, Educational Inequality, and Income Distribution: Some Recent Evidence. *Journal of Development Economics, 14,* 419-428.

Ram, R., (1989). Can Educational Expansion Reduce Income Inequality in Less-Developed Countries? *Economics of Education Review, 8*(2), 185-189.

Rao, V. (1993). The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India. *Journal of Political Economy, 101*, 666-677.

Ravallion, M. (2007). Inequality is bad for the poor. In J. Micklewright & S. Jenkins (Eds.), *Inequality and Poverty Re-Examined*. Oxford: Oxford University Press.

Ravallion, M., & Datt, G. (2002). Why has economic growth been more pro-poor in some states of India than others? *Journal of Development Economics, 68*, 381-400.

Rawls, J. (1999). *A Theory of Justice, rev. ed.* Cambridge, MA: Belknap.

Ray, D. (1998). *Development economics*. Princeton, N.J., Chichester: Princeton University Press.

Retherford, R. D., & Roy, T. K. (2003). *Factors Affecting Sex-selective Abortion in India and 17 Major States.* Retrieved from http://scholarspace.manoa.hawaii.edu/bitstream/handle/10125/3488/NFHSsubjrpt021.pdf?sequence1

Ringen, S. (2006). *Reflections on inequality and equality.* Social Science Research Center Berlin (WZB). Retrieved from https://core.ac.uk/download/files/153/6628019.pdf

Ritter, K. (2016, May 31). *Weshalb es immer weniger Schwedinnen gibt*. Die Welt. Retrieved from http://www.welt.de/wissenschaft/article155832857/Weshalb-es-immer-weniger-Schwedinnen-gibt.html

Robertson, J. A. (2001). Preconception Gender Selection. *American Journal of Bioethics, 1*(1), 2–9.

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth, 9*, 131-165.

Roine, J., Vlachos, J., & Waldenström, D. (2009). The long-run determinants of inequality: What can we learn from top income data? *Journal of Public Economics, 93*, 974-988.

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy, 94*(5), 1002–1037.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy, 98,* 71-102.

Romer, P. M. (1994). The Origins of Endogenous Growth. *The Journal of Economic Perspectives, 8*(1), 3–22.

Rose, E. (1999). Consumption smoothing and excess female mortality in rural India. *Review of Economics and Statistics, 81*(1), 41–49.

Rosenblum, D. (2013). *Economic Incentives for Sex Selective Abortion in India*. Canadian Centre for Health Economics. Retrieved from http://myweb.dal.ca/dn686159/Selective%20Abortion%20Rosenblum%208_13_13.pdf

Ross, M. L. (2008). Oil, Islam, and Women. *American Political Science Review, 102*(1), 107–123.

Rueness, J., Vatten, L., & Eskild, A. (2012). The human sex ratio: effects of maternal age. *Human Reproduction, 27*(1), 283–287.

Saardchom, N., & Lemaire, J. (2005). Causes of increasing ages at marriage: An international regression study. *Marriage & family review, 37*(3), 73-97.

Sachs, J. D. (2000). Notes on a New Sociology of Economic Development. In L. E. Harrison and S. P. Huntington (Eds.), *Culture Matters: How Values Shape Human Progress.* New York, NY: Basic Books.

Sachs, J. D. (2001). *Tropical Underdevelopment*. NBER Working Paper No. 8119.

Sachs, J. D., & Warner, A. (1995). *Economic Convergence and Economic Policies*. NBER Working Paper No. 5039.

Saez, E., Slemrod, J., & Giertz, S. H. (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature, 50*(1), 3-50.

Sala-i-Martin, X. (2006). The World Distribution of Income: Falling Poverty and … Convergence, Period. *Quarterly Journal of Economics, 121*(2), 351-397.

Sala-i-Martin, X., Doppelhofer, G., & Miller R. I. (2004). Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review, 94*(4), 813-835.

Sarmadi, D. (2015). *Oxfam Germany director: 'TTIP would intensify inequality in developing countries'.* Retrieved from http://www.euractiv.com/section/development-policy/interview/oxfam-germany-director-ttip-would-intensify-inequality-in-developing-countries/

Savvides, A., & Stengos, T. (2008). *Human capital and economic growth*. Stanford: Stanford University Press.

SBS Radio. (2015). *The news "it's a girl", still unwelcome in some cultures in Australia*. Retrieved from http://www.sbs.com.au/radio/storystream/news-its-girl-still-unwelcome-some-cultures-australia

Schultz, T. P. (2002). Why governments should invest more to educate girls. *World Development, 30*(2), 207-225.

