

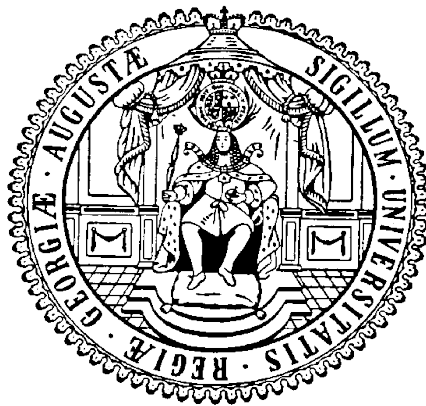
---

# ESSAYS ON GENDER, MIGRATION, AND DEVELOPMENT

---

Dissertation in order to acquire the doctoral degree  
from the Faculty of Economic Sciences,  
at the Georg-August-Universität Göttingen

Submitted by  
Manuel Pedro Duarte Santos Silva  
Born in Porto, Portugal



Göttingen, 2018



# Acknowledgment

Many people contributed, in one way or another, to this volume. I have learned a lot from Stephan Klasen and benefited greatly from the outstanding research environment at his chair. I remain today as impressed by his knowledge and scientific drive as when I first attended his lectures, as a master student, in the fall of 2012. I also want to thank the other two members of my advisory committee, Krisztina Kis-Katos and Amy Alexander, for their support.

I was fortunate to be a member of the Research Training Group 1723 “Globalization and Development” funded by the German Research Foundation (DFG). For chapters 1 and 2, funds came from the Growth and Economic Opportunities for Women (GrOW) initiative, multi-funded by the UK’s Department for International Development, the Hewlett Foundation, and the International Development Research Centre. I have wonderful memories of the GrOW workshops in New Delhi and Stellenbosch. While working on chapter 2, I visited Janneke Pieters at Wageningen University, in November 2017; I am grateful to her and her colleagues for making those weeks such a nice experience. For chapter 3, I received financial support from Lower Saxony’s Ministry for Science and Culture. The views expressed herein do not necessarily represent those of any funding institution.

I worked with several other co-authors for parts of this volume and wish to thank in particular Lisa Höckel, Janneke Pieters, Tobias Stöhr, and Christian Welzel. It was a pleasure working with all of you.

The development economics group in Göttingen has gathered, over the years, a terrific group of young people. I am sure that, many years from now, I will enjoy seeing some of you side-by-side on this page. In random order: Lennart Kaplan, Esther Heesemann, Cara Erbert, Anna Minasyan, Hendrik Kruse, Atika Pasha, Nathalie Scholl, Laura Barros, Marcello Perez, Jana Lenze, Sarah Khan, and Slava Yakubenko. A separate word is due to Bruno Witzel-Souza, with whom I shared an office throughout the whole thing. It was a lucky match. A transcript of our conversations would result, I suspect, in a equally interesting dissertation (once the “colorful” language is weeded out).

Now, the heart beats faster. Lisa was patient and supportive at every stage. She endured unreadable first drafts and navigated through horrendous regression tables, spotting many mistakes of style and argument. But, beyond all that, it was her love and company that kept me going. Thank you.

Finally, I thank my family for their constant support—my parents, Isabel and Augusto, and my sisters, Rita and Teresa. And to my nieces, Luísa and Teresa, and nephew, Francisco: if I had your curiosity and energy, this thesis would be undoubtedly better.

All remaining errors are my own.



# Contents

<b>Acknowledgments</b>	<b>i</b>
<b>List of Tables</b>	<b>v</b>
<b>List of Figures</b>	<b>vii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Gender Inequality as a Barrier to Economic Growth</b>	<b>7</b>
1.1 Introduction . . . . .	8
1.2 A simple efficiency argument: men and women . . . . .	10
1.3 Unitary households: parents and children . . . . .	13
1.4 Intra-household bargaining: husbands and wives . . . . .	19
1.5 Household formation patterns . . . . .	24
1.6 Beyond the household: openness, politics, and corruption . . . . .	27
1.7 Conclusion . . . . .	31
<b>2 What Drives Female Labor Force Participation?</b>	<b>35</b>
2.1 Introduction . . . . .	36
2.2 Data and empirical model . . . . .	40
2.2.1 Data . . . . .	40
2.2.2 Descriptives . . . . .	43
2.2.3 Modeling female labor force participation . . . . .	49
2.3 Results . . . . .	51
2.3.1 Selection into marriage and urban areas . . . . .	59
2.3.2 Selection into education . . . . .	60
2.4 Decomposition analysis . . . . .	63
2.4.1 Decomposing changes over time within countries . . . . .	64
2.4.2 Decomposing differences between countries . . . . .	67
2.5 Conclusion . . . . .	71
2.A Appendix . . . . .	73
<b>3 The Roots of Female Emancipation</b>	<b>107</b>
3.1 Introduction . . . . .	108
3.2 Theoretical discussion . . . . .	110
3.2.1 Original sources of gender (in)equality . . . . .	110
3.2.2 Historical household formation patterns . . . . .	112

3.2.3	Implications for gender equality . . . . .	114
3.2.4	Origins of household formation patterns . . . . .	117
3.3	Data and descriptives . . . . .	122
3.4	Empirical strategy . . . . .	130
3.5	Results . . . . .	131
3.5.1	Ages at first marriage and gender equality . . . . .	131
3.5.2	Cool Water breeds late-marriage societies . . . . .	134
3.5.3	Cool Water and historic late marriages . . . . .	144
3.5.4	Cool Water and contemporary gender equality . . . . .	146
3.6	Conclusion . . . . .	148
3.A	Appendix . . . . .	149
<b>4</b>	<b>Can Parental Migration Reduce Petty Corruption in Education?</b>	<b>163</b>
4.1	Introduction . . . . .	164
4.2	Related literature . . . . .	165
4.3	Moldova and corruption in education . . . . .	167
4.4	Data and descriptives . . . . .	169
4.5	Empirical strategy . . . . .	172
4.6	Main results . . . . .	174
4.7	Transmission channels and robustness . . . . .	179
4.8	Conclusion . . . . .	184
4.A	Appendix . . . . .	186
	<b>Bibliography</b>	<b>205</b>

# List of Tables

2.1	Estimation results: overview . . . . .	53
2.A.1	Data overview . . . . .	85
2.A.2	South Africa: sample means . . . . .	87
2.A.3	Brazil: sample means . . . . .	87
2.A.4	Jordan: sample means . . . . .	88
2.A.5	India: sample means . . . . .	89
2.A.6	Bolivia: sample means . . . . .	89
2.A.7	Indonesia: sample means . . . . .	90
2.A.8	Vietnam: sample means . . . . .	90
2.A.9	Tanzania: sample means . . . . .	91
2.A.10	South Africa: average marginal effects . . . . .	92
2.A.11	Brazil: average marginal effects . . . . .	93
2.A.12	Jordan: average marginal effects . . . . .	94
2.A.13	India: average marginal effects . . . . .	95
2.A.14	Bolivia: average marginal effects . . . . .	96
2.A.15	Indonesia: average marginal effects . . . . .	97
2.A.16	Vietnam: average marginal effects . . . . .	98
2.A.17	Tanzania: average marginal effects . . . . .	99
2.A.18	Trends in sample inclusion criteria over time . . . . .	100
2.A.19	India and Jordan: decomposition of FLFP . . . . .	101
2.A.20	Brazil and South Africa: decomposition of FLFP . . . . .	101
2.A.21	Indonesia and Bolivia: decomposition of FLFP . . . . .	102
2.A.22	Tanzania and Vietnam: decomposition of FLFP . . . . .	102
2.A.23	South Africa: decomposition of FLFP, 1995–2014 . . . . .	103
2.A.24	Educational attainment and social group: common variables for all countries . . . . .	104
2.A.25	Decomposition of FLFP between countries: first year . . . . .	105
2.A.26	Decomposition of FLFP between countries: last year . . . . .	106
3.1	Descriptive statistics for selected variables . . . . .	123
3.2	Determinants of gender gaps: ages at first marriage . . . . .	132
3.3	Determinants of gender gaps: ages at first marriage; subsample analysis . . . . .	135
3.4	Determinants of ages at first marriage . . . . .	137
3.5	Determinants of ages at first marriage: subsample analysis . . . . .	138
3.6	Determinants of ages at first marriage: additional controls . . . . .	142
3.7	Assessing unobservable selection: estimates of $\delta$ . . . . .	144
3.8	Europe: historical female ages at first marriage . . . . .	145

3.9	Determinants of gender gaps: reduced form estimates . . . . .	147
3.A.1	Description of variables used and their source . . . . .	150
3.A.2	Additional gender equality outcomes and ages at first marriage . . . . .	156
3.A.3	Determinants of ages at first marriage: geo-climatic variables . . . . .	157
3.A.4	Determinants of ages at first marriage: ancestry-adjustment . . . . .	158
3.A.5	Determinants of ages at first marriage: excluding Northwest Europe and Western offshoots . . . . .	159
3.A.6	Determinants of ages at first marriage: robustness to inclusion of addi- tional controls . . . . .	160
3.A.7	Gapminder: historical female ages at first marriage . . . . .	161
4.1	Selected summary statistics . . . . .	171
4.2	The effect of migration on private education inputs . . . . .	175
4.3	The effect of migration on private education inputs: controlling for house- hold assets . . . . .	178
4.4	Detailed school funding, infrastructure controls and school fixed effects	181
4.A.1	School enrollment rates in Moldova (2010) . . . . .	188
4.A.2	School budgets: descriptive statistics . . . . .	190
4.A.3	The effect of migration on private education inputs: reduced form esti- mates . . . . .	191
4.A.4	The effect of migration on private education inputs: first stage IV regres- sion . . . . .	192
4.A.5	The effect of migration on private education inputs . . . . .	193
4.A.6	The effect of migration on education expenditures: average marginal effects after IV probit . . . . .	194
4.A.7	The effect of migration on private education inputs: OLS regressions . .	195
4.A.8	Does migration predict perceived barriers to education? . . . . .	196
4.A.9	Main problems with local public schools: community leaders' perceptions	197
4.A.10	The effect of migration on private education inputs: controlling for public school expenditures . . . . .	197
4.A.11	Determinants of caregiver time: OLS results; sample split by migration status . . . . .	198
4.A.12	Household-level IV: second stage . . . . .	199
4.A.13	Second stage on perceived school quality . . . . .	200
4.A.14	The effect of migration on private education inputs: when the caregiver is a biological parent . . . . .	201
4.A.15	Determinants of grade point average (GPA) . . . . .	202
4.A.16	Summary statistics: sample selection from matching . . . . .	203



# List of Figures

1	Rising scholarship on gender equality . . . . .	2
2.1	Selected country indicators for the first and last years in our dataset . . .	41
2.2	Selected gender indicators with respect to per capita income in 2014 . . .	42
2.3	Labor force participation by gender and marital status . . . . .	44
2.4	Distribution of educational attainment over time . . . . .	45
2.5	FLFP by education level . . . . .	47
2.6	Average number of children in a urban married woman's household . .	48
2.7	Average marginal effects of the woman's own education . . . . .	56
2.8	Average marginal effects of log household per capita earnings (excluding woman's own earnings) . . . . .	57
2.9	Average marginal effects of the number of children . . . . .	58
2.10	Education selectivity . . . . .	62
2.11	Decompositions within countries over time . . . . .	65
2.12	Decompositions between countries . . . . .	69
2.13	Real FLFP vs. FLFP simulated at Brazil's coefficients . . . . .	70
2.A.1	Distribution of female workforce across industries, by education . . . . .	86
2.A.2	FLFP differences with respect to Brazil: with and without survey weights	88
3.1	Service prevalence by age and sex for selected Northwest European countries	116
3.2	Historical and contemporary female SMAM for 28 European countries .	128
3.3	CW-condition and male-female difference in SMAM . . . . .	129
3.A.1	Jamaica: ages at first marriage. 1970 is an outlier . . . . .	155
3.A.2	World distribution of the CW-condition . . . . .	156
3.A.3	CW index comparison: with and without coastal borders . . . . .	159
4.1	Kernel density plots of the household asset index in 1999 and 2011 . . .	170
4.A.1	Histogram of private education inputs . . . . .	188
4.A.2	Correlation of bribe payments to teachers and migration rate . . . . .	189
4.A.3	Map of communities with above and below median values of the network- growth IV . . . . .	189
4.A.4	Schools' wage bills in BOOST proxy the number of teachers . . . . .	190



# Introduction

I assume that the reader and I agree on a basic premise: that, above all else, equality of opportunities between men and women is an issue of elementary justice and, therefore, an end in itself. Let us not bother John Rawls or Amartya Sen; let us save our veils of ignorance for other, thornier occasions. In this sense, gender equality *is* development. As put by Harriet Martineau, in 1837, “[i]f a test of civilization be sought, none can be so sure as the condition of that half of society over which the other half has power,—from the exercise of the right of the strongest” (Martineau, 1837, p. 105).

Besides being a topic with a clear normative goal, gender inequality is a fascinating field of economic research. This dissertation focuses, in particular, on the relationship between gender equality and economic development, broadly defined. Since the seminal work of Ester Boserup (1970), gender equality is widely understood as being both an *instrument for* and a *consequence of* economic development.<sup>1</sup> The first three chapters of this thesis revolve around two fundamental questions. What are the consequences of gender inequality for economic development? (Chapter 1.) And what are its causes—i.e., why do we observe gender inequality in all societies, although to different extents? (Chapters 2 and 3.) To paraphrase Robert Lucas, “once one starts to think about [these questions], it is hard to think about anything else” (Lucas, 1988, p. 5).<sup>2</sup> As we shall see, social and cultural norms on gender roles will contribute a great deal to how we think about these issues. In the final chapter (Chapter 4), I investigate an episode of rapid change in social norms, in a context of mass emigration. Even though that essay does not deal with gender equality directly, it suggests that migration may be a powerful vehicle for cultural change in developing countries.<sup>3</sup>

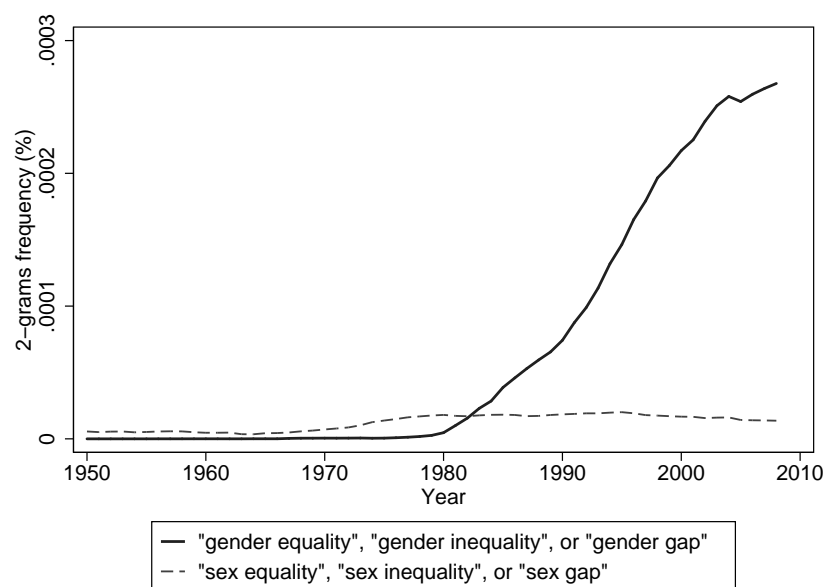
Since 1980, academic work on gender equality has risen dramatically. Figure 1 plots the combined frequency of the three 2-grams “gender equality”, “gender inequality”,

---

<sup>1</sup>For recent surveys of this two-way relationship, see World Bank (2001, 2011); Duflo (2012); Jayachandran (2015).

<sup>2</sup>Lucas’s original remark concerned variation in per capita income across countries.

<sup>3</sup>The transfer of cultural values from migrants back to their countries of origin is usually known as “social remittances”, a term coined by the sociologist Peggy Levitt (Levitt, 1998, 2001). There is growing research in economics documenting that such cultural transfers exist and are quantitatively important (e.g., Batista and Vicente, 2011; Tuccio and Wahba, 2015; Barsbai et al., 2017; Ivlevs and King, 2017).



**FIG. 1:** Rising scholarship on gender equality

*Notes:* Author's calculations from Google Books Ngram Viewer, corpus of English books, 2012 version (persistent identifier: googlebooks-eng-all-20120701). See Michel et al. (2011) for details. Data retrieved from <http://books.google.com/ngrams> [accessed on May 24, 2018]. The full line is the sum of the frequency of the following 2-grams: "gender equality", "gender inequality", and "gender gap". The dashed line is the sum of the frequency of the following 2-grams: "sex equality", "sex inequality", and "sex gap". Searches were case-insensitive. Time period is 1950–2008; 3-year moving averages are shown.

and "gender gap" in a large corpus of English books. From virtually no usage in books before 1980, their frequency has been growing fast ever since. Perhaps, this trend simply reflects a change in terminology, if the term "gender" (i.e., the social identity assigned to each biological sex) was replacing the term "sex" (i.e., the biological difference between men and women). But this is not the case: as shown in the figure, the rise of "gender" equality does not coincide with a decline of "sex" equality.

While scholarly interest on gender equality was rising, developing countries were making extraordinary progress in reducing gender gaps in several socio-economic dimensions.<sup>4</sup> Since 1980, gender gaps in enrollment rates have been eliminated for primary and secondary schooling in all but a few countries, whereas for tertiary education the gap is now in favor of women (e.g., Ganguli et al., 2014; World Bank, 2011, pp. 106–116). Fertility rates also fell dramatically in this period, with the speed of this decline being much faster than in the past experiences of today's advanced economies (World Bank,

<sup>4</sup>During the same period, advanced economies also witnessed progress towards gender equality (e.g., Kleven and Landais, 2017). Claudia Goldin provides a grand narrative for the US (Goldin, 1990, 2006, 2014).

2011, p. 63). Finally, legal constraints on women's rights have been greatly reduced, although not yet fully eliminated (Doepke et al., 2012; Hallward-Driemeier et al., 2013).

Yet, this remarkable progress did not happen in all areas. Gender gaps in labor force participation and time use have closed slowly (World Bank, 2011; Ganguli et al., 2014; Klasen and Pieters, 2015; Klasen, 2018b); whereas the gaps remain nearly unchanged with respect to occupational and sectoral segregation (World Bank, 2011; Borrowman and Klasen, 2017) and unexplained wage differentials (Weichselbaumer and Winter-Ebmer, 2005; Oostendorp, 2009).<sup>5</sup>

In the first chapter, co-authored by Stephan Klasen, we review theories where gender equality is instrumental for economic development. The vast majority of theories reviewed suggest that gender inequality is a barrier to development, particularly over the long run. Among the many plausible mechanisms through which inequality between men and women affects the aggregate economy, the role of women for fertility decisions and human capital investments is particularly important. In general, the fact that women are largely responsible for childrearing and domestic work has profound consequences for the value of women's time. First, if women are excluded from market activities, the time costs of childrearing in terms of foregone earnings are relatively low, and fertility is relatively high. Second, if parents anticipate that the returns to female education are low because women carry the burden of domestic activities, they will invest relatively less in the education of girls. In turn, the resulting gender gap in education creates incentives for adult women to specialize in domestic activities, thereby creating a vicious cycle. In sum, a society with norms inhibiting women's participation in the labor market will have higher fertility and lower human capital levels—and, consequently, worse economic performance—than a more egalitarian society.

Because women's labor supply to the market is a crucial mechanism for the positive developmental consequences of gender equality, it is important to understand what drives female participation in the labor force. In the second chapter, co-authored by Stephan Klasen, Janneke Pieters, and Le Thi Ngoc Tu, we investigate the determinants of labor force participation of urban married women, since the early 2000s, in eight developing and emerging economies: Bolivia, Brazil, India, Indonesia, Jordan, South Africa, Tanzania, and Vietnam. Using large household and labor force surveys, we build a unified empirical framework that allows for comparative analyses across time and space. We find that the coefficients to women's characteristics differ substantially across countries, and this explains most of the between-country differences in participation

---

<sup>5</sup>For overviews of gender inequality and discrimination in the labor market, see, among others, Altonji and Blank (1999); Bertrand (2011); Bertrand and Duflo (2017).

rates. In particular, the relationship between a woman's education and her participation in the labor force varies from being positive and linear (Brazil and South Africa) to being U- or J-shaped (India, Jordan, and Indonesia), or a mixture of both (Bolivia, Vietnam, and Tanzania). Overall, the economic, social, and institutional constraints that shape women's labor force participation are still largely country-specific. Nonetheless, we see that rising education levels and declining fertility consistently increased participation rates, while rising household incomes contributed negatively in relatively poorer countries, indicating that a substantial share of women work out of economic necessity.

In that chapter, we show that if the characteristics of women and their families were equalized across the eight countries, most of the large variation in participation rates would remain. What explains differences in participation between countries is, instead, the structure of the labor market, and society in general, in which married women operate. It is therefore relevant to understand the root causes of these structural differences.

In the third chapter, co-authored by Amy C. Alexander, Stephan Klasen, and Christian Welzel, we argue that part of the answer is hidden in a society's distant past. Reviewing the literature on the deep historic roots of gender inequality, we theorize and provide evidence for a trajectory that (1) originates in a climatic configuration called the "Cool Water" (CW-) condition, leading to (2) late female marriages in preindustrial times, which eventually pave the way towards (3) various gender egalitarian outcomes today. The CW-condition is a specific climatic configuration that combines periodically frosty winters with mildly warm summers under the ubiquitous accessibility of fresh water. It embodies opportunity endowments that significantly reduce fertility pressures. The resulting household formation patterns empowered women and reduced gender inequality.

The existence of cultural persistence does not necessary imply that societies are trapped in paths pre-determined by their history. There are moments of rapid and widespread change, where long-held norms and customary social behaviors suddenly collapse. International migration is one of those forces that transfer norms across countries. In the fourth chapter, co-authored by Lisa Sofie Höckel and Tobias Stöhr, we provide evidence for one such instance of widespread social change caused by international migration: a reduction in petty corruption in the education system of the origin country. The income generated from parental migration can increase funds available for children's education. Therefore, in countries where informal payments to teachers are common, migration could increase petty corruption in education. To test this hypothesis, we investigate the effect of migration on educational inputs. We use an instrumental variable approach on survey data and matched administrative records from Moldova, one of the countries with the highest emigration rates. Contrary to the

positive income effect, we find that the strongest migration-related response in private education expenditure is a substantial decrease in informal payments to public school teachers. Any positive income effect due to migration must hence be overcompensated by some payment-reducing effects. We argue that these effects probably reflect migrants' changing views on the value and acceptability of corruption, which spillover to their country of origin.





# 1 Gender Inequality as a Barrier to Economic Growth: a Review of the Theoretical Literature<sup>\*</sup>

**Abstract:** In this chapter, we survey the theoretical literature investigating the role of gender inequality in economic development. The vast majority of theories reviewed suggest that gender inequality is a barrier to development, particularly over the long run. Among the many plausible mechanisms through which inequality between men and women affects the aggregate economy, the role of women for fertility decisions and human capital investments is particularly important. Yet, we believe the body of theories could be expanded in several directions.

---

<sup>\*</sup>This chapter is co-authored with Stephan Klasen. We gratefully acknowledge funding from the Growth and Economic Opportunities for Women (GrOW) initiative, a multi-funder partnership between the UK's Department for International Development, the Hewlett Foundation and the International Development Research Centre. All views expressed here and remaining errors are our own.

## 1.1 Introduction

Theories of long-run economic development have increasingly relied on two central forces: population growth and human capital accumulation. Both forces depend on decisions made primarily within households: population growth is determined by households' fertility choices (e.g., Becker and Barro, 1988), while human capital accumulation is partially dependent on parental investments in child education and health (e.g., Lucas, 1988).

In an earlier survey of the literature linking family decisions to economic growth, Grimm (2003) laments that “[m]ost models ignore the two-sex issue. Parents are modeled as a fictive asexual human being” (p. 154).<sup>1</sup> Since then, however, economists are increasingly recognizing that gender plays a fundamental role in how households reproduce and care for their children. As a result, many models of economic growth are now populated with men and women. The “fictive asexual human being” is a dying species. In this chapter, we survey this rich new landscape in theoretical macroeconomics, reviewing, in particular, theories where gender inequality affects economic development.

Many articles review the literature on gender inequality and economic growth.<sup>2</sup> Typically, both the theoretical and empirical literature are discussed, but, in almost all cases, the vast empirical literature receives most of the attention. In addition, some of the surveys examine both sides of the two-way relationship between gender inequality and economic growth: gender equality as a cause of economic growth *and* economic growth as a cause gender equality. As a result, most surveys end up only scratching the surface of each of these distinct strands of literature.

There is, by now, a large and insightful body of theories exploring how gender equality affects economic growth. In our view, these theories merit a separate review.<sup>3</sup> Moreover, they have not received sufficient attention in empirical work, which has largely developed independently (see also Cuberes and Teignier, 2014). By reviewing the theoretical literature, we hope to motivate empirical economists in finding new ways of putting these theories to test.

---

<sup>1</sup>See Echevarria and Moe (2000) for a similar complaint that “theories of economic growth and development have consistently neglected to include gender as a variable” (p. 77).

<sup>2</sup>A non-exhaustive list includes World Bank (2001, 2011); Stotsky (2006); Sinha et al. (2007); Duflo (2012); Bandiera and Does (2013); Kabeer and Natali (2013); Cuberes and Teignier (2014); Kabeer (2016); Klasen (2018a).

<sup>3</sup>Doepke and Tertilt (2016) review the theoretical literature that incorporates families on macroeconomic models, without focusing exclusively on models that include gender inequality, as we do. Greenwood et al. (2017), in turn, review the theoretical literature from the opposite direction; they study how macroeconomic models can explain changes in family outcomes. Finally, Doepke et al. (2012) survey the political economy of women's rights, but without focusing explicitly on their impact on economic development.

Although the incorporation of gender in macroeconomic models of economic growth is a recent development, the main “gendered”-ingredients of those models are not new. They were developed in at least two strands of literature. First, since the 1960s, the “new home economics” has applied the analytical toolbox of rational choice theory to decisions being made within the boundaries of the family (see, e.g., Becker, 1960, 1981). A second literature strand, mostly based on empirical work at the micro level in developing countries, described clear patterns of gender-specific behavior within households that differed across regions of the developing world (see, e.g., Boserup, 1970). As we shall see, most of the (micro-founded) macroeconomic models reviewed in this chapter use several analytical mechanisms from the “new home economics”; these mechanisms can typically rationalize several of the gender-specific regularities observed in early studies of developing countries. The growth theorist is then left to explore the aggregate implications for economic development.

The first models that incorporated gender in a theory of economic growth did so at a very aggregated level. The main concern was that men and women were imperfect substitutes in aggregate production, and, as a consequence, gender inequality (as long as determined by non-market processes) would misallocate production factors (e.g., Knowles et al., 2002; Esteve-Volart, 2004). We review these arguments in section 1.2.

Over time, the household became the unit of analysis. The first articles in this tradition modeled the household as a unitary entity with joint preferences and interests, and with an efficient and centralized decision making process.<sup>4</sup> These theories posited how men and women specialize into different activities and how parents interact with their children. Section 1.3 reviews this literature.

From there, the literature has moved to intra-household dynamics. Now, family members are allowed to have different preferences and interests; they bargain, either cooperatively or not, over family decisions.<sup>5</sup> Now, the theorist recognizes power asymmetries between family members and analyzes how spouses bargain over decisions. These articles are surveyed in section 1.4.

A few articles explore how households are formed (Edlund and Lagerlöf, 2006; Tertilt, 2005, 2006). They show how different marriage market institutions and family formation patterns influence gender outcomes and long-run development. We review these theories in section 1.5.

---

<sup>4</sup>Examples of this unitary household approach are Galor and Weil (1996); Zhang et al. (1999); Lagerlöf (2003).

<sup>5</sup>Examples of models with non-unitary households are De la Croix and Vander Donckt (2010); Diebolt and Perrin (2013); Doepke and Tertilt (2014); Prettnner and Strulik (2017); Strulik (2018).

In addition to this descent—from aggregate production factors to households, and then to household members—the analysis has also expanded horizontally, by considering new arenas in which gender inequality has relevant consequences for economic development. Examples are international trade (Seguino, 2000; Blecker and Seguino, 2002), foreign direct investment (Rees and Riezman, 2012), and politics (Besley et al., 2017). Section 1.6 discusses this literature.

The vast majority of theories reviewed suggest that gender inequality is a barrier to economic development, particularly over the long run. In most models, irrespectively of the underlying source of differences between the genders (e.g., biology, socialization, discrimination), the opportunity cost of women's time is lower than that of men. This gender gap in the value of time affects economic growth through two main mechanisms. First, when women's time is relatively low, women will be in charge of childrearing and domestic work in the family. A low value of female time means that children are cheap. Fertility will be high, and economic growth will be low, both because population growth has a direct negative impact on long-run economic performance and because human capital accumulates at a slower pace (through the quantity-quality trade-off). Second, if parents expect low returns to female education, due to women specializing in domestic activities, they will invest relatively less in the education of girls. In the words of Harriet Martineau, one of the first to describe this mechanism, "as women have none of the objects in life for which an enlarged education is considered requisite, the education is not given" (Martineau, 1837, p. 107). In the long run, lower human capital investments (on girls) retard economic development.

We conclude, in section 1.7, by examining the limitations of the current literature and pointing ways forward.

## 1.2 A simple efficiency argument: men and women

In this section, we review three prominent arguments making the case that gender inequality in productive capabilities generates aggregate inefficiencies. According to this view, more equality between men and women leads to static efficiency gains in the short run. Yet, other authors warn that gender gaps in different dimensions interrelate, and addressing a gap in isolation may have ambiguous short term effects on economic performance.

The simplest argument for why gender inequality harms economic growth rests on two premises: (1) men and women are separate inputs in the economy-wide production of goods and services, and (2) each input has positive and diminishing marginal products.

An example of this setup is the Solow-type neoclassical growth model of Knowles et al. (2002), where male and female education are imperfect substitutes in production. A gender gap emerges in the *level* of the education input if men are more educated than women (or vice-versa). A gender gap emerges in the *returns* to the education input if its output elasticity differs between genders, such that, at any education level, the marginal products of education are also different.

In the following, by a reduction of the gender gap, we typically have in mind some sort of re-distribution between the genders to a more egalitarian outcome. Of course, if men are more educated than women, another way of reducing the gender gap would be to increase female education, keeping male education constant. Because average education in society goes up, there is an obvious positive *level* effect on per capita output. But the interesting question is whether, keeping average education constant, smaller gender gaps in education are more conducive to growth than bigger gaps, i.e., whether there is a *distribution* effect.

In the case where output elasticities are the same for both genders, an economy where men and women contribute equally to aggregate production will maximize real output. Because men and women are imperfect substitutes, gender inequalities in how productive capacities are distributed are inefficient. Simply put, if men contribute more than women, the marginal product to the male input will be lower than the marginal product to the female input. Closing these gender gaps (in education, health, capital access, etc) would boost economic growth.