Schultz, T. W. (1961). Investment in Human Capital. *American Economic Review, 51*(1), 1–17.

Sedgley, N., & Elmslie, B. (2005). Discrimination and growth: The distribution across skill matters. *Economic Letters, 90,* 194-199.

Seguino, S., & Floro, M. S. (2003). Does gender have any effect on aggregate saving? An empirical analysis. *International Review of Applied Economics, 17*(2), 147-166.

Sekhri, S., & Debnath, S. (2014). Intergenerational Consequences of Early Age Marriages of Girls: Effect on Children's Human Capital. *The Journal of Development Studies, 50*(12), 1670-1686.

Sen, A. (1985a). *Commodities and Capabilities*. Amsterdam: Elsevier.

Sen, A. (1985b). The Standard of Living: Lecture I - Concepts and Critiques. In G. Hawthorne (Ed.), *The Standard of Living.* Cambridge: Cambridge University Press.

Sen, A. (1990). More Than 100 Million Women Are Missing. *The New York Review of Books, 37*(20), 61–66.

Sen, A. (1992). Missing women: Social inequality outweighs women's survival advantage in Asia and North Africa. *BMJ: British Medical Journal, 304*(6827), 587.

Sen, A. (1995). *Inequality Reexamined*. New York: Russell Sage Foundation.

Sen, A. (1999). *Development as Freedom*. New York: Knopf.

Sernau, S. (2013). *Social Inequality in a Global Age*. Thousand Oaks, CA: SAGE Publications.

Shepherd, M. E. (2008). *The Selective Abortion in India: The Impact on Child Mortality*. Amherst, New York: Cambria Press.

Singh, N., Pripp, A., Brekke, T., & Stray-Pedersen, B. (2010). Different sex ratios of children born to Indian and Pakistani immigrants in Norway. *BMC Pregnancy and Childbirth, 10*(1), 40.

Sinno, S. M., & Killen, M. (2009). Moms at work and dads at home: Children's evaluations of parental roles. *Applied Developmental Science, 13*(1), 16-29.

Smith, A. ([1776] 1976). *An Inquiry into the Nature and Causes of the Wealth of Nations.* Chicago: Chicago University Press.

SOEP. (2015). *Version 30 of the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW),* Berlin. Retrieved from https://www.diw.de/en/soep

Solow, R. M. (1986). Unemployment: Getting the questions right. *Economica, 53* (210), 23-34.

Sommerfelt, A. E., & Arnold, F. (1998). Sex differentials in the nutritional status of young children. In *Too Young to Die: Genes or Gender?* (133-153). New York: United Nations.

Sonfield, A., Hasstedt, K., Kavanaugh, M. L., & Anderson, R. (2013). *The social and economic benefits of women's ability to determine whether and when to have children.* Guttmacher Institute, New York. Retrieved from www.guttmacher.org/pubs/social-economic-benefits.pdf

Spilimbergo, A., Londoño, J.L., & Székely, M. (1999). Income distribution, factor endowments, and trade openness. *Journal of Development Economics*, *59*(1), 77–101.

Srinivas, M. N. (2000). *Social change in modern India*. New Delhi: Orient Longman India.

Staiger, D., & Stock, J.H. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*(3), 557-586.

Statistics Sweden (2014). *Women and men in Sweden 2014*. Retrieved from http://jamda.ub.gu.se/bitstream/1/833/1/scb_eng_2014.pdf

Stimpfle, A., & Stadelmann, D. (2015). *The Impact of Fundamental Development Factors on Different Income Groups: International Evidence*. Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Trade, finance and institutions, No. E08-V1. Retrieved from https://www.econstor.eu/bitstream/10419/113128/1/VfS_2015_pid_254.pdf

Stimpfle, A., & Stadelmann, D. (2016a). *Marriage Age Affects Educational Gender Inequality: International Evidence*. Retrieved from http://www.crema-research.ch/papers/2016-02.pdf

Stimpfle, A., & Stadelmann, D. (2016b). *Does Central Europe Import the Missing Women Phenomenon?* Retrieved from http://www.crema-research.ch/papers/2016-04.pdf

Stock, J. H., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In Identification and Inference for Econometric Models. In D. W. K. Andrews & J.H. Stock (Eds.), *Essays in Honor of Thomas Rothenberg* (80-108). New York: Cambridge University Press.

Stump, D. (2011). *Prenatal Sex Selection.* Council of Europe, Committee on Equal Opportunities for Women and Men, Strasbourg. Retrieved from https://www.ecoi.net/file_upload/1226_1317113911_edoc12715prenatal.pdf

Sudha, S., & Rajan, S. I. (1999). Female Demographic Disadvantage in India 1981–1991: Sex Selective Abortions and Female Infanticide. *Development and Change, 30,* 585–618.

Sudha, S., Khanna, S., Rajan, S. I., & Srivastava, R. (2007). Is Son Preference Emerging among the Nayars of Kerala in South India? In I. Attané & C. Z. Guilmoto (Eds.), *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia* (267–294). Paris: Committee for International Cooperation in National Research in Demography.

Sun, A., & Zhao, Y. (2011). *Divorce, Abortion and Sex Ratio at Birth: The Effect of the Amended Divorce Law in China* (Working Paper). Brown University. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.594.2956&rep=rep1&type=pdf

Sun, R. (2002). Old age support in contemporary urban China from both parents' and children's perspectives. *Research on Aging, 24*(3), 337–359.

Sweetland, S. R. (1996). Human Capital Theory: Foundations of a Field of Inquiry. *Review of Educational Research, 66*(3), 341–359.

Takahashi, E. (1954). The Effects of the Age of the Mother on the Sex Ratio at Birth in Japan. *Annals of the New York Academy of Sciences, 57*(5), 531–550.

The Guardian. (2016). *Turkey hits back after Sweden accuses it of legalising child sex*. Retrieved from https://www.theguardian.com/world/2016/aug/15/sweden-accuses-turkey-legalising-child-sex

Thurow, L. C. (1987). A Surge in Inequality. *Scientific American, 256*, 30–37.

Timaeus, I., Harris, K., & Fairbairn, F. (1998). Can Use of Health Care Explain Sex Differentials in Mortality in the Developing World? In United Nations (Ed.), *Too Young to Die – Genes or Gender?*. New York: United Nations.

Todaro, M. P., & Smith, S. (2014). *Economic Development 12th Edition*. Boston, New York: Addison Wesley.

Too, Y. L. (2001). *Education in Greek and Roman antiquity*. Leiden: Brill.

Tzannatos, Z. (1999). Women and Labor Market Changes in the Global Economy: Growth Helps: Inequalities Hurt and Public Policy Matters. *World Development, 27*(3), 1010–1046.

UNDP. (2013). *Gender Inquality Index*. Retrieved from http://hdr.undp.org/en/data

UNDP. (2016). *Sustainable Development Goals. UNDP Support to the Implementation of Sustainable Development Gola 10.* Retrieved from http://www.undp.org/content/dam/undp/library/Sustainable%20Development/10_Reducing_Inequality_Jan15_digital.pdf?download

UNESCO. (2013). *Institute for Statistics (UIS).* Retrieved from http://data.uis.unesco.org/.

UNFPA. (2012a). *Sex Imbalances at Birth in Albania, Tirana.* Retrieved from http://www.unfpa.org/resources/sex-imbalances-birth-albania

UNFPA. (2012b). *Sex Imbalances at Birth: Current trends, consequences and policy implications.* UNFPA Asia and the Pacific Regional Office, Bangkok. Retrieved from http://www.unfpa.org/publications/sex-imbalances-birth

United Nations Population Division. (2015). *World Population Prospects: The 2015 Revision.* United Nations, New York. Retrieved from http://esa.un.org/unpd/wpp/index.htm

United Nations. (2013). *World Marriage Data 2012*. Department of Economic and Social Affairs, Population Division. Retrieved from http://www.un.org/esa/population/publications/WMD2012/MainFrame.html

United Nations. (2015). *The World's Women 2015 Chapter 3.* Retrieved from http://unstats.un.org/unsd/gender/downloads/Ch3_Education_info.pdf

United Nations. (2016). *International Migration Report 2015.* Department of Economic and Social Affairs, Population Division. Retrieved from http://www.un.org/en/development/desa/population/migration/publications/migrationreport/docs/MigrationReport2015_Highlights.pdf

UNU-WIDER. (2014a). *World Income Inequality Database (WIID V3.0B).* Retrieved from http://www.wider.unu.edu/research/WIID3-0B/en_GB/wiid/

UNU-WIDER. (2014b). *WIID – World Income Inequality Database version 3a. User Guide and Data Sources.* Retrieved from http://www.wider.unu.edu/research/WIID3-0B/en_GB/WIID-documentation/_files/92393308398364306/default/User%20guide.pdf

Utomo, A. J. (2014). Marrying Up? Trends in Age and Education Gaps Among Married Couples in Indonesia. *Journal of Family Issues, 35*, 1683-1706.