On top of this argument, Knowles et al. (2002) hypothesize that women's output elasticity of education is larger than men's. Although they do not explicitly model why this is the case, they justify the hypothesis with positive externalities of female education in reducing fertility and infant mortality, and improving the quantity and quality of children's education. If the output elasticity of female education is relatively large, a gender gap unfavorable to women reduces per capita output in the long run. In fact, the most efficient outcome would be a gender gap in the opposite direction, i.e., unfavorable to men.

The reasoning of Knowles et al. (2002) can easily be extended to other productive capabilities beyond education, such as health and access to capital. Often, female output elasticities are assumed to be larger than male elasticities, due to intergenerational externalities linked to woman's role as the primary caretaker in the family.

A second related argument for why gender inequality leads to aggregate inefficiency concerns the allocation of talent. Assume that talent is randomly distributed in the population. Then, an economy that curbs women's access to education, market em-

ployment, or certain occupations draws talent from a smaller pool than an economy without such restrictions (Klasen, 2002). Gender inequality can thus be viewed as a distortionary tax on talent (Dollar and Gatti, 1999). Indeed, occupational choice models with heterogeneous talent show that exogenous barriers to women's participation in the labor market and entrepreneurial occupations reduce aggregate productivity and per capita output (Esteve-Volart, 2004; Cuberes and Teignier, 2016, 2018).

Thus, if women have lower education, their marginal return to education would be higher than men's. Similarly, if women are more credit-constrained than men, female returns to capital should be higher than male returns, and so on. The problem with this type of reasoning is that it considers inequalities in separate dimensions as being independent from each other. In many cases, however, these inequalities are complementary (Duflo, 2012; Bandiera and Does, 2013; Kabeer, 2016). For example, if credit-constrained women face weak property rights, are unable to access certain markets, and have mobility and time constraints, then the marginal return to capital may nevertheless be larger for men. Similarly, the return to male education may well be above the female return if demand for female labor is low or concentrated in sectors with low productivity. In sum, "the fact that women face multiple constraints means that relaxing just one may not improve outcomes" (Duflo, 2012, p. 1076). When applied to a particular productive endowment in isolation, the efficiency argument for gender equality may not hold.

A third important economic distortion is discrimination against women in the form of lower wages, holding male and female productivity constant. Cavalcanti and Tavares (2016) estimate the aggregate effects of wage discrimination using a model-based general equilibrium representation of the US economy. In their model, households are unitary and, within the household, women are assumed to be more productive in childrearing than men, so they pay the full time cost of this activity. In the labor market, even though men and women are equally productive, women receive only a fraction of the male wage rate—this is the wage discrimination assumption. Wage discrimination works as a tax on female labor supply. Because women work less than they would without discrimination, there is a negative level effect on per capita output. In addition, there is a second negative effect of wage discrimination operating through endogenous fertility. Since lower wages reduce women's opportunity costs of childrearing, fertility is relatively high, and output per capita is relatively low. The authors calibrate the model to US steady state parameters and estimate large negative output costs of the gender wage gap. Reducing wage discrimination against women by 50 percent would raise per capita income by 35 percent, in the long run.

To sum up, when men and women are imperfect substitutes in production and

women's output elasticity is not smaller than men's, male-bias in production factors causes an efficiency loss for the economy. Likewise, when talent is randomly distributed in the population, barriers to women's education, employment, or occupational choice effectively reduce the pool of talent. And when men and women are equally productive, wage discrimination against women acts as a tax on their labor supply, which, once again, depresses aggregate economic activity. All these channels suggest that more gender equality can have an immediate positive effect on economic growth. However, policies should recognize that gender gaps in separate dimensions complement and reinforce one another. A naïve policy targeting a single gap in isolation is unlikely to have substantial growth effects in the short-run.

### 1.3 Unitary households: parents and children

In this section, we review models built upon unitary households. A unitary household maximizes a joint utility function subject to pooled household resources. Decisions are efficient by construction; the household is effectively a black-box. In this class of models, gender inequality stems from a variety of sources. It is rooted in differences in physical strength (Galor and Weil, 1996; Kimura and Yasui, 2010; Hiller, 2014) or health (Bloom et al., 2015); it is embedded in social norms (Lagerlöf, 2003; Hiller, 2014) or son preference (Zhang et al., 1999). In all of the models, gender inequality is a barrier to long-run economic development.

Galor and Weil (1996) model an economy with three factors of production: capital, physical labor ("brawn"), and mental labor ("brain"). Men and women are equally endowed with brains, but men have more brawn. In economies starting with very low levels of capital per worker, women fully specialize in childrearing because their opportunity cost is lower than men's. Over time, the stock of capital per worker builds up due to exogenous technological progress. The degree of complementarity between capital and mental labor is higher than that between capital and physical labor; as the economy accumulates capital per worker, the returns to brain rise relative to the returns to brawn. As a result, the relative wages of women rise, increasing the opportunity cost of childrearing. This negative substitution effect dominates the positive income effect on the demand for children and fertility falls.<sup>6</sup> As fertility falls, capital per worker accumulates faster creating a positive feedback loop that generates a fertility transition

---

<sup>6</sup>This is not the classic Beckerian quantity-quality trade-off because parents cannot invest in the quality of their children. Instead, the mechanism is built by assumption in the household's utility function. When women's wages increase relative to male wages, the substitution effect dominates the income effect.

and kick starts a process of sustained economic growth.

The model has multiple stable equilibria. An economy starting from a low level of capital per worker is caught in a Malthusian poverty trap of high fertility, low income per capita, and low relative wages for women. In contrast, an economy starting from a sufficiently high level of capital per worker will converge to a virtuous equilibrium of low fertility, high income per capita, and high relative wages for women. Through exogenous technological progress, the economy can move from the low to the high equilibrium.

Gender inequality in labor market access or returns to brain can slow down or even prevent the escape from the Malthusian equilibrium. Wage discrimination or barriers to employment would work against the rise of relative female wages and, therefore, slow down the takeoff to modern economic growth.

The Galor and Weil model predicts how female labor supply and fertility evolve in the course of development. First, (married) women start participating in market work and only afterwards does fertility start declining. Historically, however, in the US and Western Europe, the decline in fertility occurred *before* women's participation rates in the labor market started their dramatic increase. In addition, these regions experienced a mid-twentieth century baby boom which seems at odds with Galor and Weil's theory.

Both stylized facts can be generated by adding home production to the model, as do Kimura and Yasui (2010). In their article, as capital per worker accumulates, the market wage for brains rises and the economy moves through four stages of development. In the first stage, with a sufficiently low market wage, both husband and wife are fully dedicated to home production and childrearing. The household does not supply labor to the market; fertility is high and constant. In the second stage, as the wage rate increases, men enter the labor market (supplying both brawn and brain), whereas women remain fully engaged in home production and childrearing. But as men partially withdraw from home production, women have to replace them. As a result, their time cost of childrearing goes up. At this stage of development, the negative substitution effect of rising wages on fertility dominates the positive income effect. Fertility starts declining, even though women have not yet entered the labor market. The third stage arrives when men stop working in home production. There is complete specialization of labor by gender; men only do market work, and women only do home production and childrearing. As the market wage rises for men, the positive income effect becomes dominant and fertility increases; this mimics the baby-boom period of the mid-twentieth century. In the fourth and final stage, once sufficient capital is accumulated, women enter the market sector as wage-earners. The negative substitution effect of rising female



opportunity costs dominates once again, and fertility declines. The economy moves from a “breadwinner model” to a “dual-earnings model”.

Human capital accumulation plays no role in Galor and Weil (1996) and Kimura and Yasui (2010). Each person is exogenously endowed with a unit of brains. The fundamental trade-off in the two models is between the income and substitution effects of rising wages on the demand for children. When Lagerlöf (2003) adds education investments to a gender-based model, an additional trade-off emerges: that between the quantity and the quality of children.

Lagerlöf (2003) models gender inequality as a social norm: on average, men have higher human capital than women. Confronted with this fact, parents play a coordination game in which it is optimal for them to reproduce the inequality in the next generation. The reason is that parents expect the future husbands of their daughters to be, on average, relatively more educated than the future wives of their sons. Because, in the model, parents care for the total income of their children’s future households, they respond by investing relatively less on daughters’ human capital. Here, gender inequality does not arise from some intrinsic difference between men and women. It is instead the result of a coordination failure: “[i]f everyone else behaves in a discriminatory manner, it is optimal for the atomistic player to do the same” (Lagerlöf, 2003, p. 404).

With lower human capital, women earn lower wages than men and are therefore solely responsible for the time cost of childrearing. But if, exogenously, the social norm becomes more gender egalitarian over time, the gender gap in parental educational investment decreases. As better educated girls grow up and become mothers, their opportunity costs of childrearing are higher. Parents trade-off the quantity of children by their quality; fertility falls and human capital accumulates. However, rising wages have an offsetting positive income effect on fertility because parents pay a (fixed) “goods cost” per child. The goods cost is proportionally more important in poor societies than in richer ones. As a result, in poor economies, growth takes off slowly because the positive income effect offsets a large chunk of the negative substitution effect. As economies grow richer, the positive income effect vanishes (as a share of total income), and fertility declines faster. That is, growth accelerates over time even if gender equality increases only linearly.

The natural next step is to model how the social norm on gender roles evolves endogenously during the course of development. Hiller (2014) develops such a model by combining two main ingredients: a gender gap in the endowments of brawn (as in Galor and Weil, 1996) generates a social norm, which each parental couple takes as given (as in Lagerlöf, 2003). The social norm evolves endogenously, but slowly; it tracks

the gender ratio of labor supply in the market, but with a small elasticity. When the male-female ratio in labor supply decreases, stereotypes adjust and the norm becomes less discriminatory against women.

The model generates a U-shaped relationship between economic development and female labor force participation.<sup>7</sup> In the preindustrial stage, there is no education and all labor activities are unskilled, i.e., produced with brawn. Because men have a comparative advantage in brawn, they supply more labor to the market than women, who specialize in home production. This gender gap in labor supply creates a social norm that favors boys over girls. Over time, exogenous skill-biased technological progress raises the relative returns to brains, inducing parents to invest in their children's education. At the beginning, however, because of the social norm, only boys become educated. The economy accumulates human capital and grows, generating a positive income effect that, in isolation, would eventually drive up parental investments in girls' education.<sup>8</sup> But endogenous social norms move in the opposite direction. When only boys receive education, the gender gap in returns to market work increases, and women withdraw to home production. As female relative labor supply in the market drops, the social norm becomes more discriminatory against women. As a result, parents want to invest relatively less in their daughters' education.

In the end, initial conditions determine which of the forces dominates, thereby shaping long-term outcomes. If, initially, the social norm is very discriminatory, its effect is stronger than the income effect; the economy becomes trapped in an equilibrium with high gender inequality and low per capita income. If, on the other hand, social norms are relatively egalitarian to begin with, then the income effect dominates, and the economy converges to an equilibrium with gender equality and high income per capita.

In the models reviewed so far, human capital or brain endowments can be understood as combining both education and health. Bloom et al. (2015) explicitly distinguish these two dimensions. Health affects labor market earnings because sick people are out of work more often (participation effect) and are less productive per hour of work (productivity effect). Female health is assumed to be worse than male health, implying that women's effective wages are lower than men's. As a result, women are solely responsible for childrearing.<sup>9</sup>

<sup>7</sup>The hypothesis that female labor force participation and economic development have a U-shaped relationship—known as the feminization-U hypothesis—goes back to Boserup (1970). See also Goldin (1995). Recently, Gaddis and Klasen (2014) find only limited empirical support for the feminization-U.

<sup>8</sup>The model does not consider fertility decisions. Parents derive utility from their children's human capital (social status utility). When household income increases, parents want to "consume" more social status by investing in their children's education—this is the positive income effect.

<sup>9</sup>Bloom et al. (2015) build their main model with unitary households, but show that the key conclusions

The model produces two growth regimes: a Malthusian trap with high fertility and no educational investments; and a regime of sustained growth, declining fertility, and rising educational investments. Once wages reach a certain threshold, the economy goes through a fertility transition and education expansion, taking off from the Malthusian regime to the sustained growth regime.

Female health promotes growth in both regimes, and it affects the timing of the takeoff. The healthier women are, the earlier the economy takes off. The reason is that a healthier woman earns a higher effective wage and, consequently, faces higher opportunity costs of raising children. When female health improves, the rising opportunity costs of children reduce the wage threshold at which educational investments become attractive; the fertility transition and mass education periods occur earlier.

In contrast, improved male health slows down economic growth and delays the fertility transition. When men become healthier, there is only a income effect on the demand for children, without the negative substitution effect (because male childrearing time is already zero). The policy conclusion would be to redistribute health from men to women. However, the policy would impose a static utility cost on the household. Because women's time allocation to market work is constrained by childrearing responsibilities (whereas men work full-time), the marginal effect of health on household income is larger for men than for women. From the household's point of view, reducing the gender gap in health produces a trade-off between short-term income maximization and long-term economic development.

In an extension of the model, the authors endogenize health investments, while keeping the assumption that women pay the full time cost of childrearing. Because women participate less in the labor market (due to childrearing duties), it is optimal for households to invest more in male health. A health gender gap emerges from rational household behavior that takes into account how time-constraints differ by gender; assuming taste-based discrimination against girls or gender-specific preferences is not necessary.

Until now, parents invest in their children's human capital for purely altruistic reasons. This is captured in the models by assuming that parents derive utility directly from the quantity and quality of children. This is the classical representation of children as durable consumption goods (e.g., Becker, 1960). In reality, of course, parents may also have egoistic motivations for investing in child quantity and quality. A typical example is that, when parents get old and retire, they receive support from their children. The quantity and quality of children will affect the size of old-age transfers and parents are robust to a collective representation of the household.

internalize this in their fertility and childcare behavior. According to this view, children are best understood as investment goods.

Zhang et al. (1999) build an endogenous growth model that incorporates the old-age support mechanism in parental decisions. Another innovative element of their model is that parents can choose the gender of their children. The implicit assumption is that sex selection technologies are freely available to all parents.

At birth, there is a gender gap in human capital endowment, favoring boys over girls.<sup>10</sup> In adulthood, a child's human capital depends on the initial endowment and on the parents' human capital. In addition, the probability that a child survives to adulthood is exogenous and can differ by gender.

Parents receive old-age support from children that survive until adulthood. The more human capital children have, the more old-age support they provide to their parents. Beyond this egoistic motive, parents also enjoy the quantity and the quality of children (altruistic motive). Son preference is modeled by boys having a higher relative weight in the altruistic-component of the parental utility function. In other words, in their enjoyment of children as consumer goods, parents enjoy "consuming" a son more than "consuming" a girl. Parents who prefer sons want more boys than girls. A larger preference for sons, a higher relative survival probability of boys, and a higher human capital endowment of boys positively affect the sex ratio, because, in the parents' perspective, all these forces increase the marginal utility of boys relative to girls.

Zhang et al. (1999) show that, if human capital transmission from parents to children is efficient enough, the economy grows endogenously. When boys have a higher human capital endowment than girls, and the survival probability of sons is not smaller than the survival probability of daughters, then only sons provide old-age support. Anticipating this, parents invest more on the human capital of their sons than on the human capital of their daughters. As a result, the gender gap in human capital at birth widens endogenously.

When only boys provide old-age support, an exogenous increase in son preference harms long-run economic growth. The reason is that, when son preference increases, parents enjoy each son relatively more and demand less old-age support from him. Other things equal, parents want to "consume" more sons now and less old-age support later. Because parents want more sons, the sex ratio increases; but because each son provides less old-age support, human capital investments per son decrease (such that the gender gap in human capital narrows). At the aggregate level, the pace of human capital

---

<sup>10</sup>This assumption does not necessarily mean that boys are more talented than girls. It can be also interpreted as a reduced-form way of capturing labor market discrimination against women.

accumulation slows down and, in the long run, economic growth is lower. Thus, an exogenous increase in son preference increases the sex ratio, and reduces human capital accumulation and long-run growth (although it narrows the gender gap in education).

In summary, in growth models with unitary households, gender inequality is closely linked to the division of labor between family members. If women's time is less valued, they specialize in childrearing and home production, while men specialize in market work. And precisely due to this division of labor, the returns to female educational investments are relatively low. These household behaviors translate into higher fertility and lower human capital and thus pose a barrier to long-run development.

## 1.4 Intra-household bargaining: husbands and wives

In this section, we review models populated with non-unitary households, where decisions are the result of bargaining between the spouses. There are two broad types of bargaining processes: non-cooperative, where spouses interact in a non-cooperative game that often leads to inefficient outcomes (e.g., Doepke and Tertilt, 2014); and cooperative, where the spouses are assumed to achieve an efficient outcome. In a cooperative model, bargaining is either explicitly modeled as a function of an individual's outside option (namely divorce), or proxied by bargaining weights, taken as exogenous by the spouses (known as collective household models; see Chiappori, 1988, 1992).<sup>11</sup>

When preferences differ by gender, bargaining between the spouses matters for economic growth. If women care more about child quality than men do and human capital accumulation is the main engine of growth, then empowering women leads to faster economic growth (Prettner and Strulik, 2017). If, however, men and women have similar preferences but are imperfect substitutes in the production of household public goods, then empowering women has an ambiguous effect on economic growth (Doepke and Tertilt, 2014).

A separate channel concerns the intergenerational transmission of human capital and woman's role as the main caregiver of children. If the education of the mother matters more than the education of the father in the production of children's human capital, then empowering women will be conducive to growth (Diebolt and Perrin, 2013).

The idea that women might have stronger preferences for child-related expenditures than men can be easily incorporated in a Beckerian model of fertility. The necessary assumption is that women place a higher weight on child quality (relative to child

---

<sup>11</sup>See also Echevarria and Moe (2000), who discuss the advantages of modeling households as non-unitary entities for two-sex models of fertility and human capital accumulation.

quantity) than men do. Prettner and Strulik (2017) build a unified growth theory model with collective households. Men and women have different preferences, but they achieve efficient cooperation based on (reduced-form) bargaining parameters. The authors study the effect of two types of preferences: (i) women are assumed to have a relative preference for child quality, while men have a relative preference for child quantity; and (ii) parents are assumed to have a relative preference for the education of sons over the education of daughters. In addition, it is assumed that the time cost of childcare borne by men cannot be above that borne by women (but it could be the same).

When women have a relative preference for child quality, increasing female empowerment speeds up the economy's escape from a Malthusian trap of high fertility, low education, and low income per capita. When female empowerment increases (exogenously), a woman's relative preference for child quality has a higher impact on household's decisions. As a consequence, fertility falls, human capital accumulates, and the economy starts growing. The model also predicts that the more preferences for child quality differ between husband and wife, the more effective is female empowerment in raising long-run per capita income, because the sooner the economy escapes the Malthusian trap. This effect is not affected by whether parents have a preference for the education of boys relative to that of girls. If, however, men and women have similar preferences with respect to the quantity and quality of their children, then female empowerment does not affect the timing of the transition to the sustained growth regime.

Strulik (2018) goes one step further and endogeneizes why men seem to prefer having more children than women. The reason is a different preference for sexual activity: other things equal, men enjoy having sex more than women.<sup>12</sup> When cheap and effective contraception is not available, a higher male desire for sexual activity explains why men also prefer to have more children than women. In a traditional economy, where no contraception is available, fertility is high, while human capital and economic growth are low. When female bargaining power increases, couples reduce their sexual activity, fertility declines, and human capital accumulates faster. Faster human capital accumulation increases household income and, as a consequence, the demand for contraception goes up. As contraception use increases, fertility declines further. Eventually, the economy undergoes a fertility transition and moves to a modern regime with low fertility, widespread use of contraception, high human capital, and high economic growth. In the modern regime, because contraception is widely used, men's desire for sex is de-

---

<sup>12</sup>There are lots of empirical studies in line with this assumption, which is rooted in evolutionary psychology. See Strulik (2018) for references. There are several other evolutionary arguments for men wanting more children (including with different women). See, among others, Penn and Smith (2007); Mulder and Rauch (2009); von Rueden and Jaeggi (2016).

coupled from fertility. Both sex and children cost time and money. When the two are decoupled, men prefer to have more sex at the expense of the number of children. There is a reversal in the gender gap in desired fertility. When contraceptives are not available, men desire more children than women; once contraceptives are widely used, men desire fewer children than women. If women are more empowered, the transition from the traditional equilibrium to the modern equilibrium occurs faster.

Both Prettner and Strulik (2017) and Strulik (2018) rely on gender-specific preferences. In contrast, Doepke and Tertilt (2014) are able to explain gender-specific expenditure patterns *without* having to assume that men and women have different preferences. They set up a non-cooperative model of household decision making and ask whether more female control of household resources leads to higher child expenditures and, thus, to economic development.<sup>13</sup>

In their model, household public goods are produced with two inputs: time and goods. Instead of a single home-produced good (as in most models), there is a continuum of household public goods whose production technologies differ. Some public goods are more time-intensive to produce, while others are more goods-intensive. Each specific public good can only be produced by one spouse—i.e., time and good inputs are not separable. Women face wage discrimination in the labor market, so their opportunity cost of time is lower than men's. As a result, women specialize in the production of the most time-intensive household public goods (e.g., childrearing activities), while men specialize in the production of goods-intensive household public goods (e.g., housing infrastructure). Notice that, because the household is non-cooperative, there is not only a division of labor between husband and wife, but also a division of decision making, since ultimately each spouse decides how much to provide of his or her public goods.

When household resources are redistributed from men to women (i.e., from the high-wage spouse to the low-wage spouse), women provide more public goods, in relative terms. It is ambiguous, however, whether the *total* provision of public goods increases with the re-distributive transfer. In a classic model of gender-specific preferences, a wife increases child expenditures and her own private consumption at the expense of the husband's private consumption. In Doepke and Tertilt (2014), however, the rise in child expenditures (and time-intensive public goods in general) comes at the expense of male consumption *and* male-provided public goods.

Parents contribute to the welfare of the next generation in two ways: via human capital investments (time-intensive, typically done by the mother) and bequests of physical capital (goods-intensive, typically done by the father). Transferring resources to women

<sup>13</sup>They do not model fertility decisions. So there is no quantity-quality trade-off.

increases human capital, but reduces the stock of physical capital. The effect of such transfers on economic growth depends on whether the aggregate production function is relatively intensive in human capital or in physical capital. If aggregate production is relatively human capital intensive, then transfers to women boost economic growth; if it is relatively intensive in physical capital, then transfers to women may reduce economic growth.

There is an interesting paradox here. On the one hand, transfers to women will be growth-enhancing in economies where production is intensive in human capital. These would be more developed, knowledge intensive, service economies. On the other hand, the positive growth effect of transfers to women increases with the size of the gender wage gap, that is, *decreases* with female empowerment. But the more advanced, human capital intensive economies are also the ones with more female empowerment (i.e., lower gender wage gaps). In other words, in settings where human capital investments are relatively beneficial, the contribution of female empowerment to human capital accumulation is reduced. Overall, Doepke and Tertilt's (2014) model predicts that female empowerment has at best a limited positive effect and at worst a negative effect on economic growth.

Diebolt and Perrin (2013) assume cooperative bargaining between husband and wife, but do not rely on sex-specific preferences or differences in ability. Men and women are only distinguished by different uses of their time endowments, with females in charge of all childrearing activities. In line with this labor division, the authors further assume that only the mother's human capital is inherited by the child at birth. On top of the inherited maternal endowment, individuals can accumulate human capital during adulthood, through schooling. The higher the initial human capital endowment, the more effective is the accumulation of human capital via schooling.

A woman's bargaining power in marriage determines her share in total household consumption and is a function of the relative female human capital of the previous generation. An increase in the human capital of mothers relative to that of fathers has two effects. First, it raises the incentives for human capital accumulation of the next generation, because inherited maternal human capital makes schooling more effective. Second, it raises the bargaining power of the next generation of women and, because women's consumption share increases, boosts the returns on women's education. The second effect is not internalized in women's time allocation decisions; it is an intergenerational externality. Thus, an exogenous increase in women's bargaining power would promote economic growth by speeding up the accumulation of human capital across overlapping generations.



De la Croix and Vander Donckt (2010) contribute to the literature by clearly distinguishing between different gender gaps: a gap in the probability of survival, a wage gap, a social and institutional gap, and a gender education gap. The first three are exogenously given, while the fourth is determined within the model.

By assumption, men and women have identical preferences and ability, but women pay the full time cost of childrearing. As in a typical collective household model, bargaining power is partially determined by the spouses' earnings potential (i.e., their levels of human capital and their wage rates). But there is also a component of bargaining power that is exogenous and captures social norms that discriminate against women—this is the social and institutional gender gap.

Husbands and wives bargain over fertility and human capital investments for their children. A standard Beckerian result emerges: parents invest relatively less in the education of girls, because girls will be more time-constrained than boys and, therefore, the female returns to education are lower in relative terms.

There are at least two regimes in the economy: a corner regime and an interior regime. The corner regime consists of maximum fertility, full gender specialization (no women in the labor market), and large gender gaps in education (no education for girls). Reducing the wage gap or the social and institutional gap does not help the economy escaping this regime. Women are not in labor force, so the wage gap is meaningless. The social and institutional gap will determine women's share in household consumption, but does not affect fertility and growth. At this stage, the only effective instruments for escaping the corner regime are reducing the gender survival gap or reducing child mortality. Reducing the gender survival gap increases women's lifespan, which increases their time budget and attracts them to the labor market. Reducing child mortality decreases the time costs of kids, therefore drawing women into the labor market. In both cases, fertility decreases.

In the interior regime, fertility is below the maximum, women's labor supply is above zero, and both boys and girls receive education. In this regime, with endogenous bargaining power, reducing all gender gaps will boost economic growth.<sup>14</sup> Thus, depending on the growth regime, some gender gaps affect economic growth, while others do not. Accordingly, the policy-maker should tackle different dimensions of gender inequality at different stages of the development process.

In the bargaining models reviewed so far, men are passive observers of women's empowerment. Doepke and Tertilt (2009) set up an interesting political economy model

---

<sup>14</sup>De la Croix and Vander Donckt (2010) show this with numerical simulations, because the interior regime becomes analytically intractable.

of women's rights, where men make the decisive choice. Their model is motivated by the fact that, historically, the economic rights of women were expanded *before* their political rights. Because the granting of economic rights empowers women in the household, and this was done before women were allowed to participate in the political process, the relevant question is why did men willingly share their power with their wives?

Doepke and Tertilt (2009) answer this question by arguing that men face a fundamental trade-off. On the one hand, husbands would vote for their wives to have no rights whatsoever, because husbands prefer as much intra-household bargaining power as possible. But, on the other hand, fathers would vote for their daughters to have economic rights in their future households. In addition, fathers want their children to marry highly educated spouses, and grandfathers want their grandchildren to be highly educated. By assumption, men and women have different preferences, with women having a relative preference for child quality over quantity. Accordingly, men internalize that when women become empowered, human capital investments increase, making their children and grandchildren better-off.

Skill-biased (exogenous) technological progress that raises the returns to education over time can shift male incentives along this trade-off. When the returns to education are low, men prefer to make all decisions on their own and deny all rights to women. But once the returns to education are sufficiently high, men voluntarily share their power with women by granting them economic rights. As a result, human capital investments increase and the economy grows faster.

In summary, gender inequality in labor market earnings often implies power asymmetries within the household, with men having more bargaining power than women. If preferences differ by gender and female preferences are more conducive to development, then empowering women is beneficial for growth. When preferences are the same and the bargaining process is non-cooperative, the implications are less clear-cut, and more context-specific. If, in addition, women's empowerment is curtailed by law (e.g., restrictions on women's economic rights), then it is important to understand the political economy of women's rights, in which men are crucial actors.

## 1.5 Household formation patterns

Two-sex models of economic growth have largely ignored how households are formed. The marriage market is not explicitly modeled: spouses are matched randomly, marriage is universal and monogamous, and families are nuclear. In reality, however, household formation patterns vary substantially across societies, with some of these differences

extending far back in history. For example, Hajnal (1965, 1982) described a distinct household formation pattern in preindustrial Northwestern Europe (usually referred as the “European Marriage Pattern”) characterized by: (i) late ages at first marriage for women, (ii) most marriages done under individual consent, and (iii) neolocality (i.e., upon marriage, the bride and the groom leave their parental households to form a new household). In contrast, marriage systems in China and India consisted of: (i) very early female ages at first marriage, (ii) arranged marriages, and (iii) patrilocality (i.e., the bride joins the parental household of the groom).

Economic historians argue that the “European Marriage Pattern” empowered women, encouraging their participation in market activities and reducing fertility levels. While some view this as one of the deep-rooted factors explaining Northwestern Europe’s earlier takeoff to sustained economic growth (e.g., Hartman, 2004; De Moor and Van Zanden, 2010; Carmichael et al., 2016), others have downplayed the long-run significance of this marriage pattern (e.g., Ruggles, 2009; Dennison and Ogilvie, 2014). Despite this lively debate, the topic has been largely ignored by growth theorists. The few exceptions are Edlund and Lagerlöf (2006) and Tertilt (2005, 2006).

Edlund and Lagerlöf (2006) study how rules of consent for marriage influence long-run economic development. In their model, marriages can be formed according to two types of consent rules: individual consent or parental consent. Under individual consent, young people are free to marry whomever they wish, while, under parental consent, their parents are in charge of arranging the marriage. Depending on the prevailing rule, the recipient of the bride-price differs. Under individual consent, a woman receives the bride-price from her husband, whereas, under parental consent, her father receives the bride-price from the father of the groom.<sup>15</sup> In both situations, the father of the groom owns the labor income of his son and, therefore, pays the bride-price, either directly, under parental consent, or indirectly, under individual consent. Under individual consent, the father needs to transfer resources to his son to nudge him into marrying. Thus, individual consent implies a transfer of resources from the old to the young and from men to women, relative to the rule of parental consent. Redistributing resources from the old to the young boosts long-run economic growth. Because the young have a longer timespan to extract income from their children’s labor, they invest relatively more in the human capital of the next generation. In addition, under individual consent, the reallocation of resources from men to women can have additional positive effects on growth, by increasing women’s bargaining power (see section 1.4), although this

<sup>15</sup>The bride-price under individual consent need not be paid explicitly as a lump-sum transfer. It could, instead, be paid to the bride implicitly in the form of higher lifetime consumption.

channel is not explicitly modeled in Edlund and Lagerlöf (2006).

Tertilt (2005) explores the effects of polygyny on long-run development through its impact on savings and fertility. In her model, parental consent applies to women, while individual consent applies to men. There is a competitive marriage market where fathers sell their daughters and men buy their wives. As each man is allowed (and wants) to marry several wives, a positive bride-price emerges in equilibrium.<sup>16</sup> Upon marriage, the reproductive rights of the bride are transferred from her father to husband, who makes all fertility decisions on his own and, in turn, owns the reproductive rights of his daughters. From a father's perspective, daughters are investment goods; they can be sold in the marriage market, at any time. This feature generates additional demand for daughters, which increases overall fertility, and reduces the incentives to save, which decreases the stock of physical capital. Under monogamy, in contrast, the equilibrium bride-price is negative (i.e., a dowry). The reason is that maintaining unmarried daughters is costly for their fathers, so they are better-off paying a (small enough) dowry to their future husbands. In this setting, the economic returns to daughters are lower and, consequently, so is the demand for children. Fertility decreases and savings increase. Thus, moving from polygyny to monogamy lowers population growth and raises the capital stock in the long run, which translates into higher output per capita in the steady state.