Van de Walle, D., & Gunewardena, D. (2001). Sources of Ethnic Inequality in Viet Nam. *Journal of Development Economics, 65*(1), 177–207.

Van der Vleuten, L. (2013). *Mind the gap! The persistence of gender educational gaps in developing countries, 1950-2005*. Retrieved from http://vkc.library.uu.nl/vkc/seh/Lists/Events/Attachments/33/vleuten_mind.pdf

Van Zanden, J.L. (2011). The Malthusian Intermezzo. Women's wages and human capital formation between the Late Middle Ages and the Demographic Transition of the 19th century. *The History of the Family, 16*(4), 331-342.

Vella, F., & Collins, S. (1990). The value of youth: equalizing age differentials in marriages. *Applied Economics, 22*(3), 359-373.

Verropoulou, G., & Tsimbos, C. (2010). Differentials in sex ratio at birth among natives and immigrants in Greece: An analysis employing nationwide micro-data. *Journal of Biosocial Science, 42*, 425–430.

Viner, J. (1953). *International Trade and Economic Development*. Oxford: Oxford University Press.

Visaria, P. (1971). Provisional population totals of the 1971 census: some questions and research issues. *Economic and Political Weekly, 6*(29), 1459–1465.

Wade, R. (2001) *Is Globalization Making World Income Distribution More Equal?* DESTIN Working Paper No.1. London: LSE Development Studies Institute. Retrieved from http://www.lse.ac.uk/internationalDevelopment/pdf/WP/WP10.pdf

Weil, D. N. (2007). Accounting for the Effect of Health on Economic Growth. *Quarterly Journal of Economics, 122*(3), 1265-1306.

Westoff, C. F. (2003). *Trends in marriage and early childbearing in developing countries.* Comparative Reports No. 5. ORC Macro, Calverton, Maryland, USA. Retrieved from http://dhsprogram.com/pubs/pdf/CR5/CR5.pdf

Westoff, C. F. (2005). *Recent trends in abortion and contraception in 12 countries.* ORC Macro, Calverton, Maryland. Retrieved from http://dhsprogram.com/pubs/pdf/AS8/AS8.pdf

White, H., & Anderson, E. (2002). Growth versus Distribution: Does the Pattern of Growth Matter? *Development Policy Review, 19*(3), 267-289.

Wolf, A. (2002). *Does Education Matter? Myths About Education and Economic Growth.* London: Penguin.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

World Bank. (2011). *World Development Report 2012. Gender Equality and Development,* Washington DC. Retrieved from https://siteresources.worldbank.org/INTWDR2012/Resources/7778105-1299699968583/7786210-1315936222006/Complete-Report.pdf

World Bank. (2014). *World Development Indicators (WDI) December 2014.* Retrieved from http://data.worldbank.org/data-catalog/world-development-indicators.

World Economic Forum. (2014). *The Global Gender Gap Report 2014*. Geneva. World Economic Forum. Retrieved from http://www3.weforum.org/docs/GGGR14/GGGR_CompleteReport_2014.pdf

World Health Organization Europe. (2015). *European Health Information Gateway.* Retrieved from http://portal.euro.who.int/en/indicators/hfa-indicators/hfa_20-0064-number-of-live-births-males/

Yount, K. M. (2001). Excess Mortality of Girls in the Middle East in the 1970s and 1980s: Patterns, Correlates and Gaps in Research. *Population Studies, 55*(3), 291–308.

Zak, P. J., & Knack, S. (2001). Trust and Growth. *The Economic Journal, 111*(470), 295–321.

Zeng, Y., Ping, T., Baochang, G., Yi, X., Bohua, L., & Yongpiing, L. (1993). Causes and Implications of the Recent Increase in the Reported Sex Ratio at Birth in China. *Population and Development Review, 19*(2), 283–302.

Zhang, H. (2014) *Marriage Age Patterns: A Unifying Theory and Global Evidence.* Working Paper. Retrieved from http://www.gtcenter.org/Downloads/Conf/Zhang1867.pdf

Zhang, X. (2014). Fecundity and Husband-Wife Age Gap at First Marriage – Cross-Country Analysis. *Theoretical Economics Letters, 4,* 598-607.