Instead of enforcing monogamy in a traditionally polygynous setting, an alternative policy is to transfer marriage consent from fathers to daughters. Tertilt (2006) shows that when individual consent is extended to daughters, such that fathers do not receive the bride-price anymore, the consequences are qualitatively similar to a ban on polygyny. If fathers stop receiving the bride-price, they save more physical capital. In the long run, per capita output is higher when consent is transferred to daughters.

In summary, the rules regulating marriage and household formation carry relevant theoretical consequences for economic development. While the few studies on this topic have focused on consent rules and polygyny, other features of the marriage market remain largely unexplored.

---

<sup>16</sup>In Tertilt (2005), all men are similar (except in age). Widespread polygyny is possible because older men marry younger women and population growth is high. This setup reflects stylized facts for Sub-Saharan Africa. It differs from models that assume male heterogeneity in endowments, where polygyny emerges because a rich male elite owns several wives, while poor men remain single (e.g., Lagerlöf, 2005; Gould et al., 2008).

## 1.6 Beyond the household: openness, politics, and corruption

In this section, we review theories that explore how gender inequality affects economic growth in three areas: small open economies, politics, and corruption.

**Opening to the world** All the models reviewed so far considered closed economies. In open economies, however, gender inequality can interact with trade and international capital flows.

Seguino (2000) argues that wage discrimination against women promotes economic growth in countries where exports are the main engine of growth and where the export sector is female-intensive. Higher wage discrimination against women—i.e., an increase in the gender wage gap that is unrelated to productivity differences between the genders—increases the country's export competitiveness. But for an increase in wage discrimination not to result in women leaving to other sectors, a sufficient degree of job segregation is needed, such that women are effectively “trapped” in the export sector. Blecker and Seguino (2002) formalize this argument in a short-run demand-side model. When the price elasticity of export demand is relatively large, a reduction in the gender wage gap will hurt export-led growth.

Rees and Riezman (2012) model the effect of globalization on economic growth, through the impact of foreign direct investment on gender equality. Men and women differ in their preferences for child quality, which are relatively higher for women. Women are also assumed to bear the full burden of childrearing. Husbands and wives bargain cooperatively, with the bargaining power of each spouse being a positive function of his or her wage rate. Globalization creates job opportunities in a high productivity sector (factory) for either men or women, who otherwise work in a low productivity sector (farm). If globalization creates job opportunities for women, their bargaining power increases and households trade off child quantity by child quality. Fertility falls, human capital accumulates, and long-run per capita output is high. If, on the other hand, globalization creates jobs for men, their intra-household power increases; fertility increases, human capital decreases, and steady-state income per capita is low. Thus, whether global capital flows generate jobs primarily in female or male intensive sectors matters for long-run growth.

**Women in politics** The extent to which women participate in representative politics can affect economic growth through three different channels: the provision of public goods, role model effects, and politician quality.

The first channel concerning the provision of public goods hinges on the assumption that men and women have different preferences on public expenditures. Women would rather spend on goods that they would otherwise produce within the household. In the case that the public goods preferred by women are more conducive to development, or in the case that they liberate women's time for market activities (relative to male-preferred public goods), then increasing female representation among policy makers would foster economic growth (e.g., Stotsky, 2006; Duflo, 2012).

Besides gender-specific preferences in the *composition* of public expenditures, another common assumption is that women prefer a higher *level* of public spending, as well as a larger role of the government in redistributing income and providing social security. In this case, whether a smaller gender gap in political representation improves economic performance depends on whether a larger government is good or bad for growth, which might be highly context specific. In any case, if these political preferences are driven by women's lower economic status, then more gender equality in the economic domain reduces the importance of gender inequality in politics. If, on the other hand, differences in political behavior between men and women reflect differences in deep-rooted preferences, then the effect of empowering women as political actors is independent of gender equality in other economic and social dimensions.

The insights from Doepke and Tertilt's (2014) household model can be used to explore the effects of different demand for public goods at the political level. Recall that in their model no gender-specific preferences are needed; behavioral differences are driven by the time allocation of home production. If women provide household public goods that are time intensive, while men provide household public goods that are goods-intensive, then control over resources at the government level may affect the provision of public goods. Female politicians are likely to invest relatively more in female-provided household public goods; the opposite occurring for male politicians. The effects on economic growth of empowering women in the political arena may depend on whether the production function technology is intensive in human capital or physical capital.

Another, more indirect, channel concerns role model or aspiration effects. If female empowerment at the household level is conducive to development, then female politicians might have a positive effect on long-run growth, if they contribute to more gender equality in the next generations—either through role model effects that increase the aspiration of young girls, or through a reduction of social biases against women in general (Beaman et al., 2009).

A third channel concerns exogenous increases in female representation through the use of gender quotas for political positions. One of the most common arguments against

gender quotas in politics is that they may decrease the average quality of politicians, if lower quality women replace higher quality men. Besley et al. (2017), however, develop a model of local party politics where the opposite occurs. The introduction of a gender quota leads to an increase in average politician quality because, among *male* politicians, lower quality men are replaced by men of higher quality.

According to the model, the leader of a local party faces a fundamental trade-off. The higher the quality (*competence*) of the party's candidates, the more likely the party is to win the election, but the less likely is the leader to survive a contest from the party's candidates. In other words, the higher the quality of the remaining candidates, the tougher is the internal competition faced by the leader. Leaders come in two types, high or low quality, depending on whether their competence is above or below the median competence of party members. Leaders derive utility from ego rents (that is why they want to be leaders) and from the party winning the election (irrespective of being the leader). The result of the trade-off between electoral success and internal survival is that high quality leaders choose higher quality candidates, while low quality leaders choose lower quality candidates.

When gender is added to the model, an additional trade-off emerges. Now, voters care both about candidate quality *and* gender equality. They prefer parties with equal gender representation in their ballots. Then, a male leader who picks a higher share of female candidates increases the electoral outcomes of his party. On the other hand, a higher share of female candidates threatens the survival of male leaders—the assumption here being that female candidates prefer a female leader, because policy preferences are, to a certain extent, gender-specific. As a result, male leaders face a gendered trade-off between electoral success and internal survival. A higher share of women candidates increases the former but decreases the latter.<sup>17</sup>

When a leader is forward-looking, he will resign before the election if he predicts that a new leader has a higher chance of winning. Thus, low quality leaders will have an incentive to resign if they predict that a higher quality leader will replace them. For the high quality leader, this incentive is weaker. Since the high quality leader can pick higher quality candidates (due to a relatively smaller threat from internal competition), he will have a better chance of winning the election.

Against this background, the gender quota exogenously increases the threat to survival of all male leaders. Low quality leaders will be the first to resign, because they face the

---

<sup>17</sup>A gender gap in political representation emerges endogenously. Because a female candidate threatens the survival of a male leader more than an equally competent male candidate would, the optimal share of female candidates is below one-half. That is why a gender quota is needed.

largest threat in any case. When they resign, they are, on average, replaced by more competent leaders, who then select more competent candidates. Average politician quality goes up.<sup>18</sup>

The mechanisms of Besley et al. (2017) rest of the assumption of a democratic electoral process with (1) voters deriving utility only from politician competence and gender equality in candidate representation, (2) party leaders facing internal competition from other party candidates (they are all competing for ego rents), (3) more competent candidates posing a greater survival threat to a party leader, and (4) a higher share of women candidates posing a greater survival threat to a male leader. On the other hand, in less democratic electoral processes, in places where gender equality is less valued by society, or when party structures are highly centralized and not really open to internal dispute—in all these contexts—the mechanisms of the model would break down.

**Women and corruption** Women's underrepresentation in leadership positions, either in politics or business, may also affect economic performance via its effect on corruption. There is suggestive evidence that women engage in less corruption than men (e.g., Dollar et al., 2001; Swamy et al., 2001; Beaman et al., 2009; Brollo and Troiano, 2016).<sup>19</sup>

Swamy et al. (2001) review several hypotheses explaining this gender difference in corruption with factors that can be expected to persist over time. Women may avoid corruption because they are more risk-averse than men, or because they are more honest—since honesty is a trait they want to pass on to their children (for whose rearing they are mainly responsible)—, or because they put a greater preference on obeying the law—since the law disproportionately benefits the physically weak.

But an alternative set of explanations involves factors that result from women's historical underrepresentation in positions of power. Women may have fewer *opportunities* to engage in corruption (Goetz, 2007). For example, they may be excluded from corruption networks or have less knowledge on how these operate. These differences are likely to erode, as female representation increases over time, and powerful women become exposed to (and familiar with) corruption practices. Thus, the underlying cause for the gender-differential in corruption will determine whether increasing women's representation will reduce corruption in the short or in the long run.

<sup>18</sup>In the model, resignation is voluntary; it comes from the leader trading off ego rents with the utility from the party winning the election. But, in addition, there could be "social pressure" on low quality leaders after the introduction of a quota. Such pressure would reinforce the model's conclusions.

<sup>19</sup>There is a broader debate in economics about the effect of corruption on economic growth. The controversy is on whether corruption "greases" or "sands" the wheels of economic growth. See, among others, Shleifer and Vishny (1993); Bardhan (1997); Méon and Sekkat (2005).



In summary, if indeed women engage in less corruption than men, it is important to know the underlying cause for this gender difference. If the difference is driven by evolutionary or socialization forces, then more women in leadership positions will likely reduce corruption in the long run. If, on the other hand, the difference stems from a history of underrepresentation, the positive impact of more female leaders is likely to be short-lived.

## 1.7 Conclusion

In this chapter, we surveyed the theoretical literature linking gender inequality to economic development. This literature offers many plausible mechanisms through which inequality between men and women affects the aggregate economy. Yet, we believe the body of theories could be expanded in several directions. We discuss them below and finish by suggesting ways in which the dialogue between theory and empirics on this topic can be improved.

The first direction for future research concerns control over fertility. In models where fertility is endogenous, households are always able to achieve their preferred number of children (see Strulik, 2018, for an exception). The implicit assumption is that there is a free and infallible method of fertility control available for all households—a view rejected by most demographers. The gap between desired fertility and achieved fertility can be endogeneized at two levels. First, at the societal level, the diffusion of particular contraceptive methods may be influenced by cultural and religious norms. Second, at the household level, fertility control may be object of non-cooperative bargaining between the spouses, in particular, for contraceptive methods that only women perfectly observe (Ashraf et al., 2014; Doepke and Kindermann, 2016). More generally, the role of asymmetric information within the household is not yet explored.

A second direction worth exploring concerns gender inequality in a historical perspective. In models with multiple equilibria, an economy's path is often determined by its initial level of gender equality. Therefore, it would be useful to develop theories explaining why initial conditions varied across societies. In particular, there is a large literature on economic and demographic history documenting how systems of marriage and household formation differed substantially across preindustrial societies (e.g., Hajnal, 1965, 1982; Hartman, 2004; Ruggles, 2009; De Moor and Van Zanden, 2010). In our view, more theoretical work is needed to explain both the origins and the consequences of these historical systems.

A third avenue for future research concerns the role of technological change. In

several models, technological change is the exogenous force that ultimately erodes gender gaps in education or labor supply (e.g., Galor and Weil, 1996; Doepke and Tertilt, 2009; Bloom et al., 2015). For that to happen, technological progress is assumed to be skill-biased, thus raising the returns to education—or, in other words, favoring brain over brawn. As such, new technologies make male advantage in physical strength ever more irrelevant, while making female time spent on childrearing and housework ever more expensive. Moreover, recent technological progress increased the efficiency of domestic activities, thereby relaxing women's time constraints (e.g., Greenwood et al., 2005; Cavalcanti and Tavares, 2008). These mechanisms are plausible, but other aspects of technological change need not be equally favorable for women. In many countries, for example, the booming science, technology, and engineering sectors tend to be particularly male-intensive.

Even if current technological progress is assumed to weaken gender gaps, historically, technology may have played exactly the opposite role. If technology today is more complementary to brain, in the past it could have been more complementary to brawn. An example is the plow that, relative to alternative technologies for field preparation (e.g., hoe, digging stick), requires upper body strength, on which men have a comparative advantage over women (Boserup, 1970; Alesina et al., 2013). Another, even more striking example, is the invention of agriculture itself—the Neolithic Revolution. The transition from a hunter-gatherer lifestyle to sedentary agriculture involved a relative loss of status for women (Dyble et al., 2015; Hansen et al., 2015). One explanation is that property rights on land were captured by men, who had an advantage on physical strength and, consequently, on physical violence. Thus, in the long view of human history, technological change appears to have shifted from being male-biased towards being female-biased. Endogeneizing technological progress and its interaction with gender inequality is a promising avenue for future research.

A final point concerns the role of men in this literature. In most models, gender inequality is not the result of an active male project that seeks the domination of women. Instead, inequality emerges as a rational best response to some underlying gender gap in endowments or constraints. Then, as the underlying gap becomes less relevant—for example, due to skill-biased technological change—, men passively relinquish their power (see Doepke and Tertilt, 2009, for an exception). There is never a male backlash against the short-term power loss that necessarily comes with female empowerment. In reality, it is more likely that men actively oppose losing power and resources towards women (Kabeer, 2016). This possibility has not yet been explored in formal models, even though it could threaten the typical virtuous cycle between gender equality and growth.

If men are forward-looking, and the short run losses outweigh the dynamic gains from higher growth, they might ensure that women never get empowered to begin with. For example, Eswaran and Malhotra (2011) set up a household decision model where men use domestic violence against their wives as a tool to enhance male bargaining power. Thus, future theories should recognize more often that men have a vested interest on the process of female empowerment.

Turning now to the empirical literature, we notice two main challenges. First, most empirical studies focus on the Solow-model type of efficiency arguments, thus estimating reduced-form aggregate differences in the output elasticities of male and female inputs. At the same time, there is a large gap in the empirical literature for studies testing the type of household-based mechanisms that are at the heart of most theories surveyed in this article. For example, in a recent survey of the deep determinants of long-run economic development, gender inequality is not mentioned (Spolaore and Wacziarg, 2013). Second, the theoretical literature suggests that the timing of effects vastly differs. For some mechanisms, such as talent misallocation or export-competitiveness, the effects of gender inequality are quite short-term. But for other chains of causality, such as fertility and human capital accumulation, some of the effects are intergenerational. Empirical work needs to consider these different timings more explicitly.



## 2 What Drives Female Labor Force Participation? Comparable Micro-level Evidence from Eight Developing and Emerging Economies\*

**Abstract:** We investigate the micro-level determinants of labor force participation of urban married women in eight low- and middle-income economies: Bolivia, Brazil, India, Indonesia, Jordan, South Africa, Tanzania, and Vietnam. In order to understand what drives changes and differences in participation rates since the early 2000s, we build a unified empirical framework that allows for comparative analyses across time and space. We find that the coefficients of women's characteristics differ substantially across countries, and this explains most of the between-country differences in participation rates. In particular, the relationship between a woman's education and her participation in the labor force varies from being positive and linear (Brazil and South Africa) to being U- or J-shaped (India, Jordan, and Indonesia), or a mixture of both (Bolivia, Vietnam, and Tanzania). Overall, the economic, social, and institutional constraints that shape women's labor force participation remain largely country-specific. Nonetheless, rising education levels and declining fertility consistently increased participation rates, while rising household incomes contributed negatively in relatively poorer countries, suggesting that a substantial share of women work out of economic necessity.

---

\*This chapter is co-authored with Stephan Klasen, Janneke Pieters, and Le Thi Ngoc Tu. We are grateful to Esther Heesemann, Lisa Höckel, Bruno Witzel-Souza, and participants at the 26<sup>th</sup> IAFPE Conference and seminars of the Universities of Goettingen, Hannover, and Stellenbosch for comments and suggestions. For assistance with the Tanzanian data, we thank Novati Buberwa, James Mbongo, and Titus Mwisomba from Tanzania's National Bureau of Statistics. For assistance with the Indonesian data, we thank Krisztina Kis-Katos, Christoph Kubitz, and Robert Sparrow. Friederike Schilling provided excellent research assistance. We gratefully acknowledge funding from the Growth and Economic Opportunities for Women (GrOW) initiative, a multi-funder partnership between the UK's Department for International Development, the Hewlett Foundation, and the International Development Research Centre. The views expressed herein do not necessarily represent those of IDRC or its Board of Governors.

## 2.1 Introduction

Worldwide, the current labor force participation rate for women (age 15+) stands at 49 percent, compared to a participation rate of 76 percent for men (ILO, 2017). In the developing world, recent progress in closing this gender gap has been disappointing. In the past two decades, female labor force participation (FLFP, henceforth) rates have increased only modestly, on average, though there is considerable heterogeneity across countries and regions. Female participation rates are lowest in the Middle East and North Africa and in South Asia. South Asia also performed worst in terms of trends, with a declining share of women in the labor force. In contrast, female participation rates increased substantially in Latin America and the Caribbean.

The heterogeneity in female participation rates is observed against a background of rising female education, declining fertility, and robust economic growth in almost all developing countries. Women in developing countries have been accumulating skills at an unprecedented pace, while declining fertility reduced the burdens of childrearing and domestic work. Combined with economic growth, one would expect more educated and less time constrained women to enter an expanding labor market. Even if long held gender norms on women working outside the home fail to adjust as quickly, rising opportunity costs in foregone earnings should eventually boost women's participation rates. But this expectation did not materialize everywhere.

In this chapter, we use comparable microdata from eight low and middle-income economies—Bolivia, Brazil, India, Indonesia, Jordan, South Africa, Tanzania, and Vietnam—to analyze how women's individual and household characteristics are associated with FLFP and what are the key commonalities and differences across countries. The period covered is 2000–2014. We further ask which factors drive FLFP changes over time *within* countries, and which factors account for differences in FLFP rates *between* countries.

A large literature studies FLFP in the developing world (see Klasen, 2018b, for a review). At the macro level, the feminization-U hypothesis posits that, at low income levels, FLFP declines with economic development but, at some point, as countries get richer, the relationship turns positive (Boserup, 1970; Goldin, 1990, 1995). However, Gaddis and Klasen (2014) find only weak empirical evidence in support of this hypothesis in a large panel of countries. Instead, countries' idiosyncratic factors explain most of the worldwide variation in FLFP. Similarly, there is no evidence, at the country level, that closing the gender gap in education reduces the gender gap in labor force participation

(Ganguli et al., 2014).<sup>1</sup> Studying 101 countries over a long period of time, Aaronson et al. (2017) find large negative effects of fertility on mothers' labor supply, but only for sufficiently rich countries. At low levels of income, however, the effect of fertility is either small or zero.<sup>2</sup>

In work closely related to ours, Gasparini and Marchionni (2015) analyze microdata from 18 Latin American countries to investigate changes in FLFP between 1992 and 2012. They conclude that increased education, reduced marriage and fertility, and structural change towards more female-intensive activities contributed significantly to rising female participation throughout this period. However, these factors cannot account for the slowdown in the growth of female labor supply since the 2000s, which the authors link to the decade's strong economic growth. By improving overall conditions, economic growth "may have reduced the urgency of vulnerable women [rural, low educated, with children and low-earnings spouses] to take low quality jobs" (Gasparini and Marchionni, 2015, p. 13).

Several other papers investigate recent trends in FLFP for single countries. Assaad et al. (2014) offer a demand-side explanation for stagnating female participation rates in Jordan since 2000. As public sector hiring tightened since the adjustment policies of the 1980s, so have women's labor market opportunities; the reason being that women are disproportionately employed in education and health activities.<sup>3</sup> In Vietnam, very high FLFP is typically explained by the country's socialist legacy<sup>4</sup>, and, to a smaller extent, by excess male mortality during the Vietnam War (Kreibaum and Klasen, 2015). For South Africa, Ntuli and Wittenberg (2013) decompose the increase in the participation rate of black women from 1995 to 2004. They find that changing returns to women's labor market characteristics account for most of the FLFP increase. Klasen and Pieters (2015) ask why FLFP stagnated in India since the late 1980s and show that rising incomes and

---

<sup>1</sup>Ganguli et al. (2014) analyze census data from 40 countries. At the micro level, the authors show that if the education gender gap, the marriage gap (LFP gap between married and single women), and the motherhood gap (LFP gap between mothers and childless women) were to close everywhere, a large unexplained gender gap in participation rates would still remain for most countries. However, Ganguli et al. (2014) assume that education and FLFP are linearly related. As we will show in this chapter, the shape of the education-participation relationship is nonlinear in some countries.

<sup>2</sup>Aaronson et al. (2017) instrument fertility with twin birth (Rosenzweig and Wolpin, 1980) and sibling sex composition (Angrist and Evans, 1998). Using infertility shocks as a different source of exogenous variation for 26 developing countries, Agüero and Marks (2011) find no effect of fertility on mothers' labor force participation. Priebe (2010) argues that, in poor settings, child costs push women into the labor market; as fertility declines, this type of distress-driven FLFP falls. The author shows causal evidence of this mechanism operating in Indonesia.

<sup>3</sup>In the Jordanian context, jobs in public education and health are among the few deemed socially appropriate for married women.

<sup>4</sup>See Ganguli et al. (2014, p. 184) and Klasen (2018b, p. 15) for further evidence.

male education levels reduced married women's labor supply. Rising female education, on the other hand, contributed less than expected due to a U-shaped relationship between a woman's education and her labor force participation. They also point at the lack of employment growth in manufacturing and white-collar services as a factor obstructing women's entry into the labor force.<sup>5</sup> The correlates of FLFP in Indonesia resemble those in India, including a U-shaped education-participation relationship (Schaner and Das, 2016).

In this chapter, we draw on labor force and household surveys covering roughly the period 2000–2014. The population of interest consists of prime age (25–54) urban married women.<sup>6</sup> The final dataset contains nearly 800,000 women from 32 surveys across eight countries. We estimate country- and year-specific determinants of FLFP using a unified empirical framework based on Klasen and Pieters (2015). The covariates capture two groups of supply-side factors: *women's own characteristics*—education, age, ethnicity or religion— and *family circumstances*—household income, education of the household head, presence of a man with salaried employment (to capture income security), and number of children aged 0–4 and 5–14 in the household. Our estimates are best understood as reduced form correlates. We do not model own wage effects, due to the well-known lack of robustness of current identification strategies (e.g., Klasen and Pieters, 2015)<sup>7</sup>, but capture labor demand conditions by region fixed effects. Using the estimates from the FLFP models, we first decompose changes in participation rates over time *within* countries, and then decompose differences in participation rates *between* countries.

Our approach has several advantages. We have richer data than in typical macro-level cross-country analyses, allowing us to study heterogeneous effects across space and time in much detail. In addition, the unified empirical framework allows us to draw direct comparisons between countries and over time. In that sense, we contribute to a diverse collection of country case studies whose different methodologies and populations of interest preclude systematic comparisons. By establishing FLFP correlates within a unified empirical framework, over large samples and several periods, our study provides global stylized facts on the impact and relative importance of what are considered key determinants of FLFP.

<sup>5</sup>Since then, similar analyses were conducted for Bangladesh (Rahman and Islam, 2013; Mahmud and Bidisha, 2016) and Sri Lanka (Seneviratne, 2017), but since the FLFP series for Bangladesh suffer from severe data inconsistencies (Klasen, 2018b, p. 4), results should be taken with caution.

<sup>6</sup>Our definition of currently married couples includes cohabitation, i.e., those living together as husband and wife even if not formally married.

<sup>7</sup>Blundell and Macurdy (1999) review the standard neoclassical model of labor supply, its extensions, and econometric applications. Blundell et al. (2007) discuss in detail identification and estimation of labor supply models.



We first analyze the country- and year-specific correlates of labor force participation, producing three key findings. First, there is no universal relationship between a woman's educational attainment and her likelihood of being in the labor force. Instead, we find two types of patterns: (i) a strong positive, linear relationship in the two richest countries—Brazil and South Africa; (ii) a U- or J-shaped relationship in India, Indonesia, and Jordan, where, relative to that of the lowest educated women, the participation probability does not change or even *decreases* at intermediate education levels (typically, secondary schooling), and then *increases* substantially at higher attainment levels. In Bolivia, Tanzania, and Vietnam, the education-participation relationship combines features from both stylized patterns.

Second, fertility penalizes women's participation everywhere, but the effect is stronger in richer countries. The number of young children (ages 0–4) correlates negatively with labor force participation in all countries, but the number of older children (ages 5–14) only correlates negatively with participation in the four richest countries: Indonesia, Jordan, South Africa, and Brazil. These findings are consistent with causal evidence on fertility effects (e.g., Aaronson et al., 2017).

Third, we find that in the two richest countries (Brazil and South Africa) households' economic conditions—captured by household income, education of the household head, and presence of a man with salaried employment—do not correlate with FLFP. The negative income effect shrinks over time (in absolute term), disappearing by 2014 in both countries. This finding is in line with evidence of plummeting income effects on married women's labor supply in the US (Blau and Kahn, 2007; Heim, 2007). In the remaining countries, especially India, Indonesia, and Bolivia, higher household income and household head education are still strongly negatively related to FLFP.

When decomposing trends in FLFP, we find that rising female education and falling fertility contributed to increases in participation everywhere. Yet, the magnitude of these contributions differs substantially across countries, mainly reflecting differences in the participation returns to education and the effect of children. In all countries but the richest three (Jordan, South Africa, and Brazil), these positive contributions were offset by the negative effect of rising household income. The strength of this negative income effect suggests that, in poorer countries, a substantial share of women work out of economic necessity, leaving the labor force as soon as it becomes affordable. Lastly, we find, for several countries, a relatively strong (negative or positive) contribution from factors that are not explained by our model and reflect instead changes in coefficients and unobservables. The sign and size of this unexplained contribution does not appear to relate to the country's income level, or the observed level or change in FLFP rates.

In the final part of our analysis, we decompose FLFP differences *between* countries. We find that differences in covariates cannot explain gaps in participation rates between countries. Instead, the returns to covariates and unobservables account for the bulk of FLFP variation, both around 2000 and 2014. Thus, economic, social, and institutional constraints that shape women's labor force participation are still largely country-specific.

This chapter proceeds as follows. Section 2.2 presents the data, descriptive statistics, and the empirical model. Section 2.3 shows the estimation results. In section 2.4, we decompose labor force participation changes over time and between countries. Section 2.5 concludes.

## 2.2 Data and empirical model

In section 2.2.2, we describe the data sources and present descriptive statistics for the main variables of interest. The empirical model is discussed in section 2.2.3.

### 2.2.1 Data

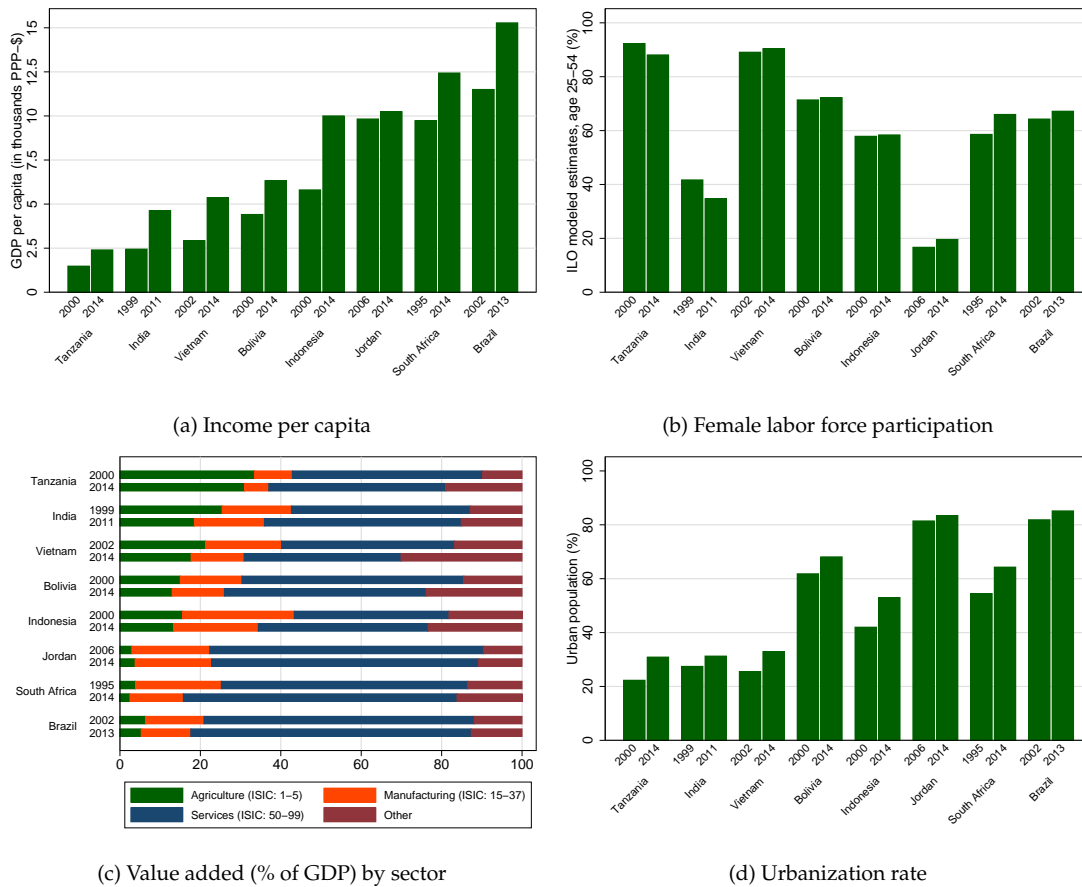
We select eight non-OECD countries with available good-quality large-scale household surveys allowing us to derive (most of) the variables used in Klasen and Pieters (2015). We purposefully choose a diverse group of countries: two upper middle income countries—Brazil and South Africa—, five lower middle income countries—Bolivia, India, Indonesia, Jordan, and Vietnam—, and one low income country—Tanzania. These countries cover a wide range of geographies, per capita incomes, FLFP rates, economic structures, and urbanization rates (Figure 2.1).<sup>8</sup>

When compared to the world, India and Jordan have less gender equality and lower FLFP than predicted by their income levels (Figure 2.2). In contrast, Tanzania and Vietnam have more gender equality and higher FLFP than predicted by income alone. For the remaining countries (Bolivia, Brazil, Indonesia, and South Africa), their relative position in the world income distribution predicts well the levels of gender equality and FLFP.

The data cover roughly the past one-and-a-half decades, from the early 2000s to the mid-2010s, with the exception of Jordan, whose available time-span is shorter: 2006–2014. For South Africa, we also include 1995 in some of our analyses, to cover the entire post-*apartheid* era.

---

<sup>8</sup>Figure 2.1 shows data for the first and last year available for each country in our dataset. We obtain similar patterns if we plot data in 2000 and 2014 for all countries.



**FIG. 2.1:** Selected country indicators for the first and last years in our dataset

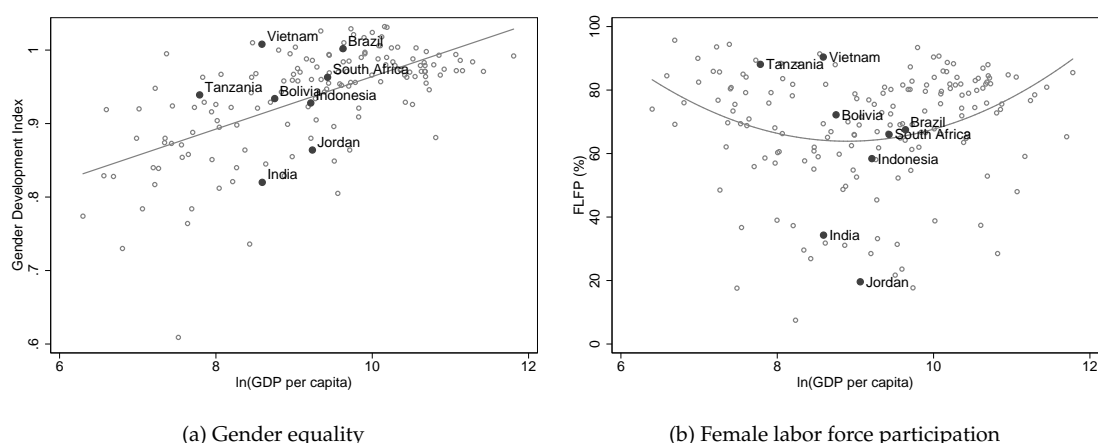
*Notes:* Sources are ILOSTAT and World Development Indicators. Countries are sorted by income per capita.

At the macroeconomic level, 2000–2014 was a period of sustained economic growth. For the eight countries, GDP per capita grew, on average, 3.2 percent per year. India and Vietnam were the best performers, with average annual growth rates of 5.3 and 5.1 percent. South Africa grew the slowest: 1.6 percent per year.<sup>9</sup> In general, our survey-years are representative of this macroeconomic period. Of 32 surveys, only two took place during recessions: Brazil, 2009, and Jordan, 2010.

We only consider urban households for two reasons.<sup>10</sup> First, the analysis requires individual earnings which are difficult to measure in rural areas, given the importance of

<sup>9</sup>Note that South Africa's GDP per capita grew much faster between 1995 and 2014, which is the period shown in Figure 2.1.

<sup>10</sup>For Jordan we consider both urban and rural areas because information on urban status is not available from the 2008 and 2014 surveys. In any case, more than 80 percent of Jordan's population lives in urban areas, in the period considered, according to data from the World Bank's World Development Indicators.



**FIG. 2.2:** Selected gender indicators with respect to per capita income in 2014

*Notes:* Sources are UNDP's Human Development Report 2016 [Panel (a)] and ILOSTAT [Panel (b)]. Panel (a)—Linear fit of the two variables shown. 154 countries included. Panel (b)—Quadratic fit of the two variables shown. 174 countries included. FLFP, for ages 25–54, is a ILO modeled estimate. GDP per capita is PPP-adjusted at 2011 international \$.

smallholder agriculture. Non-marketed agricultural output must be valued in monetary terms, but the necessary detailed price data is often unavailable.<sup>11</sup> Moreover, whenever several household members farm the same plots, or agricultural income is aggregated at the household level, it is unavoidable to impute income for each individual. In urban areas, measurement error or missing data are less severe.

Second, in settings dominated by agriculture, where many women contribute to household farming, household surveys are more likely to underreport female work. The extent of underreporting likely depends on survey methodology, which varies across countries (and sometimes within countries over time).<sup>12</sup> Focusing on urban areas, therefore, improves the comparability of labor force measurements across space and time.

The dataset includes nearly 800,000 urban married women of age 25–54. Table 2.A.1 lists the surveys, years, and sample sizes. In Appendix 2.A, we describe each data source in detail, explaining how variables were constructed and harmonized across surveys.

<sup>11</sup>For example, in its 2000 and 2006 rounds, Tanzania's Integrated Labor Force Survey only recorded agricultural income in urban areas. Other well-known practical complications are unmeasured product variety and quality.

<sup>12</sup>For example, in South Africa, the Labor Force Survey (LFS 2001–2007) is better at capturing informal casual employment than the previous October Household Survey (OHS 1995–1999) (Yu, 2007). The number of employment categories in the survey questionnaire increased from three in the OHS to eight in the LFS.

## 2.2.2 Descriptives

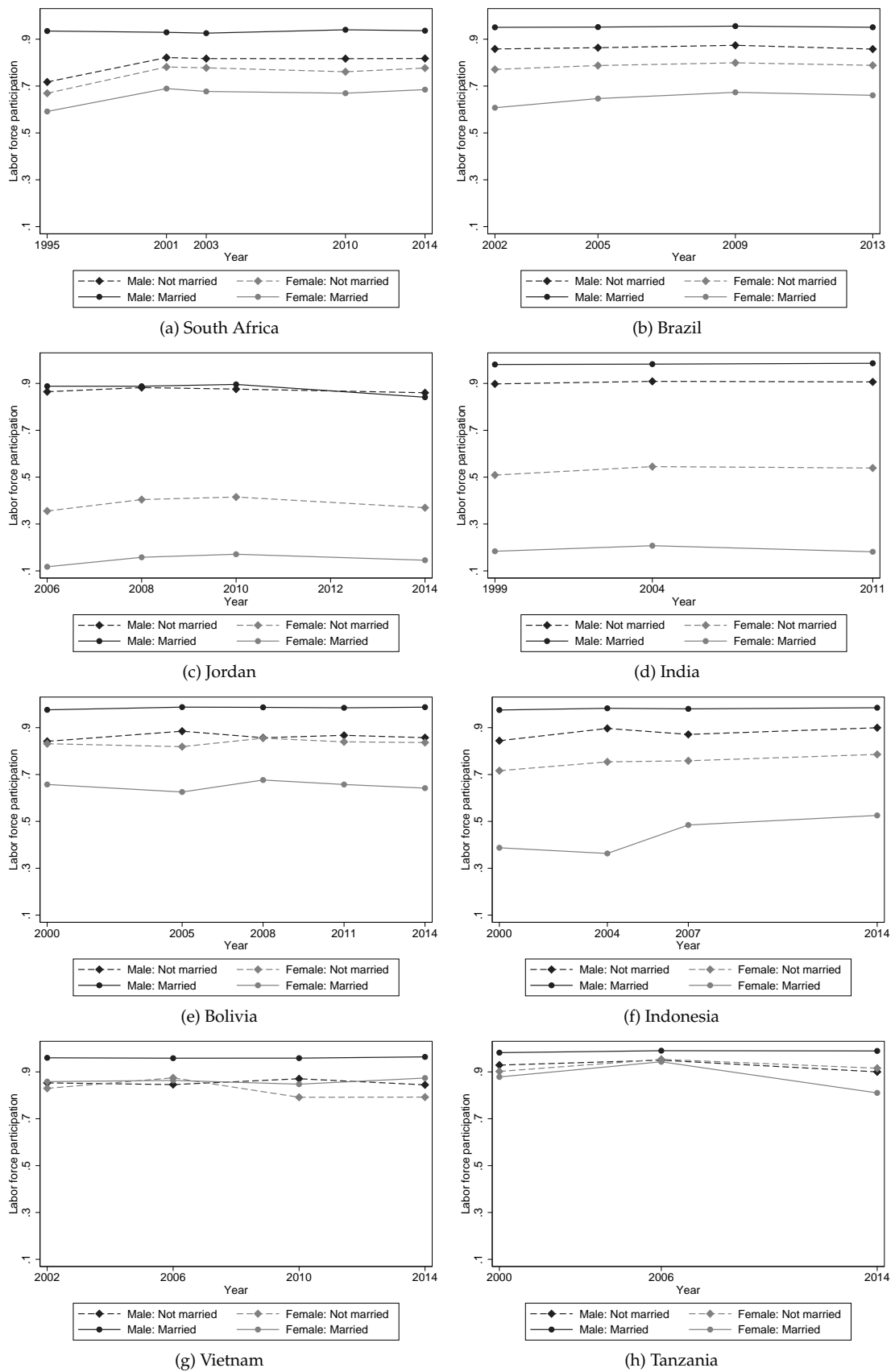
Labor force participation rates, by gender and marital status, have evolved differently across countries (Figure 2.3). Participation rates of urban married women are very high, above 80 percent, in Tanzania and Vietnam. They have fallen over time in Tanzania, remaining stable in Vietnam. Brazil, Bolivia, and South Africa follow with participation rates of 65–70 percent in 2013–14. Bolivia experienced minor fluctuations since 2000, while participation rates have increased over time in Brazil and South Africa. In the latter country, many married women entered the labor market immediately after the end of *apartheid*, between 1995 and 2001. Indonesia had the largest increase in the participation rate of married women: from 39 percent in 2000 to 53 percent in 2014. In Jordan and India, in contrast, less than 20 percent of married women participated in the labor force in 2014. In addition, trends have been disappointing: sluggish gains in Jordan—from 12 to 15 percent between 2006 and 2014<sup>13</sup>—and complete stagnation in India—18 percent in both 1999 and 2011.

Over time, the trend in participation rates is similar for currently married and not currently married women, but the former have a lower level of participation (except in Vietnam). Married men, in contrast, have extremely high participation rates in all countries and years, exceeding 90 percent in most cases.

Women's education levels have been rising in all countries (Figure 2.4). The share of married women with completed secondary schooling or some tertiary education grew, while the share of women with less than primary schooling decreased substantially. Progress was strongest in Brazil and South Africa, and weakest in Jordan.

The relationship between education and labor force participation differs across countries (Figure 2.5). In Brazil and South Africa, more educated women have higher participation rates; this relationship is strong, close to linear, and stable over time. In Bolivia, Vietnam, and Tanzania, the positive association between education and participation is much flatter and less stable over time. In Jordan, India, and Indonesia, the relationship between the two variables is U-shaped, as was reported for India by Klasen and Pieters (2015): relative to women with the lowest education level, average participation rates are lower for women with intermediate education, increasing somewhat for secondary school graduates and substantially for women with tertiary education. We return to the education-FLFP relationship in the discussion of our estimation results, which confirm

<sup>13</sup> Assaad et al. (2014) argue that even this 3 percentage point gain is illusory, resulting instead from the change in sampling frame and stratification of the Employment and Unemployment Survey in 2007. They show that FLFP rates (all women, ages 15+) were stagnant around 12 percent in 2000–2006, jumped to 15 in the first quarter of 2007, and remained stagnant thereafter (Assaad et al., 2014, Figure 1).

**FIG. 2.3:** Labor force participation by gender and marital status

Notes: See Table 2.A.1 for sources. Urban only, age 25–54; except urban *and* rural in Jordan. Common Y-axis for all subfigures.

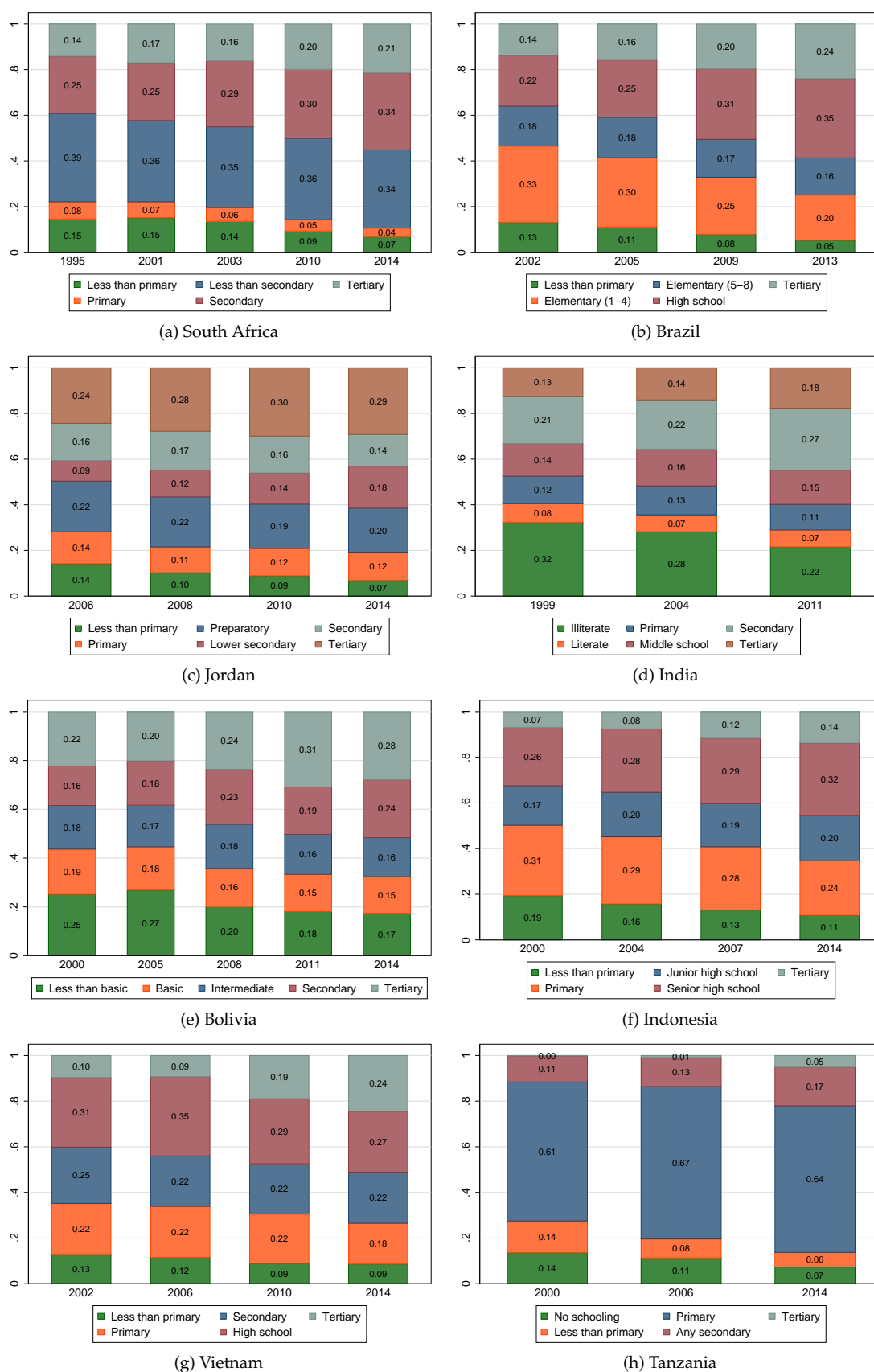


FIG. 2.4: Distribution of educational attainment over time

Notes: See Table 2.A.1 for sources. Urban married women, age 25–54; except urban *and* rural in Jordan. Common Y-axis for all subfigures.

striking differences across countries.

The average number of children in a married woman's household reflects distinct fertility and co-residence patterns across countries (Figure 2.6). Jordan and Tanzania show the highest number of children, both ages 0–4 and 5–14, per household; Brazil and Vietnam have the lowest.<sup>14</sup> Overall, most countries experienced a decline in the number of children per household over time.

In all countries, working married women are concentrated in a few industries. Most highly educated women work in white-collar services, in particular, public administration, education, and health; the majority of less educated women work in other services, in particular, wholesale and retail trade (Figure 2.A.1). In urban Tanzania, agriculture remains the most important activity for less educated women. Construction and mining employ very few married women in all countries.

Based on these descriptive statistics, we can draw several hypotheses. The different patterns we observe in terms of the education-participation relationship imply that rising education levels will have very different impacts on women's participation rates across countries. In some countries, particularly those with a strong U-shaped relationship, the impact may be limited or even negative. On the other hand, declining fertility is likely to contribute to higher participation rates everywhere, though this depends on the extent to which the presence of children is a barrier to women's participation in the different countries and how this changed over time. The distribution of female workers across industries suggests that changes in the sectoral structure of employment could have important bearings on women's likelihood of entering the labor force. While the descriptive patterns are quite similar across countries, the structure of growth may differ and could potentially explain differences in trends in participation rates. Finally, aggregate income growth has two potentially counteracting impacts: rising unearned income and rising earnings. As discussed below, we do not analyze the effect of women's own expected earnings, which will to some extent be captured by the effects of education. Increases in unearned income are likely to have a negative impact on participation rates in all countries, and here our interest mainly lies in the extent of this force.

---

<sup>14</sup>The figure for Jordan is inflated by including rural areas.



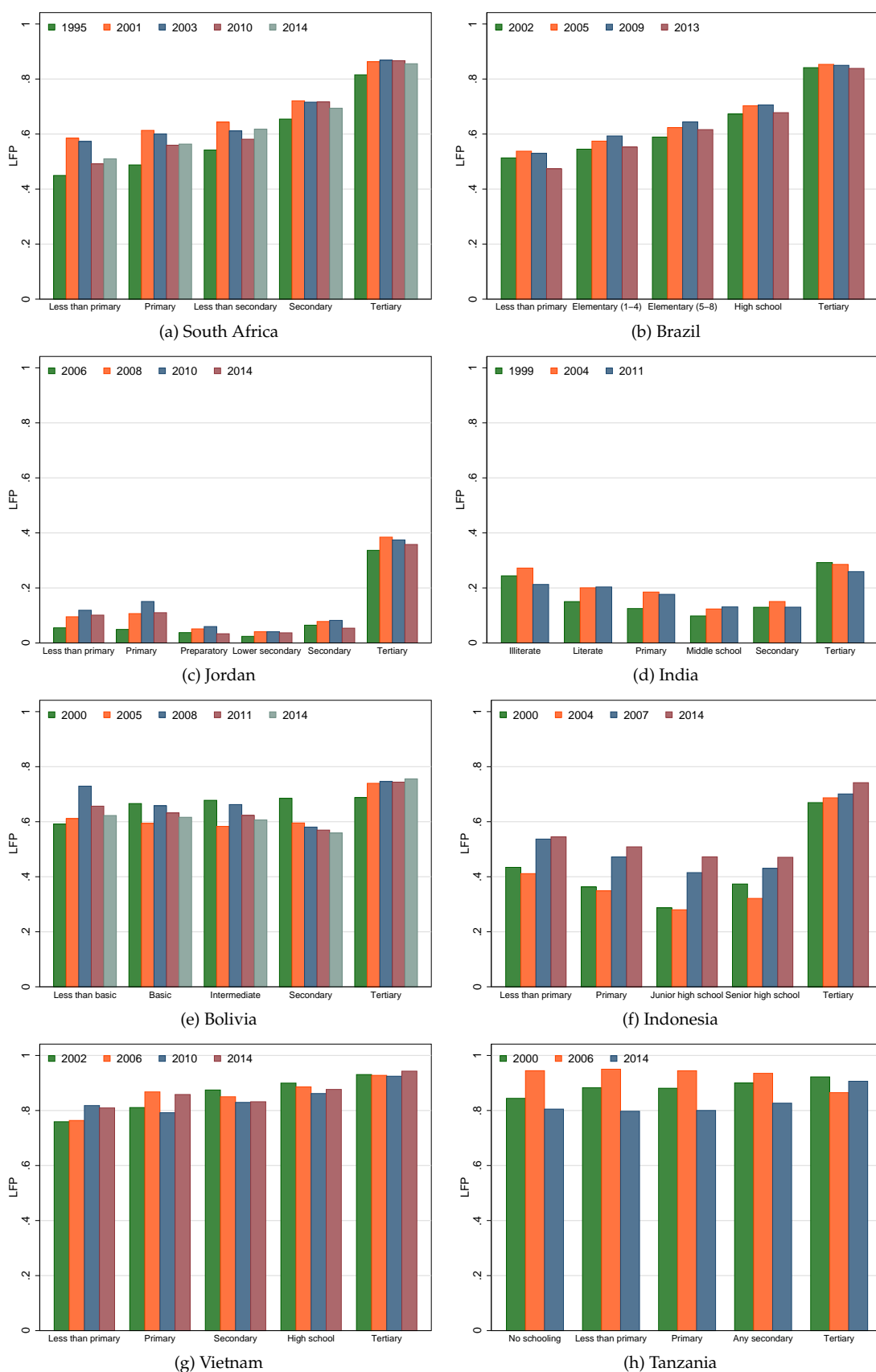
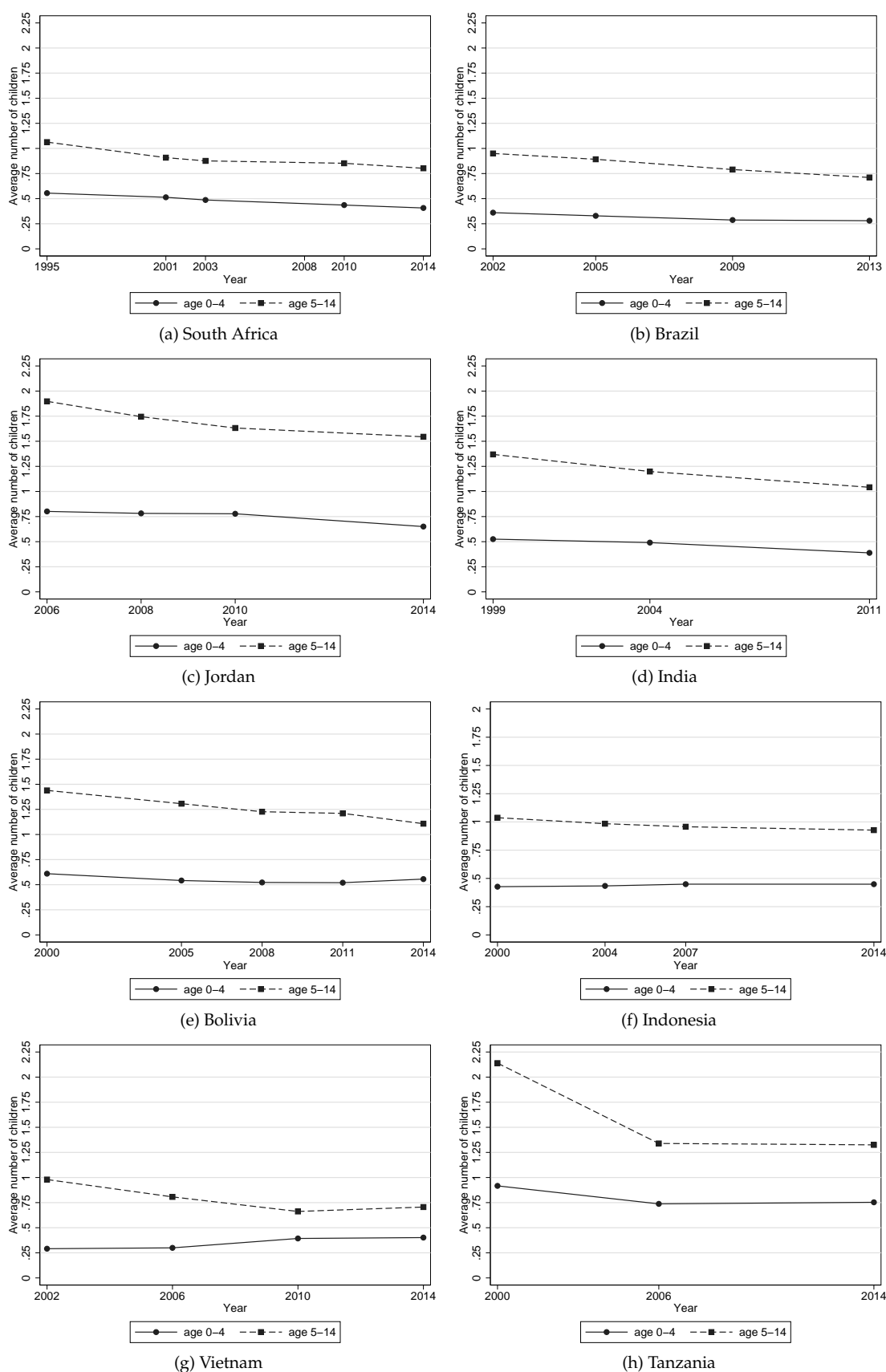


FIG. 2.5: FLFP by education level

Notes: See Table 2.A.1 for sources. Urban married women, age 25–54; except urban *and* rural in Jordan. Common Y-axis for all subfigures.



**FIG. 2.6:** Average number of children in a urban married woman's household

*Notes:* See Table 2.A.1 for sources. Urban married women, age 25–54; except urban *and* rural in Jordan. Common Y-axis for all subfigures.

### 2.2.3 Modeling female labor force participation

The empirical analysis follows the probit model of Klasen and Pieters (2015) for married women, ages 25–54, living in urban areas:

$$P(LFP_{ict} = 1) = \Phi \left( \alpha_{ct} + \sum_E \beta_{ct}^E D_{ict}^E + \mathbf{X}_{ict} \mathbf{f}_{ct} + \delta_{rct} \right), \quad (2.1)$$

where  $LFP$  is the labor force participation status of woman  $i$  in country  $c$  and year  $t$ , and  $\Phi(\cdot)$  is the standard normal CDF.<sup>15</sup> As indicated by the coefficients' subscripts, we estimate a separate probit model for each country-year pair.  $D_{ict}^E$  is a set of dummies for the woman's education attainment, whose exact definition varies across countries, but mostly captures attainment at the primary, secondary, or tertiary level.  $\mathbf{X}_{it}$  is a vector of individual and household variables. At the individual level, the vector contains the woman's age, its square, and her membership to ethnic or religious groups.<sup>16</sup> At the household level, we first capture, as two separate variables, the number of children of ages 0–4 and 5–14. Second, we add a set of education attainment dummies for the household head except whenever woman  $i$  is the head; we code those cases with a separate dummy variable.<sup>17</sup> Further, vector  $\mathbf{X}$  includes the natural log of per capita monthly household income, defined as the sum of earnings from each individual's main occupation, *excluding* woman  $i$ 's earnings. To proxy for a stable income source, we measure whether at least one adult male in the household is currently engaged in wage employment.  $\delta_{rct}$  is a set of regional fixed effects that capture demand and supply conditions at the local labor market level for each country-year. Regions, indexed by  $r$ , vary in number and dimension by country, but we always use the highest subnational level of aggregation available in each survey.<sup>18</sup> Finally,  $\alpha_{ct}$  is an intercept.<sup>19</sup> We cluster

<sup>15</sup>We obtain similar results with a logit model.

<sup>16</sup>For Indonesia and Tanzania, it was not possible to derive meaningful proxies for ethnicity or religion that were also comparable over time.

<sup>17</sup>For South Africa, however, we use an alternative definition of household head education, since the head is not identifiable from the data. As a best approximation, we use the maximum educational attainment of any adult married man of age 18+, with an additional dummy whenever no such household member exists.

<sup>18</sup>These are: provinces in South Africa, Indonesia, and Vietnam, states in Brazil and India, governorates in Jordan, departments in Bolivia, and regions in Tanzania. As a robustness check, we remove as much spatial heterogeneity as possible by adding primary sampling unit (PSU) fixed effects to the model. PSU information is not available for all surveys. For Brazil, Bolivia, South Africa and Tanzania, we find similar results with either PSU or regional fixed effects. For India and Indonesia, adding fixed effects at the second highest subnational level—districts in India, regencies (*Kabupaten*) and cities (*Kota*) in Indonesia—also produces similar results.

<sup>19</sup>We also include survey wave dummies whenever there are several survey waves per year (as in South

standard errors at the regional level.<sup>20</sup>

In an alternative specification, we analyze whether FLFP is associated with the sector in which jobs are available locally, as do Klasen and Pieters (2015) for urban India. As a result of norms about the types of work appropriate for women, discriminatory practices, and the extent to which hours and location of work are flexible within a particular occupation, employment opportunities for women may depend especially on employment growth in particular sectors. To capture the structure of local labor demand, we replace the regional fixed effects with the sectoral composition of male employment at the regional level (construction, agriculture, mining, manufacturing, white-collar services, and other services).<sup>21</sup> However, we find no clear relationship between these sectoral variables and FLFP. For this reason, we only present results for the specification with regional fixed effects.

Our estimates are best interpreted as reduced-form correlations. In this setting, endogeneity mainly stems from omitted variable bias, due to the individual or household unobservables jointly determining labor force participation, education, fertility, marital matching, and location (urban-rural). We explicitly address some of these concerns in sections 2.3.1 and 2.3.2, where we assess the importance of selection bias related to marriage, settlement in urban areas, and educational attainment. Reverse causality, on the other hand, is less of a concern. We assume that prime-age women completed their education and marriage market histories. Moreover, we assume that each woman takes the labor market status of her spouse as exogenous, since in all countries and years of our sample prime-age married men have nearly universal labor force participation rates.

We do not attempt to causally identify structural parameters for two reasons. First, there is no quasi-experimental strategy (such as an instrumental variables approach) applicable to all countries and years similarly.<sup>22</sup> Second, the prevailing methods for estimating own-wage effects are notoriously challenging and known to produce unstable results.<sup>23</sup> In addition, the quality of existing wage data varies substantially across

---

Africa after 1995, Jordan, and Tanzania).

<sup>20</sup>For more details on the construction of these variables across countries see Appendix 2.A; for sample means of the variables by country and year see Tables 2.A.2-2.A.9.

<sup>21</sup>The Indian and Indonesian surveys are representative at the second highest subnational level; this is the level of aggregation used for the regional employment share variables. For the remaining countries, we use the highest subnational administrative level to aggregate the employment shares.

<sup>22</sup>In principle, one could pursue a country and year-specific IV approach, but the resulting local average treatment effects would be hard to interpret in a unified comparative framework, as the population of compliers would vary across settings and IVs.

<sup>23</sup>See Klasen and Pieters (2015, pp. 460–461) for a discussion of the lack of robustness in estimates of own-wage effects in India, as well as a more detailed discussion of the challenges involved in such estimations.

surveys.

## 2.3 Results

We first summarize the estimation results for each country (in increasing order of GDP per capita), and then turn to a discussion of the main trends and patterns. Table 2.1 provides an overview of the relationship between key variables (or variable groups) and women's labor force participation in each country, and their changes over time.<sup>24</sup>

In Tanzania, FLFP increases linearly with education attainment in 2000 and 2014. The effect of household income is negative but small, and declines in absolute magnitude over time. The number of children aged 0–4 only has a significant (and small) negative effect in 2014. Otherwise, the number of children in the household does not correlate with FLFP. Besides a tiny negative effect of household income, none of the explanatory variables is statistically significant in 2006, which likely reflects the lack of variation in the dependent variable: the participation rate in the estimation sample is 92 percent.

India shows a clear U-shaped relationship between own education and FLFP. Relative to the reference group of illiterate women, the average marginal effects are negative and larger in magnitude with each additional level of educational attainment up to completed middle schooling—which is the level associated with the lowest participation rates in all years. The average marginal effect is still negative for complete secondary schooling, but closer to zero. For women with any tertiary education, the positive marginal effect is very large and significant, although declining over time: from 21 percentage points in 1999 to 14 percentage points in 2011. Household head education, household income, and male salaried employment (to proxy security of income) correlate negatively with participation—although the latter effect is no longer significant in 2011. The presence of young children is correlated with lower participation, and this negative effect is becoming stronger over time. For older children, the average marginal effect is actually positive after 1999, but always small. Finally, caste and religion are important correlates of FLFP as well, with lower caste and Hindu women being more active in the labor market than upper caste and Muslim women. The effect of caste is weakening over time; the effect of religion is strengthening.

In Vietnam, the relationship between education attainment and FLFP is positive and linear in 2002, but only the effect of tertiary education remains over time. The small negative income effect in 2002 becomes insignificant in the later years. The number

---

<sup>24</sup>The average marginal effect estimates for the probit models are reported in Tables 2.A.10-2.A.17.

of young children is negatively associated with FLFP after 2002; the effect is large (in absolute terms) and increases over time: in 2014, one additional young child is associated with a 6 percentage points reduction in women's participation probability. We do not find clear associations between FLFP and older children, male salaried employment, household head education, or ethnicity.

In Bolivia, education is not significantly correlated with FLFP, except for tertiary schooling, which affects participation positively. The estimate fluctuates a bit between 2000 and 2008, after which it increases until 2014. Household income and salaried employment of a male household member reduce FLFP. The effects are substantial, when compared to estimates from other countries. The presence of at least one male salaried employee in the household correlates with a 4 to 10 percentage point decline in the woman's participation likelihood, depending on the year. Young children have a sizable negative effect. Household head education was negatively related to FLFP in 2008, 2011, and 2014, with the effect getting weaker over time. Native speakers of indigenous languages are more likely to participate in the labor market.

In Indonesia, the relationship between own education and FLFP in the first year (2000) resembles the U-shape found for India, with negative effects of primary and junior high school completion (relative to the reference group of women who did not complete primary school), and positive effects of completed secondary schooling and especially tertiary education. Yet, the pattern changes: in 2014, only the positive tertiary education effect remains, and it is somewhat smaller than in 2000. Household income has a sizable negative effect on participation, and this becomes stronger over time. The estimates of male salaried employment are, likewise, negative and increasing (in absolute terms), while the negative effect of household head education decreases. There is a large negative effect of young children and a smaller negative effect of older children. Both are increasing over the years, in absolute terms.

In Jordan, tertiary education has a strong positive relationship with FLFP, and the effect is very stable over time. Across lower education levels the relationship with FLFP is flat, except for a small negative effect of lower secondary education, resulting in a J-shaped education-participation relationship. Income has a small but significant negative effect in every year.<sup>25</sup> Male salaried employment increases FLFP in the most recent years (2010 and 2014), while the positive effect of a tertiary educated household head disappears after 2010, both suggesting that income security is less relevant. We

<sup>25</sup>The small size of the income effect should be interpreted with caution. The earnings variable available from the Jordanian surveys is very roughly measured: it is the mid-point of five earning brackets. We thus suspect the average marginal effects of household income to suffer from attenuation bias.

TABLE 2.1: Estimation results: overview

Country	Own education	Household head education <sup>†</sup>	Household income	Male salaried employment	Children 0–4	Children 5–14	Population group
<b>Tanzania</b>							
Effect	weak linear increase	(0)	small (–) ↗	(0)	small (–) in 2014	(0)	not included
Time trend	→						
<b>India</b>							
Effect	strong U-shape	sizable (–) ↗	sizable (–) →	small (–) gone by 2011	small (–) ↗	tiny (+) after 1999, →	sizable: caste (religion) ↘ (↗)
Time trend	more shallow						
<b>Vietnam</b>							
Effect	weak linear increase	(0)	small (–) gone after 2002	(0)	sizable (–) after 2002, ↗	(0)	(0) ethnicity
Time trend	↗						
<b>Bolivia</b>							
Effect	weak linear increase	small (–) for tertiary educ after 2005, ↗	sizable (–) ↗	sizable (–) ↗	sizable (–) ↗	(0)	sizable: language ↗
Time trend	↗						
<b>Indonesia</b>							
Effect	U-shape	(–) ↗	sizable (–) ↗	small (–) ↗	sizable (–) ↗	small (–) ↗	not included
Time trend	more shallow						
<b>Jordan*</b>							
Effect	imprecise U-shape	(+) for tertiary educ gone by 2014	small (–) →	small (+) after 2008	small (–) →	small (–) →	sizable: nationality ↗
Time trend	slightly deeper						
<b>South Africa</b>							
Effect	strong linear increase	(0) [strong (+) for female heads] [gone after 2003]	small (–) →, but gone in 2014	(0)	sizable (–) slightly ↗	small (–) slightly ↗	sizable: race →
Time trend	↗						
<b>Brazil</b>							
Effect	strong linear increase	(–) [strong (+) for female heads] ↗ and gone after 2009 [↗]	small (–) turning (+) by 2013	small (–) gone by 2009	sizable (–) ↗	small (–) ↗	sizable: race ↗
Time trend	→						

Sources: Authors' own calculations, see Table 2.A.1. Average marginal effects by country and year are reported in Tables 2.A.10–2.A.17. Notes: (0) denotes statistical insignificance or estimates very close to zero; (+) [–] [–] denotes positive [negative] and significant estimates. ↗ (↘) denotes increasing (decreasing) coefficient estimates over time in *absolute terms*. Countries are sorted by per capita GDP. † For South Africa, household head education is proxied by the highest education level of any adult married male in the household. \* For Jordan, both urban and rural areas are included.

may rather be picking up effects of assortative matching. Nationality is a significant factor. Women from other Arab countries are significantly less likely to participate than Jordanian women (and the effects become stronger over time), while those from non-Arab countries are much more likely to be active. Finally, both young and older children reduce FLFP, and these effects are stable over time.

In South Africa, we see strong positive participation-returns to education, with the marginal effect of education increasing at each level. The effects declined between 1995 and 2001 but then increased again until 2014. Household income has a slight negative effect that is no longer significant in 2014, and we find no clear association between household head education or male salaried employment and FLFP. Differences by skin color or ethnic group are large and significant: white women and especially Indian/Asian women are less likely to participate than black and colored women. As in Jordan, both young and older children reduce FLFP. The estimates are larger in South Africa than in Jordan, especially for young children, and they become slightly stronger over time.

Finally, the results for Brazil also show a strong positive education-participation relationship, with increasing marginal returns. The returns to elementary school increase over time, whereas returns to higher education levels are stable across periods. We find a slight negative income effect from 2002 to 2009, which turns significantly positive in 2013. Male salaried employment has a negative effect in the early 2000s but no longer in 2009 and 2013. Household head education also has a negative effect in 2002 and 2005; afterwards, only tertiary educated household heads have a negative effect whose magnitude declines substantially. Female household heads, on the other hand, were 11 percentage points more likely to be active in 2002. The effect is shrinking fast over time, and is no longer statistically significant in 2013. Differences between ethnic groups declined over time, but remain noteworthy. The negative effects of children, which are larger for ages 0–4, became stronger over time.

Our results reveal two types of patterns between women's own educational attainment and their labor force participation: (i) a strong positive relationship with linearly increasing marginal participation-returns to education in Brazil and South Africa, and (ii) a U- or J-shaped relationship in India, Indonesia, and Jordan (Figure 2.7). In Bolivia, Vietnam, and Tanzania, the two patterns mix. Initially, there is a linear positive relationship that is much flatter and imprecise than in Brazil and South Africa. But, over time, the relationship turns into a J-shape in Bolivia and Vietnam, as the returns to low and intermediate education fall to zero. In South Africa, India, and Indonesia, the positive effect of secondary and tertiary education declined over time. For India, Klasen and



Pieters (2015) relate this decline to changes in the selectivity of higher education, an issue we address in section 2.3.2.

The patterns suggest that the education-participation relationship moves from weak linear in low-income countries to a U- or J-shape in middle-income countries, before becoming strongly positive in upper-middle income countries. To some extent, this is also the pattern we observe over time within Bolivia and Vietnam. Our results thus illustrate that countries growing from low-income to lower-middle-income status will not necessarily experience an increase in the participation returns to education, and therefore increases in educational attainment levels may have ambiguous effects on FLFP rates.

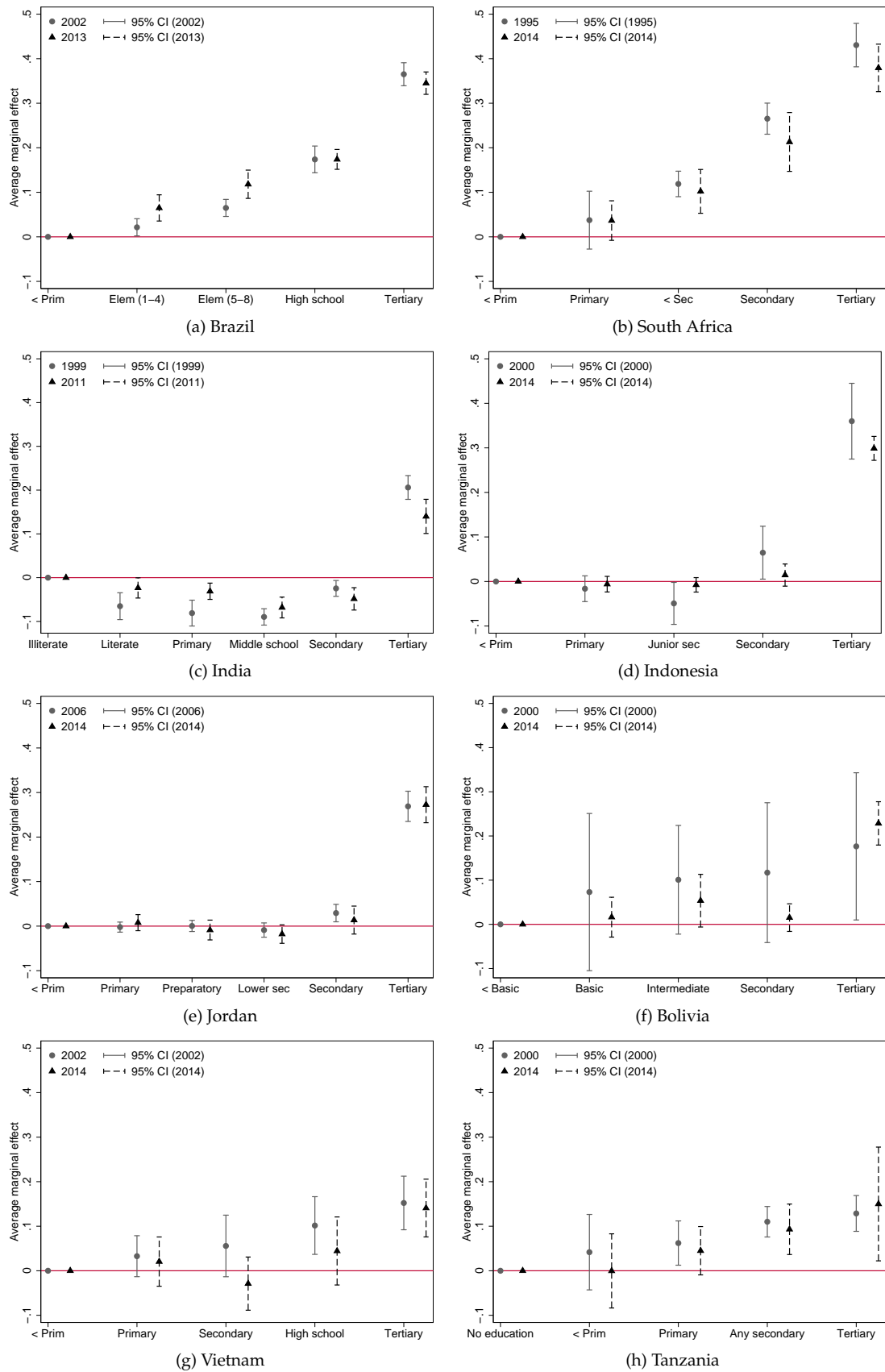
Furthermore, India, Indonesia, and Jordan are not only at the middle of the GDP per capita distribution in this sample of countries, but also form a more or less distinct group in terms of social and religious norms around women's participation in market activities. It is likely that the U-shape or J-shape at least partly reflects such norms, by which employment outside the home is not deemed appropriate for women at intermediate levels of education. In India and Indonesia, this is further corroborated by a negative relationship between household head education and FLFP, indicating that when the household's socio-economic status improves, women withdraw from the labor force.

Household income is negatively related to women's participation everywhere, but interestingly the negative effects disappeared in South Africa and Brazil by 2013–14 (Figure 2.8). In these two countries, male salaried employment and household head education have no clear relationship with FLFP either. Hence, in the richest two countries in our sample, income and income uncertainty seem to play no role in 'pushing' women to participate in the labor force, whereas their own education is a major factor.<sup>26</sup> Thus, women's own characteristics matter the most for labor force participation; household conditions, except the number of young children, have become irrelevant.

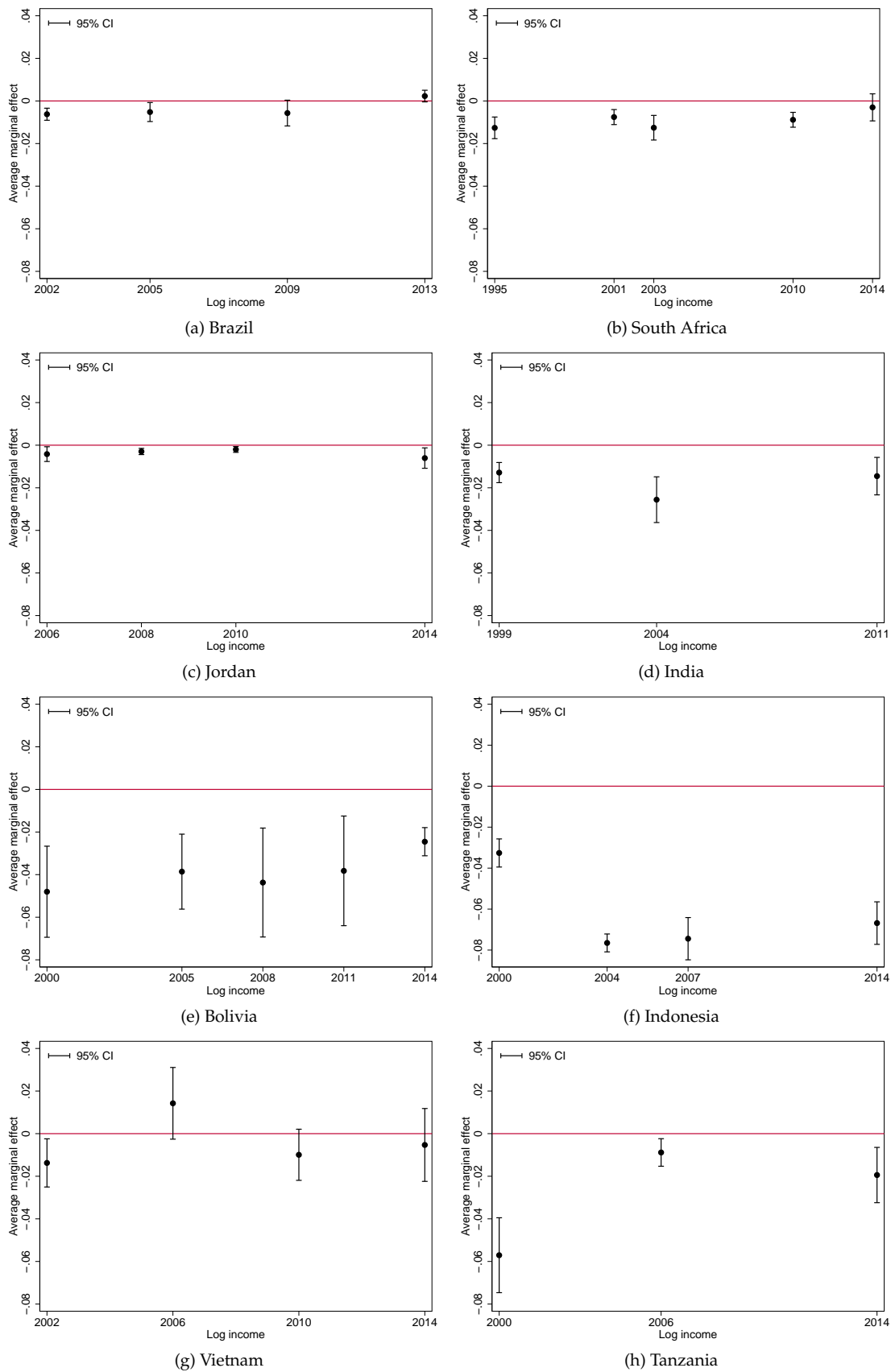
The role of children is also noteworthy (Figure 2.9). While women in households with young children are less likely to participate in the labor force in all countries and time periods, older children reduce FLFP only in the relatively high-income countries. In poorer countries we find no evidence for such a relationship, which may reflect income constraints, whereby mothers cannot afford to stay out of the labor force for long in poorer settings.<sup>27</sup>

<sup>26</sup>This finding resembles patterns that are taking place in OECD countries in the past decades. In the US, for example, Blau and Kahn (2007) and Heim (2007) show that income elasticities of married women labor supply have plummeted since the 1980s.

<sup>27</sup>See Priebe (2010) for causal evidence of this mechanism in Indonesia.

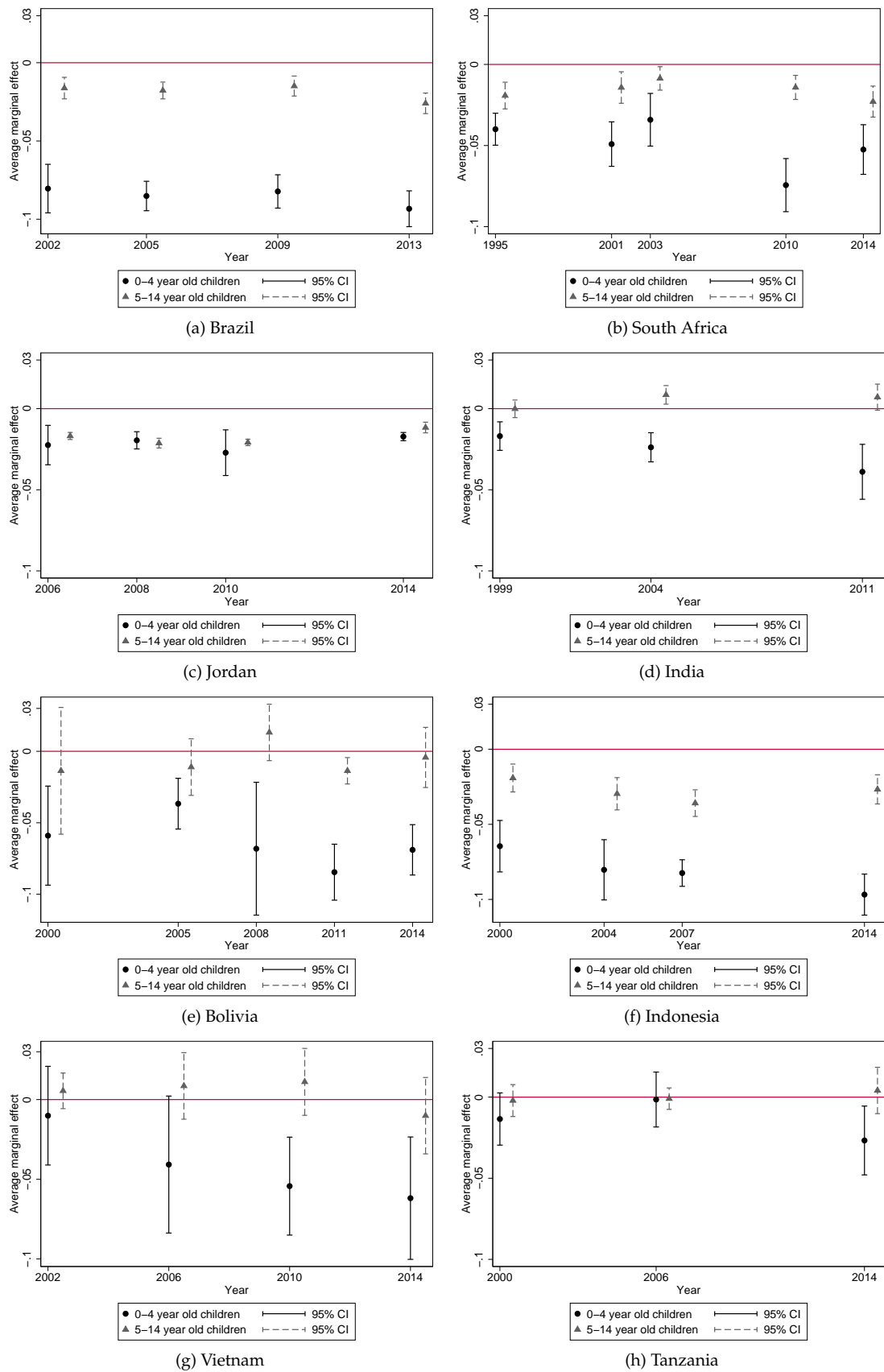
**FIG. 2.7:** Average marginal effects of the woman's own education

Notes: Common Y-axis for all subfigures. Average marginal effects of the full probit model are reported, for each country and year, in Tables 2.A.10-2.A.17.



**FIG. 2.8:** Average marginal effects of log household per capita earnings (excluding woman's own earnings)

*Notes:* Common Y-axis for all subfigures. Average marginal effects of the full probit model are reported, for each country and year, in Tables 2.A.10-2.A.17.



**FIG. 2.9:** Average marginal effects of the number of children

*Notes:* Common Y-axis for all subfigures. Average marginal effects of the full probit model are reported, for each country and year, in Tables 2.A.10-2.A.17.

In sum, the correlates of FLFP differ across (groups of) countries. In the remainder of the chapter, we investigate whether these differences can explain trends in labor force participation and gaps across countries. But first we assess whether the estimates are robust to trends in marriage and urbanization rates.

### 2.3.1 Selection into marriage and urban areas

Given our exclusive focus on urban married women, it is important to take into account trends in the incidence of marriage and urban residence. Otherwise, unobservable factors shaping selection into marriage (or urban areas) and selection into the labor force could affect how coefficient estimates evolve over time. In our data, marriage rates declined in South Africa, Brazil, and Bolivia, increased in Jordan, Indonesia, and Vietnam, and remained constant in India and Tanzania. Urbanization rates were rising in Bolivia, India, Indonesia, Tanzania, South Africa, and Vietnam, while no changes occurred in Brazil.<sup>28</sup>

We follow Blau and Kahn (2007), who control for falling marriage rates in modeling female labor supply in the US. Their procedure consists of (1) predicting an individual's marriage probability, and (2) excluding from the estimation sample the married women with lowest predicted probabilities (i.e., the least "marriage-prone" among the married) such that the resulting "adjusted"-marriage rate is equal across survey years.<sup>29</sup> We estimate marriage probabilities for each country-year from a probit model with the covariates age, age squared, ethnicity/religion, education attainment, regional dummies, and (whenever relevant) survey wave dummies. We use the same approach to control for selection into urban areas. The probit model predicting urban residence includes the same covariates as the marriage model plus the number of children in the household of ages 0–4 and 5–14. We then re-estimate our labor force participation model on the "adjusted"-samples.

Trends in marriage incidence or urbanization do not affect our findings. The results from the "adjusted"-marriage and "adjusted"-urban samples are qualitatively similar to the baseline estimates.<sup>30</sup>

Whenever possible, we directly control for rural-urban migration. For Tanzania, we

<sup>28</sup>The magnitude of these changes varies across countries. See Table 2.A.18 for the relevant descriptive statistics. Recall that there is no urban/rural information for Jordan.

<sup>29</sup>For example, the incidence of marriage among prime-age urban women in South Africa fell from 0.66 in 1995 to 0.51 in 2014. Thus, from the 1995 sample, we eliminate the 22.7 percent  $[(0.66 - 0.51)/0.66]$  least marriage-prone individuals based on their predicted marriage probabilities. In practice, there are tiny differences in the resulting "adjusted"-marriage rates due to the use of sampling weights.

<sup>30</sup>We do not report them here; available upon request.

find that women migrating more than five years before the survey are 5 percentage points more likely to be in the labor force in 2014; the effect being insignificant in the first two years. For migrants arriving less than five years before the survey, the effects are never significant. In Brazil, migration status (captured by individuals' place of birth being in a different state or different municipality than their current residence) has no significant effects. In Bolivia, a woman's migration status (a dummy variable for whether, five years before the survey, she lived outside the municipality of current residence), was associated with lower labor force participation only in the last two survey years (the average marginal effect is around minus 5 percentage points in both years).<sup>31</sup> For all three countries, adding the migration controls does not affect the average marginal effects of the remaining explanatory variables.<sup>32</sup>

In sum, trends in the rates of marriage and urban residence among prime-age women do not influence the determinants of labor force participation.

### 2.3.2 Selection into education

We now consider selection into education, not because of concerns about the robustness of our estimates, but rather out of interest in the forces driving changes in returns to education. Education levels have increased over time in all eight countries. Since, in our sample of prime-age married women, education histories are largely complete, average educational attainment increases because younger, more educated cohorts progressively replace older, less educated ones. This process raises the question of whether the selection of women into education levels varies across cohorts. If it does, trends in the estimated average marginal effects of educational attainment could be driven by changes in the sample's cohort composition, rather than by changes in the marginal effects of education.

We explore this possibility in more detail for India, Indonesia, and South Africa. The three countries experienced rising shares of highly educated women (tertiary level, see Figure 2.10) and, simultaneously, a sizable *decrease* in the (positive) average marginal effect of being highly educated (see Figure 2.7).

We would like to know how much of the decline in the effect of tertiary education could be plausibly explained by decreasing selectivity of women in terms of labor force attachment at the top of the education distribution. Klasen and Pieters (2015) propose

<sup>31</sup>Full results available upon request.

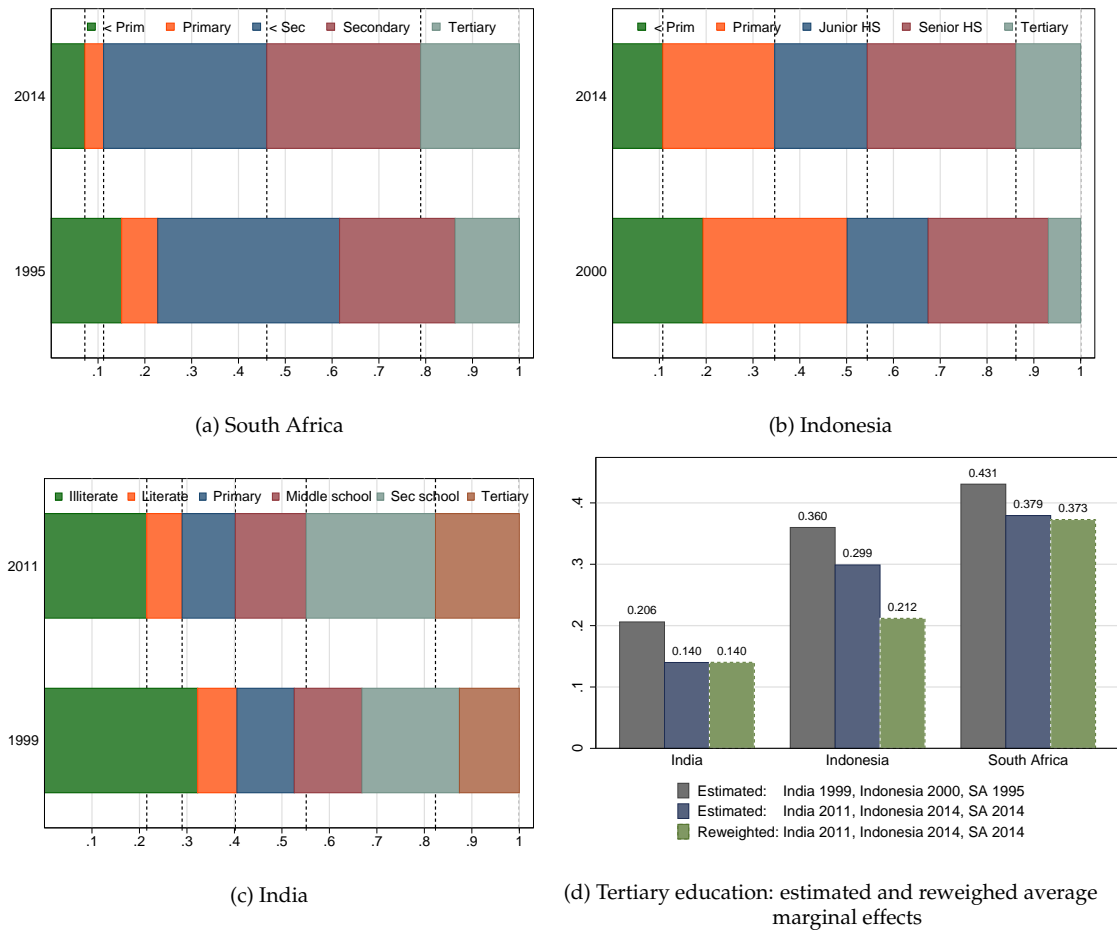
<sup>32</sup>Klasen and Pieters (2015) show that, for India, the 1999 results are robust to adding migration variables (both the woman's and her spouse's), which are themselves insignificant. There is no migration data available for 2004 and 2011.

a thought experiment to estimate an upper bound on the size of the selection effect. Imagine that the initial distribution of women's educational attainment is a one-to-one match to the distribution of unobserved labor force attachment. If there are  $K$  educational levels, there are also  $K$  attachment levels; the women achieving the highest level of education being also the ones with the highest level of labor force attachment. As a result, the average marginal effect of education on labor force participation is positively biased. Now, consider a completely supply-driven expansion of education: the government produces and offers cost-free slots of tertiary education. The new slots are filled by women below that educational level in decreasing order of labor force attachment. That is, less attached women are moving up the education ladder. As a consequence, average labor force attachment at the tertiary level is now lower than before, and the estimated effect of education on labor force participation falls.

Consider two extreme scenarios of the thought experiment. If all women have the same labor force attachment (or education and labor force attachment are completely unrelated), the education expansion would have *no* selection effect; over time, any changes in the education estimates result from changes in the effect of education itself. If, on the other hand, the education effect is *fully* driven by labor force attachment, then the post-expansion education estimates are a weighted sum of the pre-expansion estimates, where the weights are the changes in the attachment composition of each education level.

With the last scenario in mind, we can estimate an upper bound of the selection effect. Let us illustrate the procedure for South Africa. In 2014, the share of women with tertiary education was 0.21. Nearly two decades before, in 1995, that share was 0.14. Thus, in 1995, one third of the women in the top 21 percentiles of the education distribution had complete secondary schooling (see Figure 2.10). We can then estimate the average marginal effect of being in the top 21 percentiles of the education distribution in 1995 as two thirds the average marginal effect of tertiary education plus one third the average marginal effect of completed secondary schooling. If this reweighted 1995 estimate comes closer to the average marginal effect of tertiary education in 2014, then the effect of being in the 21 highest education percentiles (relative to the reference group with below primary schooling) did not change over time. What changed instead was the selectivity of women into educational attainment.

For India and South Africa, we find that the reweighted estimates closely reproduce the average marginal effects of the latest year. In theory, the selection effect is large enough to explain the declining effect of high education in the two countries (Figure 2.10). For Indonesia, the reweighted estimate is about 30 percent smaller than the average

**FIG. 2.10:** Education selectivity

*Notes:* Panels (a), (b), (c) show changes in the education distribution over time. Panel (d) shows the estimated average marginal effects from the probit models and from the reweighting procedure described in the text.

marginal effect, implying that the selection effect can account for a stronger decline in returns to higher education than actually observed. This suggests that the participation returns to tertiary education may have in fact increased in Indonesia between 2000 and 2014, even though the estimated effect on labor force participation declined. Yet, since our reweighted estimate reflects the upper bound, it is also possible that the returns to education did not change, or declined.



## 2.4 Decomposition analysis

In this section, we decompose differences in labor force participation rates using Fairlie's (2006) extension of the Blinder-Oaxaca decomposition for binary dependent variables. Consider two mutually exclusive groups of women,  $A$  and  $B$ . In our case,  $A$  and  $B$  will be either the first and last survey of a country—thereby decomposing changes in labor force participation rates over time—or two countries in a given year—thereby decomposing the gap in participation rates between two countries. Start by defining the overall mean LFP gap between group  $A$  and group  $B$  as:

$$\Delta_O \equiv \mathbb{E}[LFP_B | D_B = 1] - \mathbb{E}[LFP_A | D_A = 1],$$

with  $D_g$  being an indicator variable determining membership of group  $g$ , where  $g = A, B$ .<sup>33</sup> Then, decompose the gap between the usual covariate contribution,  $\Delta_X$ , and the unexplained (i.e., coefficients and unobservables) contribution,  $\Delta_U$ , by substituting in our probit model of LFP and rearranging:

$$\begin{aligned} \Delta_O &= (\mathbb{E}[\Phi(X\beta_A) | D_B = 1] - \mathbb{E}[\Phi(X\beta_A) | D_A = 1]) \\ &\quad + (\mathbb{E}[\Phi(X\beta_B) | D_B = 1] - \mathbb{E}[\Phi(X\beta_A) | D_B = 1]) \\ &= \Delta_X + \Delta_U, \end{aligned}$$

Replacing the expectations with their empirical counterparts gives:

$$\overline{LFP}_B - \overline{LFP}_A \approx \left[ \sum_{N_B} \frac{\Phi(X_B \hat{\beta}_A)}{N_B} - \sum_{N_A} \frac{\Phi(X_A \hat{\beta}_A)}{N_A} \right] + \left[ \sum_{N_B} \frac{\Phi(X_B \hat{\beta}_B)}{N_B} - \sum_{N_B} \frac{\Phi(X_B \hat{\beta}_A)}{N_B} \right],$$

with  $N_g$  being the size of group  $g$ .<sup>34</sup> Notice how the coefficients of group  $A$ ,  $\hat{\beta}_A$ , weigh the covariate contribution, and the covariate distribution of group  $B$ ,  $X_B$ , weighs the unexplained term. An equally valid decomposition, but leading to different results, is using  $\hat{\beta}_B$  to weigh the covariate contribution and  $X_A$  to weigh the unexplained term.

In sum, the choice of counterfactual matters. In the presence of general equilibrium effects, the parameter vector for the appropriate counterfactual might be neither  $\hat{\beta}_A$  nor  $\hat{\beta}_B$ . Accordingly, we always report results based on both counterfactuals, interpreting them as a reasonable interval containing the true effect.

<sup>33</sup>We loosely follow the notation of Fortin et al. (2011), who review decomposition methods relevant to labor economics.

<sup>34</sup>The expression holds as an exact equality for logit models that include an intercept, whereas it holds very closely for probit models (Fairlie, 2006).

In a nonlinear setting, the main challenge is decomposing the total covariate contribution,  $\Delta_X$ , into the individual contributions of each covariate because the contribution of each variable depends on the distributions of all other variables. Fairlie's (2006) method creates a series of counterfactuals by sequentially replacing the distribution of a variable with its counterpart in the comparison group, while holding constant the distribution of the other covariates.<sup>35</sup> The average difference between the observed values and each counterfactual gives the variable's contribution.

In practice, the sample sizes of groups *A* and *B* always differ in our setting. Fairlie (2006) suggests taking random subsamples of the largest group that fit the size of the smallest group. Next, predict LFP probabilities within each group, rank each individual in her group based on her predicted LFP, and match similarly ranked pairs across the two groups (i.e., the top ranked observation of group *A* with the top ranked observation of group *B*, and so on). The final step is then to average the result over the sample draws.<sup>36</sup>

### 2.4.1 Decomposing changes over time within countries

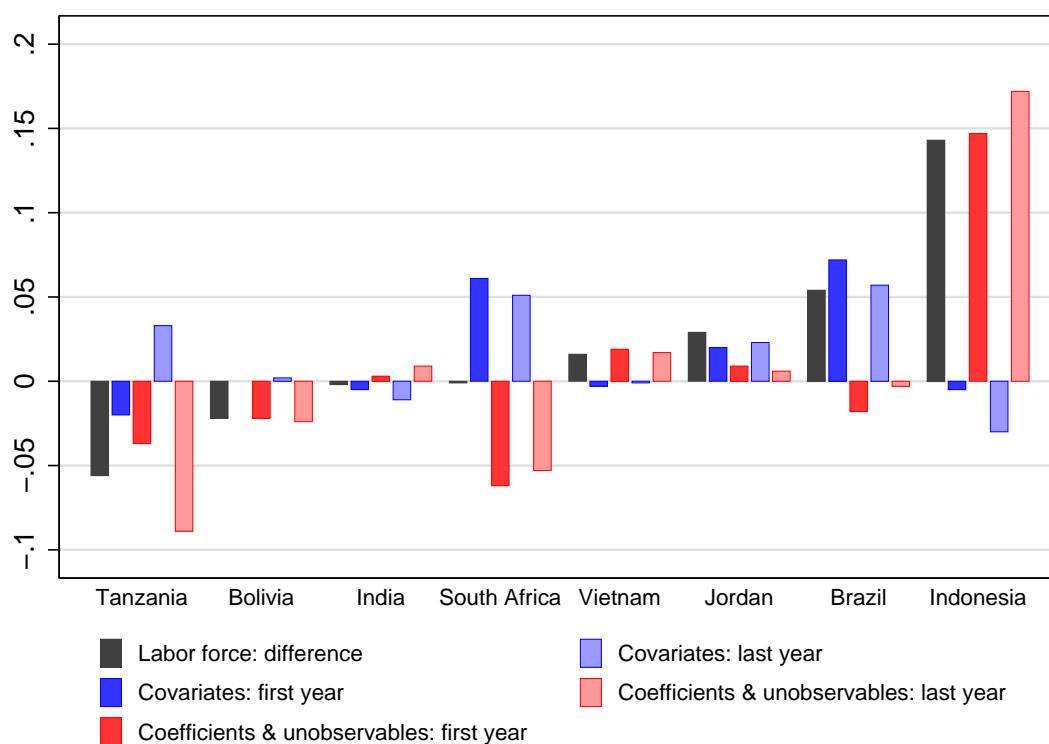
We start by decomposing the change in FLFP over time for each country. Figure 2.11 summarizes the results, showing countries in increasing order of the FLFP gap between the last and first years.<sup>37</sup> We show results for the two alternative counterfactuals: weighting the covariate contribution at first or last year coefficients. With the exception of Tanzania, the two counterfactuals produce consistent results: the contributions of covariates and of coefficients and unobservables (i.e., the unexplained term) have the same direction and order of magnitude.

The extent to which changes in covariates can account for changes in FLFP differs across countries. Covariates explain most of the change in India, Jordan, and Brazil. In India, FLFP declined slightly from 18.4 percent in 1999 to 18.2 percent in 2011. The change in covariates during that period would predict a reduction in FLFP of 0.5 to 1.1 percentage points. The negative effects of rising household head education and rising household income more than offset the positive effect of rising women's education and falling fertility. In Jordan, rising education and falling fertility drive most of the positive covariate effect, which accounts for more than two thirds of the small increase in FLFP

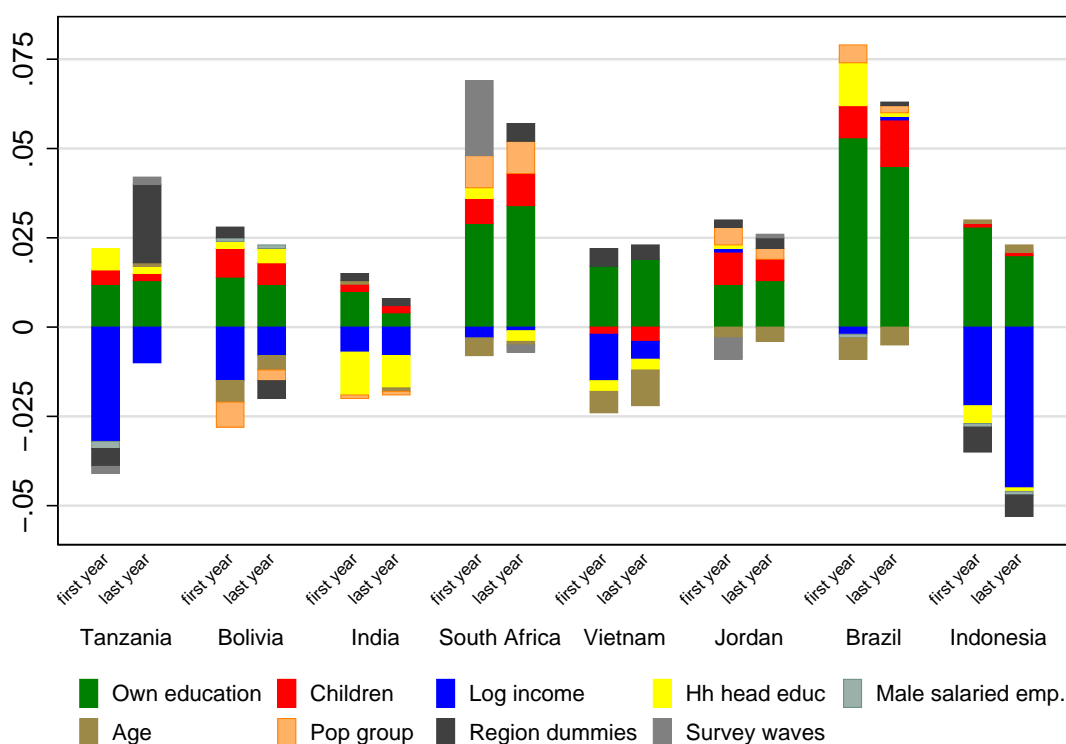
<sup>35</sup> See Fairlie (2006) for more details.

<sup>36</sup>For each decomposition, we draw 1000 random samples. In addition, at each sample draw, the ordering of the variables in the sequence of counterfactuals is randomly determined. This addresses the issue of path dependence: since individual contributions depend on the distributions of all other covariates, the ordering of the variables matters for the final result.

<sup>37</sup>For point estimates of the decompositions, see Tables 2.A.19-2.A.22.



(a) Total covariate contribution



(b) Contribution of variable groups

**FIG. 2.11:** Decompositions within countries over time

Notes: For point estimates of the decompositions, see Tables 2.A.19-2.A.22. *First year* and *last year* refer to the year of the coefficients used to compute the covariate contribution.

between 2006 and 2014. In Brazil, covariates come close to explaining the full FLFP increase between 2002 and 2013 (which was about 5 percentage points). Rising female education and, to a lesser extent, declining fertility are the main forces.

In Tanzania, Bolivia, Vietnam, and Indonesia, changes in coefficients and unobservables account for almost all of the change in participation rates.<sup>38</sup> The changes in covariates in each of these four countries contributed little or nothing to changes in FLFP rates, whether they declined (in Tanzania and Bolivia) or increased (in Vietnam and, especially, Indonesia). A positive contribution of increasing women's educational attainment (and declining fertility in Tanzania and Bolivia) was, in all cases, offset mainly by the negative contribution of rising household incomes.

Finally, in South Africa, where the participation rate was nearly constant between 2001 and 2014, the positive covariate contribution is offset by a negative contribution of coefficients and unobservables. Similar to Brazil, the covariate contribution is large and is accounted for by rising female education levels and reduced numbers of children.<sup>39</sup>

Summarizing the main findings, rising educational attainment contributed to higher FLFP in all countries, but most strongly in Brazil and South Africa, reflecting the strong participation-returns to education in these countries. In the other countries, the contribution was more limited, but still positive, despite the U- or J-shaped relationship between education and participation in Jordan, India, and Indonesia. This reflects educational attainment increasing predominantly at the highest levels of education, where the participation returns are positive.

With the exception of Vietnam, falling fertility also contributed to higher participation rates in all countries. The effect was strongest in Brazil, South Africa, and Jordan. This is mainly because children are more strongly associated with lower participation in Brazil and South Africa; hence a decline in the number of children accounts for a larger increase in the observed participation rate. Rising household incomes contributed to a decline in participation in Tanzania, Bolivia, India, Vietnam, and Indonesia. India is the only country in our sample where rising household head education made a significant

---

<sup>38</sup>For Tanzania, results depend on the choice of counterfactual. Using the 2000 coefficients, the covariate effect

drive the negative covariate effect, being partly offset by the positive effect of rising female education. In 2014, the negative average marginal effect of income shrinks by two thirds relative to 2000. As a result, the total covariate contribution becomes positive when weighted at 2014 coefficients.

<sup>39</sup>In addition, we decompose the FLFP change in South Africa for the full post-*apartheid* period: 1995–2014. Participation rates of urban married women rose substantially from 58.5 percent to 68.1 percent between 1995 and 2001. We find that women's labor market characteristics account for around 70–74 percent of this increase (Table 2.A.23). Rising education, declining fertility, and a relative increase in the share of black women (in urban areas) were powerful drivers of participation in this period.

negative contribution to FLFP rates. Finally, we also find a relatively strong (negative or positive) contribution of changes in the returns to characteristics and unobserved factors in several countries. The direction and relative importance of this component vary widely across countries, and it does not appear to be related to countries' income level or the observed level or change in FLFP rates.

## 2.4.2 Decomposing differences between countries

We next decompose FLFP differences between countries, using Brazil as the reference country. The decomposition shows the extent to which gaps in participation rates between a particular country and Brazil emanate from differences between women's observed characteristics versus differences in the returns to those characteristics (or other unobservables). We take Brazil as the counterfactual for two main reasons. First, having the highest per capita income in our sample, it constitutes a natural benchmark. Second, having the second highest increase in FLFP—entirely accounted for by changes in covariates—it is of particular interest to assess to what extent other countries' participation rates differ from Brazil's due to differences in covariates.

We run two sets of decompositions: *first year*, which uses covariates and coefficients from Brazil's 2002 survey and the other country's data from the survey year closest to 2002; and *last year*, which uses covariates and coefficients from Brazil's 2013 survey and the other country's data from the survey year closest to 2013.

The exercise requires a few data adjustments. First, we recode the educational attainment of the woman and household head into four broader categories (less than primary, primary completed, secondary completed, and any tertiary) that are identical for all countries. Similarly, we recode the social group variable—reflecting ethnicity, religion, or nationality—into a dummy variable equal to 1 for the social groups with positive average marginal effects on participation within each country and 0 otherwise.<sup>40</sup> To capture regional effects in a comparable way, we compute, for each country and period, the quartiles of the regional average marginal effects on participation. We then create a dummy variable for each quartile.<sup>41</sup> Finally, we do not use survey weights for the decompositions. Countries use different sampling strategies in their surveys, and often follow distinct approaches in calculating sample weights. We want to avoid that such methodological differences drive the results. Overall, the unweighted difference in FLFP between each country and Brazil comes very close the weighted difference (Figure

<sup>40</sup>Table 2.A.24 shows how the education and social group variables are created for each country.

<sup>41</sup>Excluding the regional dummies altogether does not change the decomposition results in any meaningful way.

2.A.2).<sup>42</sup>

The results are clear: the observed differences to Brazil's FLFP are mostly accounted for by differences in coefficients and unobservables (Figure 2.12).<sup>43</sup> The covariate contribution is always negative and relatively small, with the exception of South Africa and Vietnam, where it is positive in the first year. The largest negative contributions come from children and education: in most countries, households have more children, and women achieve lower educational attainment than in Brazil. As a result, participation rates are lower than they would be if fertility and education were at the Brazilian level. For example, if the number of children per household in Jordan would decline to the level of Brazil, FLFP would increase by 5 to 6 percentage points. Household head education (including the effect of female household headship) is also sizable and negative in Jordan, India, Indonesia, and Bolivia; in some of these countries, household heads are more educated than in Brazil (while household head education is associated with lower FLFP in Brazil), and in others there are much fewer women who head their households (while, in Brazil, female household headship is positively related to participation).

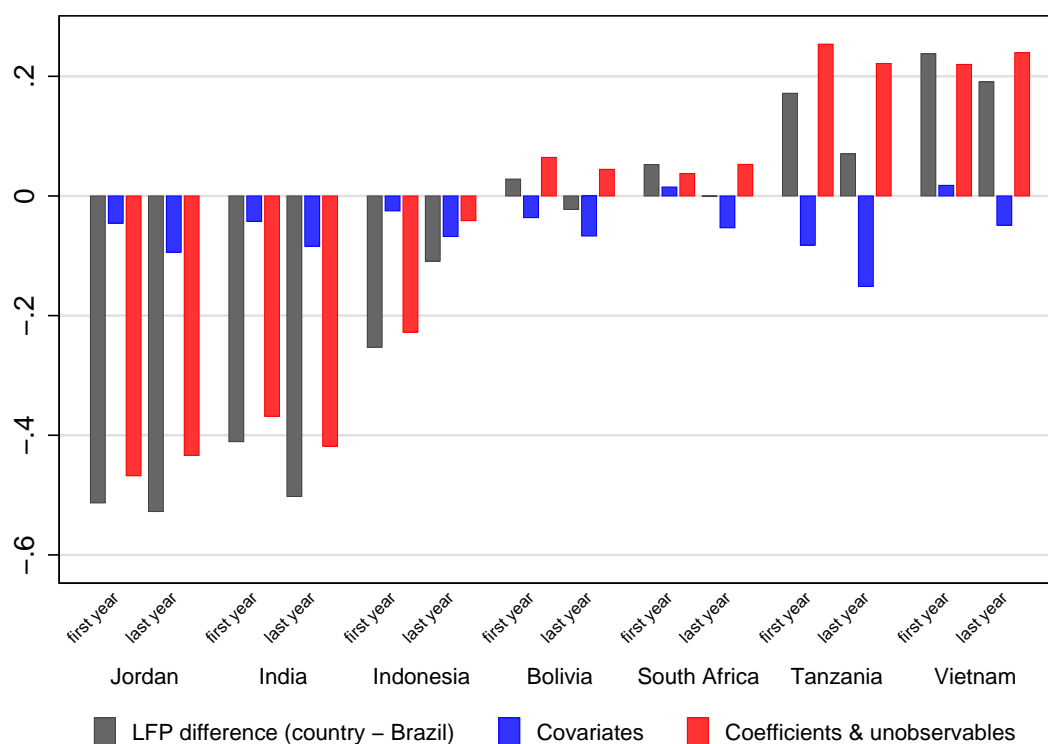
The overall negative covariate contribution is in the right direction for Jordan, India, and Indonesia (which have lower participation rates than Brazil), but much smaller than the observed difference. For the remaining countries, in at least one of the periods, the sign of the covariate contribution differs from the sign of the actual FLFP gap. In sum, differences in covariates between countries fail to explain the magnitude of differences in FLFP; in some cases, they even fail to predict the *sign* of the FLFP differences between countries.

Another way to appreciate this result is to imagine that all countries operate in a single (fictional) labor market. All women face the same returns to covariates and share the same level of any other relevant unobservables, irrespective of their country of origin: these coefficients and unobservables are equal to those of Brazil. Otherwise, each woman has her own observable characteristics as given in the data.

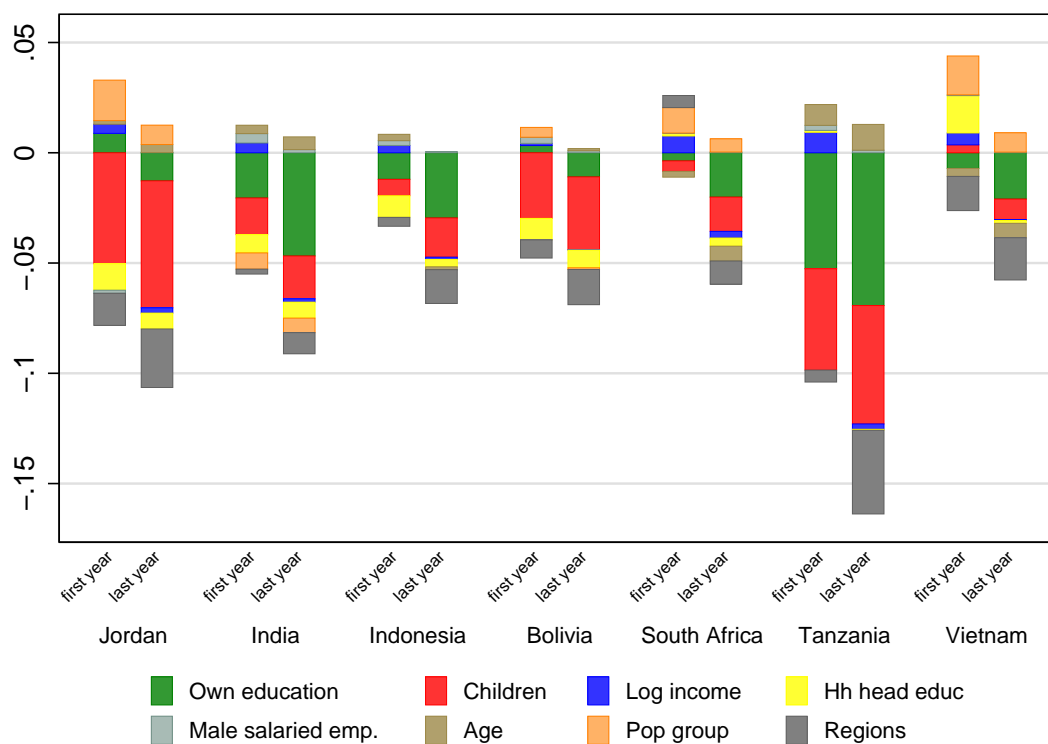
What would be the labor force participation in this fictional "Brazilian"-like labor market? Figure 2.13 shows the answer: most of the observed FLFP differences would disappear. In the first year, the lowest participation rate would be 54 percent in Tanzania; the highest would be 64.4 percent in Vietnam. The average labor force participation would be 60 percent, with a standard deviation of 3 percent. In the last year, the average FLFP would rise to 61 percent (standard deviation of 4 percent). Compare these numbers

<sup>42</sup>The only exception is Tanzania where the weighted FLFP rate is much higher than the unweighted rate. Accordingly, we interpret the results for Tanzania with caution.

<sup>43</sup>For point estimates of the between-country decompositions, see Tables 2.A.25 and 2.A.26.



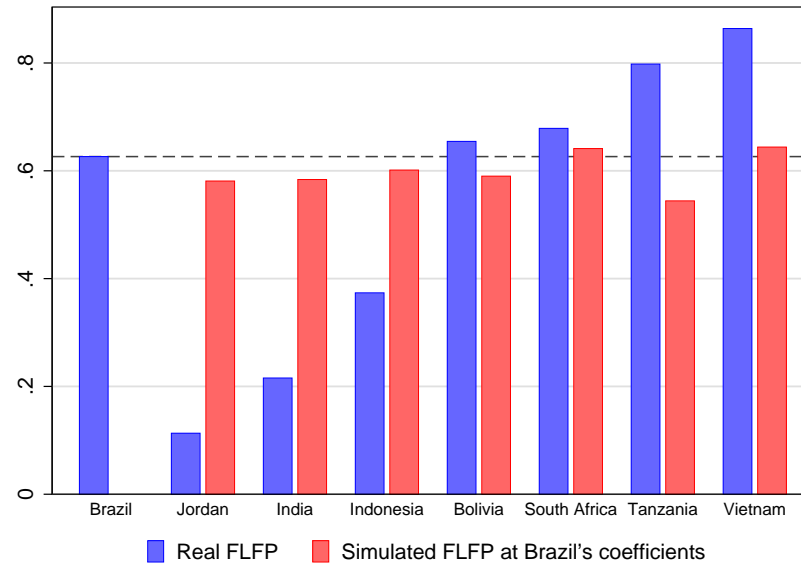
(a) Total covariate contribution



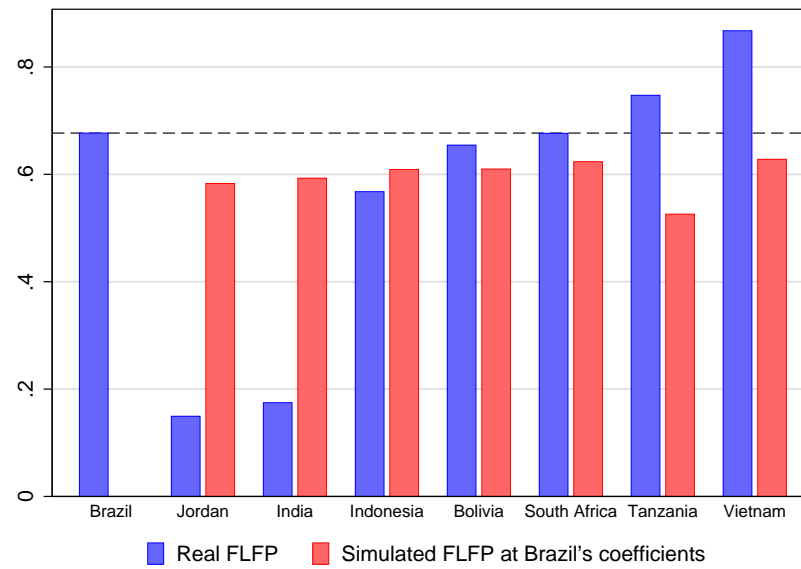
(b) Contribution of variable groups

**FIG. 2.12:** Decompositions between countries

Notes: Brazil is the reference country for all pairwise decompositions. For point estimates of the decompositions, see Tables 2.A.25 and 2.A.26. *First year* and *last year* refer to the year of the coefficients used to compute the covariate contribution.



(a) First year, around 2002



(b) Last year, around 2014

**FIG. 2.13:** Real FLFP vs. FLFP simulated at Brazil's coefficients



to reality: the mean FLFP in the first year (last year) was 54 (56) percent; ranging from 11 (15) percent in Jordan to 86 (87) percent in Vietnam. The standard deviation was 27 (26) percent.<sup>44</sup>

In the fictional “Brazilian” market, Jordanian women would have a *higher* participation rate than women from Tanzania. In reality, in 2014, the participation rate for Jordan was a staggering 67 percentage points *lower* than in Tanzania. In brief: differences in the observed characteristics of women and their households cannot account for the wide variation in FLFP between countries. Instead, most of the between-country differences result from variation in the returns to those characteristics and other unobservable factors.

## 2.5 Conclusion

Using comparable microdata from eight low and middle-income countries, this chapter sheds light on the impact and relative importance of what are considered key determinants of FLFP. We find that the participation-returns to women’s *own characteristics* and *family circumstances*—including education, income, and fertility—differ substantially across countries. In fact, heterogeneity in returns to these characteristics explains most of the between-country differences in participation rates, indicating that the economic, social, and institutional constraints that shape women’s labor force participation are still largely country-specific.

Nonetheless, some important patterns appear. Overall, rising education levels and declining fertility consistently increase FLFP, although the strength of these two forces differs across countries. At the same time, rising household incomes have a negative effect in all but the three richest countries in our sample (Jordan, South Africa, and Brazil), indicating that, in poorer countries, a substantial share of women work out of economic necessity.

In relatively poor countries with high initial participation rates (Vietnam, Tanzania, and, to a lesser extent, Bolivia), improving *family circumstances* (e.g., higher household incomes, or better educated household heads) have a moderate negative effect on women’s participation. In terms of *women’s own characteristics*, the positive participation-education gradient is flattening over time, except for relatively high participation returns occurring at the tertiary level. Future gains in female participation rates will depend on the extent to which women achieve educational attainment at the tertiary level.

---

<sup>44</sup>The FLFP rates in the paragraph are calculated *without* survey weights.

In countries with low initial participation rates and strong social barriers to women's outside-home employment (India, Jordan, and, to a lesser extent, Indonesia), *family circumstances* have a much stronger grip on women's participation. Own education has a U- or J-shape relationship with participation, such that rising attainment at intermediate education levels actually depresses FLFP. Once again, expansion of women's access to tertiary education would be required to raise FLFP further. As shown by the Indonesian experience, however, changing returns to women's labor supply characteristics can boost participation rates dramatically.

In the richest countries (Brazil and South Africa), where social barriers to women's employment are relatively small, *family circumstances* other than fertility have become largely irrelevant. With a strong positive education-participation gradient and a strong negative effect of fertility, increases in women's own education and falling fertility boosted participation in these countries. In the future, higher educational attainment and lower fertility will likely continue to translate into higher FLFP.

Finally, we find suggestive evidence of reduced selectivity of tertiary education in India, Indonesia, and South Africa (similar to the results for India in Klasen and Pieters (2015)). This may mitigate the extent to which further educational advancement will translate into higher FLFP in these countries.

While this chapter has focused on supply side factors, FLFP might be severely constrained by demand factors. For example, the unexplained portion of the gender wage gap did not decline substantially in recent decades (Weichselbaumer and Winter-Ebmer, 2005; Oostendorp, 2009), and employment sectors and occupations remain highly segregated by gender (Borrowman and Klasen, 2017). Further improvement of women's labor market characteristics will likely have a limited effect in rising FLFP rates, unless accompanied by the removal of barriers and constraints to female employment both at the household and at the labor market level.

## 2.A Appendix

### Data

In this Appendix, we describe in detail the data sources used, the coding procedures that ensure comparability across countries and years, and the limitations of the final dataset. But first we describe three coding procedures that apply to all countries, concerning household income, male salaried employment, and regional employment.

**Household income** For comparability across surveys, the real earnings variable only considers earnings from the respondent's main job. Income from secondary jobs and non-labor income are not covered. To avoid losing the observations of urban married women living in households where the sum of earnings for all the other adult members is zero, we add one dollar-PPP to all household incomes before the log-transformation. If at least one household member is employed but has missing earnings, we code the household-level earnings variable as missing.

**Male salaried employment** To capture the existence of a stable male source of income in the household, we would ideally identify male wage employment in the economy's formal sector, since the stability and certainty of formal wage contracts are higher than that of informal labor arrangements. Unfortunately, information of the type of employment sector (formal vs. informal) is not available for several countries. As a second best, we create a household-level dummy variable for the existence of at least one adult (18+) man working as a salaried employee, irrespectively of formality.

**Regional employment** To create regional employment shares, we use one-digit industrial codes from the surveys. Most surveys follow the international ICIC code classification, although there are a few minor deviations. We re-group the one-digit codes into broader categories: (1) *Construction*, corresponding to "construction"; (2) *Agriculture*, including "agriculture, hunting and forestry" and "fishing"; (3) *Mining*, corresponding to "mining and quarrying"; (4) *Manufacturing*, corresponding to "manufacturing"; (5) *White-collar services*, including "financial intermediation," "real estate, renting and business activities," "public administration and defense; compulsory social security," "education," "health and social work," "extra-territorial organization and bodies"; and (6) *Other services*, including "electricity, gas and water supply," "wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods," "hotels

and restaurants,” “transport, storage and communication,” “other community, social and personal service activities,” and “activities of households.”

### South Africa

For South Africa, we use five nationally representative labor force surveys from Statistics South Africa (StatsSA): the October Household Survey (OHS) in 1995, the Labor Force Survey (LFS) in 2001 and 2003, and the Quarterly Labor Force Survey (QLFS) in 2010 and 2014 (see Table 2.A.1). We extract the data from the Post-Apartheid Labour Market Series (PALMS) v3.1, a dataset created by Data First at the University of Cape Town (Kerr et al., 2016), which provides consistent coding for a large number of variables across South African labor force surveys between 1994 and 2015 (Kerr and Wittenberg, 2016). We selected the years as to guarantee a good time coverage, under the constraint that the relevant set of variables used in the empirical analysis exists in a consistent way across surveys. Between 2004 and 2008, the LFS does not contain urban/rural information and the QLFS does not provide continuous earnings data in 2008, 2009, and 2012. This is why we do not select surveys from those years.

Naturally, the coding of certain variables in the original surveys changes over time.<sup>45</sup> For the years 1995, 2001 and 2003, the urban/rural variable has two possible categories: *urban* or *rural*. For the QLFS 2010 and 2014, the corresponding variable is coded in four categories: urban formal, urban informal, tribal areas, and rural formal. To create a consistent variable, we assign both urban formal and urban informal to the category *urban* and tribal areas and rural formal to the category *rural*.

For the individual-level regressions, we drop the 37 urban married women of age 25–54 whose race category in the LFS 2001 and 2003 is “Other”, because this racial category does not occur in any of the other selected years.

We create the highest level of education achieved by recoding the detailed education achievement variables of the PALMS into five broader groups: less than primary; primary completed; secondary not completed; secondary completed; any tertiary. We drop from the analysis the few individuals categorized in the residual educational level “Other”. The currently married dummy variable assigns value 1 to persons officially married or “living together as husband and wife”.

Until the introduction of the QLFS in 2008, there were two clear definitions of labor force participation. The *strict* definition from the OHS and LFS includes (1) the

<sup>45</sup>Yu discusses in detail the comparability between the OHS and the LFS (Yu, 2007), and between the LFS and QLFS (Yu, 2009).

employed—those who have worked in the week before the survey interview, or are temporarily not working but will return soon (on holidays, sick leave, parental leave, strike, etc)—and (2) the unemployed—those currently without a job but who are willing to accept one within a week from the survey interview and have actively searched for work during the previous month.<sup>46</sup> In the OHS 1995, any positive amount of hours worked in the reference week counts for employment; in the LFS, employment requires a minimum of one hour of work in the reference week. Employment includes all paid employees either in cash or kind, employers and self employed, as well as unpaid workers in family businesses. Unpaid domestic services and begging for money or food do not count as employment. It is clear from the change in survey questionnaires that the LFS is better able to capture informal, casual employment relative to the OHS: there are eight detailed employment categories allowed as answers in the LFS, whereas only three (working full-time, working part-time, with a job but absent from work) are allowed as answers in the OHS. Thus, part of the rise in labor force participation from 1995 to 2001 could reflect improvements in the coverage of casual, low-income employment in the LFS relative to the OHS (Yu, 2007, pp. 17–18). The *expanded* definition is similar to the *strict* one but it additionally includes the discouraged job-seekers in the labor force, as part of the unemployed population. Discouraged job-seekers are those individuals who are currently without a job and, although they desire to work, have not actively searched for a job in the month before the survey interview. Unfortunately, with the introduction of the QLFS, the distinction between the strict and expanded concepts becomes less clear (Yu, 2009; Kerr and Wittenberg, 2016). Yu (2009) shows that comparability between the last LFS (in 2007) and the first QLFS (in 2008) is much better for the strict definition than for the expanded definition. Accordingly, we consider *strict* labor force status in our analyses.

The household labor income variable is created from the consistent real earnings variable in the PALMS dataset for wage employment and self-employment.<sup>47</sup> However, a substantial fraction (20–30 percent) of employed individuals in the OHS 1995, LFS 2001 and 2003 did not report their earnings as a point value, but rather used the earnings brackets available in the questionnaire. Individuals who are self-employed or in high skilled occupations disproportionately used the bracket option. As a result, creating an household income variable without accounting for bracket responses will underestimate

<sup>46</sup>Unfortunately, the OHS1995 metadata does not include a detailed explanation on how the employment status is derived from the survey questions (Yu, 2009, pp. 17 & 49–58). However, the derivation of employment is very similar for the later OHS (1996–1999), so we refer to it in the text.

<sup>47</sup>See Burger and Yu (2006) for more details on constructing a consistent earnings series from the OHS and LFS surveys.

the household incomes of such individuals. As a simple solution, whenever necessary, we impute earnings with the bracket midpoint and, for the top bracket, which is open-ended, we set the midpoint to be 10 percent higher than the lower bound. Von Fintel (2007) shows that, for the LFS, this “simplistic” midpoint imputation performs as well (for purposes of statistical inference) as more complex distributional assumptions (e.g. interval regressions), given that skewness is not too extreme and the share of right-censored observations is not too high.<sup>48</sup> For the years 1995, 2001 and 2003, we calculate the midpoint of the corresponding bracket variable from PALMSv3.1, convert it into monthly earnings and deflate it to 2000 Rands.<sup>49</sup> For the year 2010, Stats SA already imputes refusals and categorical responses; for 2014, only categorical responses are imputed by Stats SA. Unfortunately, the imputation methods used by Stats SA are not described in the surveys’ metadata (Kerr and Wittenberg, 2016). After the imputations described, around 7 percent of employed individuals in the pooled sample have missing earnings information. For the household-level income variable, we sum up the earnings of all individuals in the household, with the exception of those households where at least one employed individual has missing earnings, for whom we assign missing household income. Finally, we convert the values to international dollars using the World Bank’s 2011 PPP exchange rate for private consumption.

A limitation of the PALMS is the absence of information on the relationships between household members. As a result, there is no information on who is the household head. Given these limitations, we cannot capture the household’s lifetime income using the household head’s or the husband’s education as a proxy. In practice, we use the maximum education level of any adult (18+) married male household member, coding as an explicit missing category those households where no adult married men are listed.

We aggregate employment at the province level by industry using the classification of Klasen and Pieters (2015). For South Africa, this means recoding the PALMS variable *jobindcode*. When in doubt about whether a particular industry should be classified as blue or white collar, the education distribution of urban married female employees was used as an auxiliary tool: thus, the seemingly ambiguous “Services” category in the raw data was included in the white-collar services category due to the much higher

<sup>48</sup>The number of right-censored observations for each wave of the LFS 2001 and 2003 is low, with a minimum of 38 observations in the first wave (March) of LFS 2001 and a maximum of 83 observations in the second wave (September) of LFS 2003.

<sup>49</sup>For the year 1995, the lowest earning bracket for wage employment is too wide (R1 - R999). As a result, all of the observations reporting *daily* earnings and 97.23 percent of the observations reporting *weekly* earnings fall in this category (c.f. with 26 and 0.4 percent of the observations for *monthly* and *yearly* wage earnings). Therefore we do not impute a midpoint value for the 531 observations with daily or weekly reference periods; they are set to missing. We proceed similarly for self-employment earnings.

prevalence of highly educated employees (see, e.g., Figure 2.A.1a).

For the estimation of population means, average marginal effects of regression covariates, and decomposition analyses, we use the individual cross-entropy weights available in the PALMSv3.1.<sup>50</sup>

## Brazil

For Brazil, we use four yearly household surveys called Pesquisa Nacional por Amostra de Domicílios (PNAD) from the Brazilian Institute of Geography and Statistics (IBGE): PNAD 2002, 2005, 2009 and 2013. The surveys are harmonized using the Stata code created by Data Zoom at the PUC-Rio.<sup>51</sup>

A particular feature of the PNAD surveys is the distinction between different family units within a given household (see Alves, 2005). For example, multigenerational households are usually classified as different families living in a single household. We code as currently married family heads and their spouses, including couples who are officially married or living together as husband and wife. Otherwise, for consistency with surveys from the other countries, we construct all household-level variables using the household identifiers, disregarding their sub-classification into families.

There were several education reforms in the past three decades. As a result, some levels of education attainment changed names and duration. We reclassify these different levels into five broader groups: less than primary; elementary (levels 1–4); elementary (levels 5–8); high school completed; any tertiary. We proxy the household's lifetime earnings potential with the education level of the household head, creating an additional missing category for the cases when the married woman is the household head, as her education level is already captured in the own education variable.

There are two reference periods available to define labor force participation: previous week or previous year. For consistency with most of the other countries, we use labor force participation in the week of reference (last week of September) in the empirical analyses.

In the PNAD, employment status covers all individuals of age 10 or above that work in: (1) paid activities; (2) unpaid activities in support of a self-employed or employer household member in the production of primary goods; (3) unpaid activities in support of a religious institution or cooperative; (4) food production and/or construction work

<sup>50</sup>Variable `cweight2`. This is an update by Takwanisa Machemedze at DataFirst, University of Cape Town of the original cross-entropy weights created by Nicola Branson. See Branson and Wittenberg (2014) for details on the cross-entropy approach.

<sup>51</sup>Available at <http://www.econ.puc-rio.br/datazoom/english/index.html>.

for own consumption. In category (1), there is no restriction on the minimum hours worked per week, whereas for the other categories, the individual must work at least one hour in the reference week to be considered employed (IBGE, 2015).

For consistency with data from other countries, we add up real earnings from the individual's main job, excluding the woman's own earnings to construct the household income variable and convert it to 2011 PPP-\$ using the World Bank PPP exchange rate for private consumption.

The coding of employment industries follows Klasen and Pieters (2015) and is done at the federal state level.

The PNAD also allows us to partially capture internal migration. We create dummy variables recording whether the individual was born in the state of current residence, and whether the individual was born in the municipality of current residence. We use these variables as a rough control for rural-urban migration: to the extent that rural-urban movements cross state or municipal borders, they are captured in our migration variables.

## Jordan

For Jordan, we use four Employment and Unemployment Surveys (EUS) collected by Jordan's Department of Statistics and harmonized by the Economic Research Forum (ERF) covering the years 2006, 2008, 2010, and 2014 (OAMDI, 2015a,b,c,d). The surveys are collected on a quarterly basis, resulting in four waves per year. We use all waves available.

Unfortunately, there is no urban/rural information in the EUS for the years 2008 and 2014, so we cannot remove rural households as we do in all other countries. In any case, urbanization rates are very high in Jordan, in the period considered. According to the World Bank's World Development Indicators<sup>52</sup>, the urbanization rate was 81 percent in 2006 and 83 percent in 2014.

The variable currently married includes both monogamous and polygamous marriages, although the latter account for only 0.15 percent of the pooled (unweighted) sample. All monogamous marriages are also reported as being legal marriages. Informal marriages and cohabitation outside marriage are both very rare in Jordan.

We recode the detailed education variable provided by the ERF into six broader attainment categories: less than primary completed; primary completed; preparatory completed; lower secondary completed; secondary completed; any tertiary. The lifetime

---

<sup>52</sup> Accessed on November 8, 2017.



earnings potential of the household is proxied with the education level of the its head, creating an additional missing category for the cases when the married woman is the household head, as her education level is already captured in the own education variable.

All individuals above age 15 are coded as being in the labor force if they were employed for at least one hour in the week before the survey interview, or if they are unemployed (i.e., no current job, but actively searching). Those who have a job but were temporarily absent during the reference week (e.g., on holidays, sick leave, maternal leave, striking) are also part of the employed population. Employment includes employees who work for wage or in-kind payments, employers, self employed, and unpaid workers in family enterprise. Unpaid domestic services or unpaid voluntary work do not count as employment. Those without jobs and not searching one in the reference week are coded as inactive.

The sampling frame and stratification strategy of the EUS changed in the first quarter of 2007. Assaad et al. (2014) argues that this change resulted in a break of the female labor force participation series: before 2007, female labor force participation would have been underestimated by around three percentage points. To extend time coverage, we nevertheless include the EUS 2006, which has still conducted under the old methodology.

The EUS only provides categorical earnings data using five earning brackets. During the harmonization of the data, the ERF imputed a continuous earnings variable with the midpoint of each bracket. This imputation procedure is also done for some cases in the South African data, as described above, so we do not modify it. We inflate the earnings from the individual's main job using monthly CPI indexes from the Jordanian Department of Statistics, with 2010 as the base-year. Finally, we sum up the individual real earnings excluding the woman's own earnings to construct the household income variable and convert it to 2011 PPP-\$ using the World Bank PPP exchange rate for private consumption.

The coding of employment industries follows Klasen and Pieters (2015). We are able to reclassify the residual category *Other* in the raw data (variable *ind*) into the other more meaningful categories using additional variables with ISIC revision 3, 3-digit codes (variable *ind\_isic3\_3*) for the years 2006 and 2008, and ISIC revision 4, 3-digit codes (variable *ind\_isic4\_3*) for the years 2010 and 2014. Employment shares are then aggregated at the governorate level.

## Bolivia

For Bolivia, we use the household survey Encuesta de Hogares collected by the Bolivian National Institute of Statistics. We use the years 2000, 2005, 2008, 2011 and 2014.<sup>53</sup>

The variable currently married includes both married couples as well cohabiting couples that are living as husband and wife. We create a consistent educational attainment over time with five categories: less than basic; basic; intermediate; secondary; any tertiary. We proxy the household's lifetime earnings potential with the education level of the household head, creating an additional missing category for the cases when the married woman is the household head, as her education level is already captured in the own education variable.

Unfortunately, it is difficult to construct a consistent ethnicity variable across surveys due to changes in coding and lack of appropriate codebook for the year 2014. We opt for an approximation based on the self-reported languages spoken: we create a dummy dividing the surveyed individuals into those only speaking Spanish, and those speaking Spanish and an Indigenous language. Very few individuals speak no Spanish at all, so we add them to the latter category.

A labor force participant is anyone above age seven who was either employed for at least one hour in the week before the survey interview, or unemployed (i.e., no current job, but actively searching). Employment encompasses the production of goods and services for the market or the production of goods for own consumption. Those who have a job but were temporarily absent during the reference week (e.g., on holidays, sick leave, maternal leave, striking) are also part of the employed population. Unpaid domestic services, voluntary service work or other unpaid work for a salaried family members do not count as employment. Those without jobs and not actively searching are coded as inactive.

The income variable is constructed from individual-level monthly earnings from the main job, as done for the other countries. We recode the missing earnings of unpaid family workers or unpaid apprentices with zero. We then inflate the earnings to 2010 prices. Finally, we sum up the individual real earnings excluding the woman's own earnings to construct the household income variable and convert it to 2011 PPP-\$ using the World Bank PPP exchange rate for private consumption.

The coding of employment industries follows Klasen and Pieters (2015) and employment shares are aggregated at the department level.

Although information on rural-urban migration is unavailable, we measure (overall)

---

<sup>53</sup>In 2000, the survey was known as Encuesta Continua de Hogares.

migration with a dummy variable for whether the woman was living somewhere other than the municipality of current residence, five years before the survey. There is a change in the wording of the question used to derive this dummy variable. In 2000 and 2005, the question asks: *Between [5 years before survey year] and [survey year], did you live somewhere else?* After 2008, the question asks: *Where did you live five years ago?*

## Vietnam

For Vietnam, we use four years of a national representative general purpose household survey: the Vietnam Household Living Standards Survey (VHLSS) 2002, 2006, 2010, and 2014.

For the education variable, we classify the highest completed school grade into five broader education levels, based on the Vietnam education system, namely: less than primary education, completed primary, completed lower secondary, completed high school, and any tertiary education. We proxy the household's lifetime earnings potential with the education level of the household head, creating an additional missing category for the cases when the married woman is the household head, as her education level is already captured in the own education variable.

Vietnam has approximately 56 ethnic groups, but around 88 percent of household heads are Kinh. In our analysis, we create a dummy with two categories: Kinh and non-Kinh.

All individuals aged 10 or older were asked to state whether they had a job in the 12 months before the survey interview. Having a job is defined as working as a wage earner, or being self-employed in agriculture or non-farm activities. Unfortunately, after 2002, we cannot distinguish between those not working but actively searching for a job. That is, we cannot distinguish between the unemployed and the inactive. We thus define labor force participants as those having a job in the previous 12 months. For 2002, however, there is a job search question, so we can construct a labor force participation variable that classifies the unemployed as active. For this year, the difference of participation rate between the two definitions of labor force is very small (around 2 percent). Notice that the reference period for the job question (12 months) is much longer than the reference period in most other countries (usually, one week). Thus, on the one hand, excluding the unemployed from the active population will, in general, underestimate the rate of labor force participation. But, on the other hand, unemployment will be much lower with such a long reference period. In the end, at least for the year 2002, the two effects seem to cancel out.

Whenever employed individuals have missing earnings, we impute them using a simple hotdeck procedure for each year separately based on age (5-years groups, from 16–20 to 61–65, and 65+), gender, educational attainment, and rural/urban. Finally, we sum up the individual real earnings excluding the woman's own earnings to construct the household income variable and convert it to 2011 PPP-\$ using the World Bank PPP exchange rate for private consumption.

The coding of employment industries follows Klasen and Pieters (2015) and employment shares are aggregated at the province level.

### **Tanzania**

For Tanzania, we use the Integrated Labour Force Surveys 2000–01, 2005–06 and 2014 collected by the National Bureau of Statistics. These are quarterly surveys and we use all four waves for each year.

The coding of the education variables is slightly different from the one used in other countries. Given the lower education attainment levels compared with the remaining countries, we do not distinguish between completed and not completed secondary schooling. We create five attainment levels: never attended school; primary not completed; primary completed; any secondary schooling; any tertiary. We proxy the household's lifetime earnings potential with the education level of the household head, creating an additional missing category for the cases when the married woman is the household head, as her education level is already captured in the own education variable.

We did not identify any variable capturing a meaningful social identity (or discriminatory) marker such as the ethnicity or religion markers used for other countries.

A labor force participant is anyone above age 10, in 2000 and 2006, and 15, in 2014, who was either employed for at least one hour in the week before the survey interview, or unemployed (i.e., no current job, but actively searching). Employment is defined as working for cash or in-kind pay, employers and self employed, unpaid family workers in family enterprises, production of primary products for own consumption, or production of other fixed assets (including housing) for own use. Those who have a job but were temporarily absent during the reference week (e.g., on holidays, sick leave, maternal leave, striking) are also part of the employed population. Unpaid domestic services are excluded from this definition. Those without jobs and not actively searching are coded as inactive. Notice that, according to the employment definition above, individuals engaged in subsistence agriculture are part of the labor force. Even in urban areas,

agriculture accounts for a substantial share of Tanzanian workers.

The income variable is constructed from individual-level monthly earnings from the main job, as done for the other countries. We recode as zero the missing earnings of unpaid family workers, and farmers reporting no marketed produce. We then inflate the earnings to 2010 prices. Finally, we sum up the individual real earnings excluding the woman's own earnings to construct the household income variable and convert it to 2011 PPP-\$ using the World Bank PPP exchange rate for private consumption.

The coding of employment industries follows Klasen and Pieters (2015). We aggregate employment shares at the region level. Due to administrative reforms in 2003 and 2012, new regions were created. In 2014 there were five more regions than in 2000. By merging regions whenever necessary, we obtain administrative units that are consistent over time.

We can identify rural-urban migrants, as well as years since arrival at the urban area. We create two dummy variables distinguishing between rural-urban migrants whose arrival was less than five years before the survey and rural-urban migrants for which more than five years have passed since arrival.

## India

For India, we use three rounds of the NSS Employment and Unemployment Survey: round 55, 1999–2000; round 61, 2004–2005, and round 68, 2011–2012. These are three out of the six rounds used by Klasen and Pieters (2015) and, for more details on the data, see their article, including its online appendix.

We code the educational attainment into six categories: illiterate, literate, primary school completed, middle school completed, secondary school completed, any tertiary schooling. We proxy the household's lifetime earnings potential with the education level of the household head, creating an additional missing category for the cases when the married woman is the household head, as her education level is already captured in the own education variable.

To capture the individual's social group, we create a variable combining religion and caste with four categories: Hindus non-SCST<sup>54</sup>; SCST; Muslim non-SCST; and Other, which is a residual category.

Labor force participation is defined by the individual's main activity status, and its reference period is one year before the interview.

Income is measured as earnings from the main job in the reference week. The earnings

---

<sup>54</sup>SCST stands for "scheduled castes or scheduled tribes".

of self-employed individuals are missings. We impute them based on the earnings of employees (see Klasen and Pieters, 2015, footnote 11). We convert the individual earnings into monthly earnings and deflate them to 1999–2000 prices using state-specific CPIs (see Klasen and Pieters, 2015, for sources and more details). We then sum up the individual real earnings excluding the woman’s own earnings to construct the household income variable and convert it to 2011 PPP-\$ using the World Bank PPP exchange rate for private consumption.

As in Klasen and Pieters (2015), we aggregate employment shares at the district level.

## Indonesia

For Indonesia, we use the annual national household survey, Susenas, for the years 2000, 2004, 2007, and 2014.

We code the educational attainment into five categories: less than primary, primary completed; junior high school completed; senior high school, and any tertiary schooling. We proxy the household’s lifetime earnings potential with the education level of the household head, creating an additional missing category for the cases when the married woman is the household head, as her education level is already captured in the own education variable.

We did not identify any meaningful ethnicity or religion variable that was consistent across survey years.

A labor force participant is anyone above age 10 who, in the week before the interview, (i) has worked any amount of time, or (ii) was temporarily absent from a permanent job, or (iii) did not have a job but has actively searched for one. Those who do not meet any of the conditions above are classified as inactive. In 2004, besides the three conditions for labor force participation defined above, there is an additional question asking whether, in the reference week, the individual (iv) was preparing to start a business. We include (iv) as being part of the labor force. In 2007 and 2014, (iii) and (iv) are included in a single survey question. In 2000, (iv) is not explicitly asked.

In 2000, 2004, and 2007, monthly earnings are missing for the self-employed. We impute them using a simple hotdeck procedure for each year separately based on age (5-years groups, from 16–20 to 61–65, and 65+), gender, educational attainment, rural/urban, and province. Finally, we sum up the individual real earnings excluding the woman’s own earnings to construct the household income variable and convert it to 2011 PPP-\$ using the World Bank PPP exchange rate for private consumption.

The coding of employment industries follows Klasen and Pieters (2015). We aggregate

employment shares at the regency level. We use the borders as of 1998, to obtain units that are consistent over time.

## Additional Tables and Figures

**TABLE 2.A.1:** Data overview

Country	Year	Survey	N <sup>†</sup>
<i>South Africa:</i>	1995	October Household Survey	8,262
	2001	Labor Force Survey	12,862
	2003	Labor Force Survey	12,050
	2010	Quarterly Labor Force Survey	21,438
	2014	Quarterly Labor Force Survey	20,744
<i>Brazil:</i>	2002	Pesquisa Nacional por Amostra de Domicílios	46,562
	2005	Pesquisa Nacional por Amostra de Domicílios	48,637
	2009	Pesquisa Nacional por Amostra de Domicílios	49,360
	2013	Pesquisa Nacional por Amostra de Domicílios	45,423
<i>Jordan</i> <sup>‡</sup> :	2006	Employment and Unemployment Survey	26,140
	2008	Employment and Unemployment Survey	33,629
	2010	Employment and Unemployment Survey	32,993
	2014	Employment and Unemployment Survey	30,593
<i>India:</i>	1999	NSS Employment and Unemployment Survey	33,507
	2004	NSS Employment and Unemployment Survey	30,489
	2011	NSS Employment and Unemployment Survey	28,252
<i>Bolivia:</i>	2000	Encuesta Continua de Hogares (MECOVI)	1,563
	2005	Encuesta de Hogares	1,283
	2008	Encuesta de Hogares	1,183
	2011	Encuesta de Hogares	3,113
	2014	Encuesta de Hogares	3,863
<i>Indonesia:</i>	2000	Susenas	51,363
	2004	Susenas	73,447
	2007	Susenas	75,713
	2014	Susenas	87,462
<i>Vietnam:</i>	2002	Living Standard Survey	5,281
	2006	Household Living Standard Survey	1,704
	2010	Household Living Standard Survey	1,970
	2014	Household Living Standard Survey	2,043
<i>Tanzania:</i>	2000	Integrated Labour Force Survey	2,051
	2006	Integrated Labour Force Survey	2,899
	2014	Integrated Labour Force Survey	4,325

Notes: <sup>†</sup>Number of observations of urban married women age 25–54. Estimation samples are smaller due to missing covariate data. <sup>‡</sup>For Jordan, sample sizes refer to both urban and rural areas.

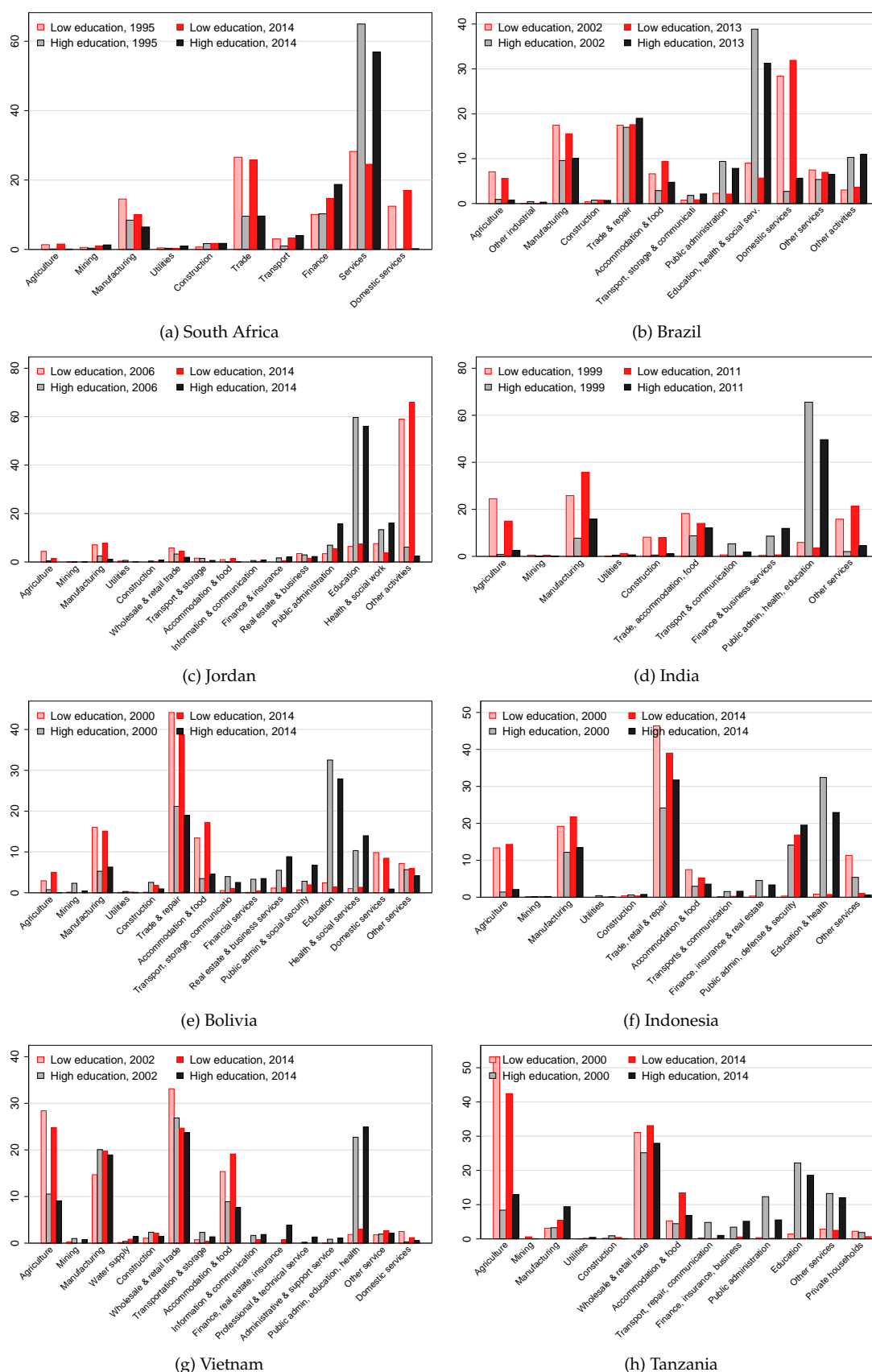


FIG. 2.A.1: Distribution of female workforce across industries, by education

Notes: See Table 2.A.1 for sources. Urban married women (25–54), except urban *and* rural in Jordan; employed only (including self-employed). Low education is below secondary schooling; high education is completed secondary or higher (*any* secondary or higher for Tanzania).



TABLE 2.A.2: South Africa: sample means

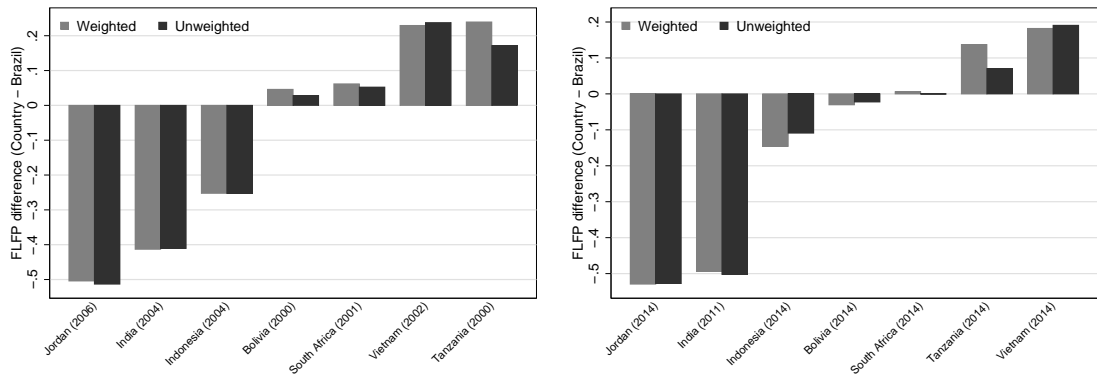
	1995	2001	2003	2010	2014
Labor force	0.58	0.68	0.67	0.67	0.68
<i>Own education:</i>					
Less than primary	0.15	0.16	0.15	0.09	0.07
Primary	0.08	0.07	0.06	0.05	0.04
Secondary not completed	0.39	0.37	0.37	0.36	0.35
Secondary completed	0.25	0.24	0.28	0.30	0.33
Tertiary	0.14	0.16	0.15	0.20	0.21
Log income	4.46	3.75	3.50	4.35	4.16
Male salaried emp.	0.77	0.65	0.63	0.66	0.65
<i>Max adult married male education:</i>					
Less than primary	0.14	0.15	0.14	0.09	0.08
Primary	0.05	0.06	0.06	0.04	0.03
Secondary not completed	0.36	0.32	0.32	0.31	0.30
Secondary completed	0.24	0.23	0.25	0.28	0.29
Tertiary	0.16	0.16	0.15	0.19	0.21
Missing: no adult married male	0.05	0.07	0.08	0.08	0.08
<i>Ethnicity:</i>					
Black	0.50	0.56	0.58	0.60	0.62
Coloured	0.15	0.14	0.15	0.15	0.15
Indian/Asian	0.06	0.06	0.06	0.05	0.05
White	0.29	0.23	0.20	0.20	0.18
Age	37.37	37.67	37.81	38.77	38.79
Children 0-4	0.56	0.52	0.50	0.44	0.41
Children 5-14	1.07	0.92	0.90	0.85	0.80
N	7,601	11,361	10,658	20,713	17,890

TABLE 2.A.3: Brazil: sample means

	2002	2005	2009	2013
Labor force	0.62	0.66	0.69	0.67
<i>Own education:</i>				
Less than primary	0.12	0.10	0.07	0.05
Elementary (1-4)	0.33	0.30	0.24	0.19
Elementary (5-8)	0.18	0.18	0.17	0.16
High school	0.23	0.26	0.31	0.35
Tertiary	0.14	0.16	0.20	0.24
Log income	4.90	4.91	5.10	5.27
Male salaried emp.	0.63	0.64	0.67	0.68
<i>Household head education:</i>				
Less than primary	0.13	0.11	0.08	0.05
Elementary (1-4)	0.32	0.30	0.23	0.18
Elementary (5-8)	0.16	0.16	0.14	0.13
High school	0.19	0.21	0.23	0.25
Tertiary	0.13	0.13	0.14	0.16
Missing: woman is hh head	0.06	0.08	0.17	0.23
<i>Ethnicity:</i>				
White	0.61	0.57	0.54	0.51
Black	0.05	0.06	0.07	0.08
Mixed	0.33	0.36	0.38	0.40
Asian	0.01	0.01	0.01	0.01
Indigenous	0.00	0.00	0.00	0.00
Age	37.80	38.14	38.38	38.63
Children 0-4	0.35	0.32	0.29	0.28
Children 5-14	0.92	0.87	0.78	0.71
N	39,193	42,189	42,855	38,596

TABLE 2.A.4: Jordan: sample means

	2006	2008	2010	2014
Labor force	0.12	0.16	0.17	0.15
<i>Own education:</i>				
Less than primary	0.14	0.10	0.09	0.07
Primary	0.14	0.11	0.12	0.12
Preparatory	0.22	0.22	0.19	0.20
Lower secondary	0.09	0.12	0.14	0.18
Secondary	0.16	0.17	0.16	0.14
Tertiary	0.24	0.28	0.30	0.29
Log income	4.30	4.40	4.47	4.12
Male salaried emp.	0.69	0.68	0.67	0.64
<i>Household head education:</i>				
Less than primary	0.11	0.09	0.08	0.06
Primary	0.15	0.12	0.12	0.13
Preparatory	0.26	0.27	0.25	0.26
Lower secondary	0.03	0.06	0.08	0.13
Secondary	0.15	0.15	0.14	0.13
Tertiary	0.27	0.29	0.30	0.26
Missing: woman is hh head	0.02	0.03	0.03	0.03
<i>Nationality:</i>				
Jordan	0.94	0.92	0.91	0.90
Iraq	0.01	0.02	0.01	0.01
Syria	0.01	0.01	0.01	0.06
Egypt	0.01	0.01	0.01	0.01
Other Arab countries	0.02	0.02	0.02	0.01
Other non-Arab countries	0.02	0.03	0.03	0.02
Age	37.41	37.31	37.39	37.82
Children 0-4	0.80	0.78	0.78	0.65
Children 5-14	1.90	1.74	1.63	1.55
N	25,834	33,219	32,691	30,458



(a) Brazil: 2002; closest year for other countries.

(b) Brazil: 2013; closest year for other countries.

**FIG. 2.A.2:** FLFP differences with respect to Brazil: with and without survey weights

Notes: See Table 2.A.1 for sources. Urban married women, ages 25-54. Except for Jordan: urban and rural areas.

TABLE 2.A.5: India: sample means

	1999	2004	2011
Labor force	0.18	0.21	0.18
<i>Own education:</i>			
Illiterate	0.32	0.28	0.22
Literate	0.08	0.07	0.07
Primary	0.12	0.13	0.11
Middle school	0.14	0.16	0.15
Secondary	0.21	0.22	0.27
Tertiary	0.13	0.14	0.18
Log income	4.11	4.22	4.60
Male salaried emp.	0.48	0.45	0.45
<i>Household head education:</i>			
Illiterate	0.18	0.16	0.14
Literate	0.10	0.09	0.08
Primary	0.11	0.13	0.11
Middle school	0.15	0.15	0.15
Secondary	0.26	0.27	0.29
Tertiary	0.18	0.17	0.21
Missing: woman is hh head	0.02	0.03	0.02
<i>Social group:</i>			
Hindu non-SCST	0.65	0.65	0.65
SCST	0.17	0.17	0.16
Muslim	0.14	0.14	0.15
Other	0.05	0.05	0.04
Age	36.13	36.68	36.86
Children 0-4	0.53	0.49	0.39
Children 5-14	1.37	1.20	1.04
N	33,462	30,463	28,247

TABLE 2.A.6: Bolivia: sample means

	2000	2005	2008	2011	2014
Labor force	0.67	0.63	0.68	0.66	0.64
<i>Own education:</i>					
Less than basic	0.25	0.27	0.20	0.18	0.17
Basic	0.19	0.18	0.16	0.15	0.15
Intermediate	0.18	0.17	0.18	0.16	0.16
Secondary completed	0.16	0.18	0.22	0.19	0.24
Tertiary	0.22	0.20	0.24	0.31	0.28
Log income	4.79	4.70	4.85	5.04	5.11
Male salaried emp.	0.53	0.57	0.54	0.56	0.52
<i>Household head education:</i>					
Less than basic	0.12	0.14	0.11	0.10	0.10
Basic	0.18	0.17	0.14	0.14	0.14
Intermediate	0.21	0.20	0.19	0.19	0.17
Secondary completed	0.17	0.18	0.22	0.20	0.21
Tertiary	0.26	0.22	0.25	0.30	0.27
Missing: woman is hh head	0.06	0.08	0.09	0.07	0.11
<i>Language:</i>					
Spanish only	0.47	0.47	0.52	0.53	0.54
Other	0.53	0.53	0.48	0.47	0.46
Age	37.74	38.05	37.69	37.67	37.78
Children 0-4	0.61	0.54	0.53	0.52	0.56
Children 5-14	1.45	1.30	1.23	1.21	1.12
N	1,517	1,245	1,151	3,057	3,771

TABLE 2.A.7: Indonesia: sample means

	2000	2004	2007	2014
Labor force	0.39	0.37	0.49	0.53
<i>Own education:</i>				
Less than primary	0.19	0.16	0.13	0.11
Primary	0.31	0.30	0.28	0.24
Junior high school	0.17	0.19	0.19	0.20
Secondary completed	0.26	0.28	0.29	0.32
Tertiary	0.07	0.08	0.12	0.14
Log income	3.95	4.35	4.30	4.64
Male salaried emp.	0.57	0.53	0.58	0.59
<i>Household head education:</i>				
Less than primary	0.16	0.14	0.13	0.12
Primary	0.27	0.25	0.24	0.22
Junior high school	0.16	0.17	0.17	0.17
Secondary completed	0.30	0.32	0.32	0.34
Tertiary	0.10	0.11	0.13	0.13
Missing: woman is hh head	0.01	0.01	0.01	0.02
Age	37.04	37.24	37.78	37.99
Children 0-4	0.43	0.43	0.45	0.45
Children 5-14	1.04	0.99	0.96	0.93
N	50,243	69,311	74,896	86,076

TABLE 2.A.8: Vietnam: sample means

	2002	2006	2010	2014
Labor force	0.86	0.86	0.85	0.87
<i>Own education:</i>				
Less than primary	0.13	0.11	0.08	0.09
Primary	0.22	0.22	0.22	0.18
Secondary	0.25	0.22	0.22	0.22
High school	0.31	0.35	0.29	0.27
Tertiary	0.10	0.09	0.19	0.24
Log income	3.97	4.39	4.72	4.96
Male salaried emp.	0.62	0.63	0.62	0.62
<i>Household head education:</i>				
Less than primary	0.14	0.13	0.10	0.10
Primary	0.17	0.15	0.14	0.15
Secondary	0.17	0.17	0.17	0.18
High school	0.20	0.26	0.25	0.25
Tertiary	0.10	0.09	0.15	0.15
Missing: woman is hh head	0.22	0.20	0.19	0.17
<i>Ethnicity:</i>				
Kinh	0.95	0.94	0.95	0.94
Other	0.05	0.06	0.05	0.06
Age	39.02	40.29	39.10	39.59
Children 0-4	0.29	0.30	0.39	0.40
Children 5-14	0.98	0.80	0.66	0.71
N	5,252	1,685	1,936	2,040

TABLE 2.A.9: Tanzania: sample means

	2000	2006	2014
Labor force	0.86	0.93	0.81
<i>Own education:</i>			
No schooling	0.12	0.09	0.07
Less than primary	0.13	0.07	0.06
Primary completed	0.62	0.68	0.64
Any secondary	0.12	0.15	0.17
Tertiary	0.00	0.01	0.05
Log income	3.05	3.72	3.44
Male salaried emp.	0.43	0.42	0.36
<i>Household head education:</i>			
No schooling	0.06	0.04	0.04
Less than primary	0.16	0.08	0.07
Primary completed	0.51	0.54	0.54
Any secondary	0.18	0.23	0.21
Tertiary	0.03	0.02	0.09
Missing: woman is hh head	0.06	0.09	0.06
Age	35.32	34.77	35.52
Children 0-4	0.95	0.73	0.76
Children 5-14	2.14	1.33	1.32
<i>N</i>	1,708	1,947	4,246

TABLE 2.A.10: South Africa: average marginal effects

Pr(Labor force)	1995	2001	2003	2010	2014
<i>Own education (Ref. = Less than primary)</i>					
Primary	0.038 (0.033)	0.021 (0.018)	0.035 (0.026)	0.050** (0.025)	0.037 (0.023)
Secondary not completed	0.119*** (0.015)	0.075*** (0.019)	0.048*** (0.011)	0.070*** (0.023)	0.102*** (0.025)
Secondary completed	0.265*** (0.018)	0.186*** (0.021)	0.194*** (0.018)	0.212*** (0.022)	0.213*** (0.034)
Tertiary	0.431*** (0.025)	0.333*** (0.023)	0.363*** (0.026)	0.388*** (0.019)	0.379*** (0.027)
Log income	-0.013*** (0.003)	-0.008*** (0.002)	-0.013*** (0.003)	-0.009*** (0.002)	-0.003 (0.003)
Male salaried emp.	0.035*** (0.011)	-0.030*** (0.009)	0.007 (0.009)	0.007 (0.009)	0.017 (0.011)
<i>Max adult married male education (Ref. = Less than primary)</i>					
Primary	0.010 (0.018)	0.035*** (0.009)	0.033 (0.023)	0.011 (0.018)	0.021 (0.029)
Secondary not completed	-0.023 (0.018)	0.028 (0.021)	0.007 (0.011)	0.004 (0.018)	0.030 (0.023)
Secondary completed	0.012 (0.020)	0.038 (0.024)	0.032 (0.023)	0.023* (0.014)	-0.004 (0.042)
Tertiary	-0.046 (0.030)	0.025 (0.018)	-0.025 (0.030)	-0.032** (0.016)	-0.025 (0.037)
Missing: no adult married male	0.097*** (0.024)	0.058** (0.024)	0.058** (0.024)	0.029 (0.027)	0.001 (0.035)
<i>Ethnicity (Ref. = Black)</i>					
Coloured	-0.025 (0.022)	-0.033*** (0.010)	-0.005 (0.016)	-0.018 (0.022)	-0.037* (0.021)
Indian/Asian	-0.220*** (0.014)	-0.219*** (0.008)	-0.216*** (0.020)	-0.138*** (0.026)	-0.225*** (0.031)
White	-0.121*** (0.012)	-0.121*** (0.017)	-0.120*** (0.016)	-0.040** (0.019)	-0.099*** (0.018)
Age	-0.002 (0.001)	-0.002** (0.001)	-0.002 (0.001)	-0.003*** (0.001)	-0.000 (0.001)
Children 0-4	-0.040*** (0.005)	-0.049*** (0.007)	-0.034*** (0.008)	-0.074*** (0.008)	-0.052*** (0.008)
Children 5-14	-0.019*** (0.004)	-0.014*** (0.005)	-0.009** (0.004)	-0.014*** (0.004)	-0.023*** (0.005)
Survey wave fixed effects	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes
N	7,601	11,361	10,658	20,713	17,890
Pseudo R-squared	0.081	0.076	0.076	0.083	0.067
FLFP-(sample)	0.557	0.679	0.663	0.654	0.676
FLFP-(survey weighted)	0.585	0.682	0.669	0.667	0.681

TABLE 2.A.11: Brazil: average marginal effects

Pr(Labor force)	2002	2005	2009	2013
<i>Own education (Ref. = Less than primary)</i>				
Elementary (1-4)	0.021** (0.010)	0.034** (0.013)	0.042*** (0.007)	0.065*** (0.015)
Elementary (5-8)	0.065*** (0.010)	0.080*** (0.012)	0.088*** (0.013)	0.118*** (0.016)
High school	0.174*** (0.015)	0.177*** (0.011)	0.155*** (0.015)	0.174*** (0.011)
Tertiary	0.365*** (0.013)	0.349*** (0.012)	0.314*** (0.011)	0.345*** (0.013)
Log income	-0.006*** (0.001)	-0.005** (0.002)	-0.006* (0.003)	0.002* (0.001)
Male salaried emp.	-0.016*** (0.006)	-0.022*** (0.006)	-0.011 (0.009)	-0.005 (0.005)
<i>Household head education (Ref. = Less than primary)</i>				
Elementary (1-4)	-0.003 (0.007)	-0.018** (0.008)	0.002 (0.011)	0.003 (0.012)
Elementary (4-8)	-0.011 (0.008)	-0.014** (0.007)	0.018 (0.014)	0.006 (0.013)
High school	-0.042*** (0.005)	-0.050*** (0.010)	-0.006 (0.021)	0.002 (0.016)
Tertiary	-0.101*** (0.014)	-0.119*** (0.011)	-0.041*** (0.013)	-0.052** (0.020)
Missing: woman is hh head	0.113*** (0.009)	0.072*** (0.008)	0.046*** (0.012)	0.016 (0.020)
<i>Ethnicity (Ref. = White)</i>				
Black	0.096*** (0.014)	0.063*** (0.012)	0.054*** (0.010)	0.051*** (0.011)
Mixed	0.041*** (0.013)	0.015** (0.006)	0.030*** (0.004)	0.012*** (0.004)
Asian	0.006 (0.033)	0.038 (0.023)	-0.068* (0.037)	-0.002 (0.037)
Indigenous	0.002 (0.047)	0.069 (0.074)	0.093* (0.050)	0.026 (0.047)
Age	-0.005*** (0.001)	-0.006*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)
Children 0-4	-0.080*** (0.008)	-0.085*** (0.005)	-0.082*** (0.005)	-0.093*** (0.006)
Children 5-14	-0.016*** (0.004)	-0.018*** (0.003)	-0.015*** (0.003)	-0.026*** (0.003)
State fixed effects	Yes	Yes	Yes	Yes
N	39,193	42,189	42,855	38,596
Pseudo R-squared	0.067	0.067	0.063	0.068
FLFP-(sample)	0.626	0.662	0.690	0.677
FLFP-(survey weighted)	0.621	0.657	0.686	0.675

TABLE 2.A.12: Jordan: average marginal effects

Pr(Labor force)	2006	2008	2010	2014
<i>Own education (Ref. = Less than primary)</i>				
Primary	-0.002 (0.006)	0.000 (0.004)	0.005 (0.008)	0.008 (0.009)
Preparatory	0.000 (0.006)	-0.019 (0.011)	-0.020 (0.016)	-0.009 (0.011)
Lower secondary	-0.009 (0.008)	-0.026*** (0.010)	-0.038** (0.017)	-0.018* (0.011)
Secondary	0.029*** (0.010)	0.009 (0.014)	0.011 (0.020)	0.014 (0.016)
Tertiary	0.269*** (0.017)	0.273*** (0.035)	0.257*** (0.042)	0.273*** (0.021)
Log income	-0.004** (0.002)	-0.003*** (0.001)	-0.002*** (0.001)	-0.006** (0.002)
Male salaried emp.	0.009 (0.006)	0.010 (0.012)	0.019*** (0.006)	0.012** (0.005)
<i>Household head education (Ref. = Less than primary)</i>				
Primary	0.004 (0.010)	-0.006 (0.006)	0.008* (0.004)	-0.014** (0.007)
Preparatory	0.003 (0.015)	0.013 (0.010)	0.024*** (0.007)	-0.008 (0.011)
Lower secondary	0.014 (0.012)	0.014 (0.018)	0.018* (0.010)	-0.008 (0.009)
Secondary	0.009 (0.012)	0.018* (0.010)	0.024 (0.017)	-0.000 (0.012)
Tertiary	0.028** (0.014)	0.041*** (0.010)	0.046*** (0.014)	0.013 (0.011)
Missing: woman is hh head	-0.000 (0.017)	0.038** (0.017)	0.051*** (0.017)	0.011 (0.022)
<i>Nationality (Ref. = Jordan)</i>				
Iraq	-0.060*** (0.018)	-0.094*** (0.004)	-0.125*** (0.003)	-0.122*** (0.002)
Syria	-0.030* (0.018)	-0.069*** (0.023)	-0.110*** (0.033)	-0.097*** (0.011)
Egypt	-0.067*** (0.015)	-0.039*** (0.010)	-0.059*** (0.020)	-0.092*** (0.025)
Other Arab countries	0.027* (0.014)	-0.043*** (0.011)	0.026 (0.036)	-0.030** (0.014)
Other non-Arab countries	0.688*** (0.058)	0.590*** (0.039)	0.624*** (0.020)	0.659*** (0.016)
Age	-0.002*** (0.000)	-0.003*** (0.000)	-0.003*** (0.001)	-0.002** (0.001)
Children 0-4	-0.022*** (0.006)	-0.020*** (0.003)	-0.027*** (0.007)	-0.017*** (0.001)
Children 5-14	-0.017*** (0.001)	-0.021*** (0.002)	-0.021*** (0.001)	-0.012*** (0.002)
Survey wave fixed effects	Yes	Yes	Yes	Yes
Governorate fixed effects	Yes	Yes	Yes	Yes
N	25,834	33,219	32,691	30,458
Pseudo R-squared	0.308	0.308	0.302	0.335
FLFP-(sample)	0.113	0.152	0.166	0.149
FLFP-(survey weighted)	0.117	0.158	0.171	0.146



TABLE 2.A.13: India: average marginal effects

Pr(Labor force)	1999	2004	2011
<i>Own education (Ref. = Illiterate)</i>			
Literate	-0.065*** (0.016)	-0.054*** (0.016)	-0.023** (0.012)
Primary	-0.081*** (0.015)	-0.067*** (0.021)	-0.031*** (0.009)
Middle school	-0.090*** (0.010)	-0.110*** (0.015)	-0.068*** (0.012)
Secondary	-0.025*** (0.009)	-0.054*** (0.013)	-0.048*** (0.013)
Tertiary	0.206*** (0.014)	0.131*** (0.027)	0.140*** (0.020)
Log income	-0.013*** (0.002)	-0.026*** (0.005)	-0.015*** (0.004)
Male salaried emp.	-0.038*** (0.011)	-0.022** (0.009)	-0.010 (0.012)
<i>Household head education (Ref. = Illiterate)</i>			
Literate	-0.048** (0.021)	-0.042*** (0.015)	0.016 (0.017)
Primary	-0.081*** (0.013)	-0.054*** (0.016)	-0.059** (0.026)
Middle school	-0.120*** (0.010)	-0.078*** (0.011)	-0.083*** (0.014)
Secondary	-0.155*** (0.009)	-0.125*** (0.011)	-0.104*** (0.011)
Tertiary	-0.161*** (0.017)	-0.117*** (0.017)	-0.133*** (0.020)
Missing: woman is hh head	0.037	-0.020	-0.030
<i>Social group (Ref.= Hindu non-SCST)</i>			
SCST	0.059*** (0.016)	0.037** (0.018)	0.034*** (0.012)
Muslim	-0.071*** (0.016)	-0.093*** (0.014)	-0.084*** (0.012)
Other	0.009 (0.015)	0.000 (0.011)	0.034 (0.021)
Age	0.003*** (0.001)	0.001** (0.001)	-0.000 (0.001)
Children 0-4	-0.017*** (0.004)	-0.024*** (0.005)	-0.039*** (0.009)
Children 5-14	-0.000 (0.003)	0.009*** (0.003)	0.007* (0.004)
State fixed effects	Yes	Yes	Yes
N	33,462	30,463	28,247
Pseudo R-squared	0.124	0.123	0.092
FLFP-(sample)	0.179	0.216	0.175
FLFP-(survey weighted)	0.184	0.208	0.182

TABLE 2.A.14: Bolivia: average marginal effects

Pr(Labor force)	2000	2005	2008	2011	2014
<i>Own education (Ref. = Less than basic)</i>					
Basic	0.073 (0.091)	0.014 (0.073)	-0.040 (0.045)	0.008 (0.027)	0.016 (0.023)
Intermediate	0.101 (0.063)	0.037 (0.040)	0.019 (0.037)	0.040 (0.030)	0.054* (0.030)
Secondary completed	0.117 (0.081)	0.076 (0.056)	-0.019 (0.024)	-0.010 (0.023)	0.015 (0.016)
Tertiary	0.177** (0.085)	0.216*** (0.073)	0.149*** (0.054)	0.178*** (0.041)	0.229*** (0.025)
Log income	-0.048*** (0.011)	-0.039*** (0.009)	-0.044*** (0.013)	-0.038*** (0.013)	-0.025*** (0.003)
Male salaried emp.	-0.044* (0.024)	-0.100*** (0.031)	-0.064 (0.043)	-0.082*** (0.016)	-0.040*** (0.015)
<i>Household head education (Ref. = Less than basic)</i>					
Basic	-0.059 (0.081)	0.020 (0.072)	-0.001 (0.026)	-0.062** (0.028)	-0.003 (0.019)
Intermediate	-0.002 (0.076)	0.016 (0.083)	-0.047 (0.033)	-0.083** (0.036)	-0.058 (0.048)
Secondary completed	-0.030 (0.064)	0.012 (0.043)	-0.082 (0.052)	-0.081*** (0.031)	-0.014 (0.019)
Tertiary	-0.004 (0.123)	0.005 (0.082)	-0.104** (0.050)	-0.065** (0.031)	-0.060** (0.027)
Missing: woman is hh head	0.006 (0.100)	-0.015 (0.063)	-0.093 (0.077)	-0.073 (0.052)	0.052 (0.033)
<i>Language (Ref. = Spanish only)</i>					
Other	0.116*** (0.028)	0.060 (0.040)	0.044** (0.020)	0.080*** (0.014)	0.063* (0.034)
Age	-0.000 (0.003)	0.007*** (0.001)	0.003 (0.002)	0.004** (0.002)	0.005*** (0.001)
Children 0-4	-0.059*** (0.018)	-0.037*** (0.009)	-0.068*** (0.024)	-0.085*** (0.010)	-0.069*** (0.009)
Children 5-14	-0.014 (0.023)	-0.011 (0.010)	0.013 (0.010)	-0.014*** (0.005)	-0.004 (0.011)
Department fixed effects	Yes	Yes	Yes	Yes	Yes
N	1,517	1,245	1,151	3,057	3,771
Pseudo R-squared	0.075	0.086	0.084	0.075	0.073
FLFP-(sample)	0.655	0.652	0.667	0.661	0.654
FLFP-(survey weighted)	0.667	0.629	0.676	0.659	0.645

TABLE 2.A.15: Indonesia: average marginal effects

Pr(Labor force)	2000	2004	2007	2014
<i>Own education (Ref. = Less than primary)</i>				
Primary	-0.016 (0.015)	-0.005 (0.007)	-0.019 (0.012)	-0.006 (0.009)
Junior high school	-0.049** (0.024)	-0.016** (0.008)	-0.031** (0.012)	-0.008 (0.008)
Secondary completed	0.065** (0.030)	0.062*** (0.013)	0.016 (0.014)	0.014 (0.013)
Tertiary	0.360*** (0.043)	0.416*** (0.024)	0.287*** (0.015)	0.299*** (0.014)
Log income	-0.033*** (0.004)	-0.077*** (0.002)	-0.074*** (0.005)	-0.067*** (0.005)
Male salaried emp.	-0.027*** (0.007)	-0.030*** (0.008)	-0.053*** (0.005)	-0.041*** (0.006)
<i>Household head education (Ref. = Less than primary)</i>				
Primary	-0.034*** (0.009)	-0.015*** (0.006)	-0.016*** (0.006)	-0.014*** (0.003)
Junior high school	-0.081*** (0.014)	-0.043*** (0.005)	-0.022*** (0.007)	-0.009 (0.009)
Secondary completed	-0.081*** (0.011)	-0.064*** (0.007)	-0.041*** (0.006)	-0.017 (0.011)
Tertiary	-0.094*** (0.019)	-0.045*** (0.012)	-0.022** (0.011)	-0.028* (0.015)
Missing: woman is hh head	0.134*** (0.031)	-0.015 (0.049)	-0.090*** (0.022)	-0.098*** (0.017)
Age	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)
Children 0-4	-0.065*** (0.009)	-0.080*** (0.010)	-0.082*** (0.005)	-0.097*** (0.007)
Children 5-14	-0.019*** (0.005)	-0.030*** (0.005)	-0.036*** (0.005)	-0.027*** (0.005)
Province fixed effects	Yes	Yes	Yes	Yes
N	50,243	69,311	74,896	86,076
Pseudo R-squared	0.086	0.109	0.081	0.076
FLFP-(sample)	0.399	0.374	0.489	0.568
FLFP-(survey weighted)	0.386	0.368	0.489	0.529

TABLE 2.A.16: Vietnam: average marginal effects

Pr(Labor force)	2002	2006	2010	2014
<i>Own education (Ref. = Less than primary)</i>				
Primary	0.033 (0.023)	0.095*** (0.028)	-0.017 (0.037)	0.021 (0.028)
Secondary	0.056 (0.035)	0.016 (0.035)	-0.016 (0.040)	-0.029 (0.030)
High school	0.102*** (0.033)	0.074** (0.030)	0.070* (0.036)	0.044 (0.039)
Tertiary	0.152*** (0.031)	0.167*** (0.038)	0.142*** (0.041)	0.141*** (0.033)
Log income	-0.014** (0.006)	0.014* (0.009)	-0.010 (0.006)	-0.005 (0.009)
Male salaried emp.	0.009 (0.012)	-0.002 (0.017)	-0.026** (0.013)	-0.001 (0.016)
<i>Household head education (Ref. = Less than primary)</i>				
Primary	-0.002 (0.015)	0.065** (0.026)	-0.018 (0.031)	-0.004 (0.026)
Secondary	0.020 (0.014)	0.055*** (0.021)	0.039 (0.044)	0.007 (0.030)
High school	-0.018 (0.016)	0.023 (0.025)	-0.023 (0.052)	-0.010 (0.035)
Tertiary	-0.073** (0.029)	-0.003 (0.069)	-0.020 (0.031)	-0.061 (0.043)
Missing: woman is hh head	-0.024 (0.020)	0.071** (0.028)	-0.018 (0.026)	0.003 (0.038)
<i>Ethnicity (Ref. = Kinh)</i>				
Other	-0.027 (0.036)	0.003 (0.032)	-0.016 (0.023)	-0.049 (0.057)
Age	-0.001 (0.001)	-0.003* (0.002)	-0.005** (0.002)	-0.004*** (0.001)
Children 0-4	-0.010 (0.016)	-0.041* (0.022)	-0.054*** (0.016)	-0.062*** (0.020)
Children 5-14	0.005 (0.006)	0.009 (0.011)	0.011 (0.011)	-0.010 (0.012)
Province fixed effects	Yes	Yes	Yes	Yes
N	4,938	1,264	1,751	1,737
Pseudo R-squared	0.117	0.123	0.143	0.126
FLFP-(sample)	0.864	0.839	0.860	0.868
FLFP-(survey weighted)	0.850	0.829	0.838	0.857

TABLE 2.A.17: Tanzania: average marginal effects

Pr(Labor force)	2000	2006	2014
<i>Own education (Ref. = No schooling)</i>			
Less than primary	0.042 (0.043)	0.030 (0.022)	-0.000 (0.042)
Primary completed	0.062** (0.025)	0.027* (0.016)	0.045 (0.028)
Any secondary	0.110*** (0.017)	0.023 (0.023)	0.093*** (0.029)
Tertiary	0.129*** (0.021)	-0.133 (0.133)	0.150** (0.065)
Log income	-0.057*** (0.009)	-0.009*** (0.003)	-0.019*** (0.007)
Male salaried emp.	0.025 (0.019)	0.002 (0.013)	-0.004 (0.014)
<i>Household head education (Ref. = No schooling)</i>			
Less than primary	0.038 (0.052)	-0.027 (0.019)	-0.017 (0.046)
Primary completed	0.079 (0.055)	-0.023 (0.023)	-0.052 (0.039)
Any secondary	0.079 (0.050)	-0.028 (0.029)	-0.004 (0.045)
Tertiary	0.114*** (0.043)	0.030 (0.049)	0.000 (0.054)
Missing: woman is hh head	-0.046 (0.084)	0.028 (0.028)	0.084*** (0.029)
Age	0.003* (0.001)	-0.000 (0.001)	0.006*** (0.001)
Children 0-4	-0.013 (0.008)	-0.001 (0.009)	-0.027** (0.011)
Children 5-14	-0.002 (0.005)	-0.001 (0.003)	0.004 (0.007)
Region fixed effects	Yes	Yes	Yes
Survey wave fixed effects	Yes	Yes	Yes
N	1,708	1,947	4,246
Pseudo R-squared	0.150	0.104	0.170
FLFP-(sample)	0.798	0.923	0.747
FLFP-(survey weighted)	0.860	0.926	0.812

**TABLE 2.A.18:** Trends in sample inclusion criteria over time

South Africa	1995	2001	2003	2010	2014
<i>Women (25-54): share currently married:</i>					
Urban	0.66	0.56	0.56	0.54	0.51
Rural	0.64	0.56	0.55	0.45	0.41
<i>Women (25-54): share in urban areas:</i>					
Married	0.55	0.63	0.62	0.71	0.73
Not married	0.53	0.63	0.61	0.63	0.64
Brazil	2002 <sup>†</sup>	2005	2009	2013	
<i>Women (25-54): share currently married:</i>					
Urban	0.67	0.66	0.65	0.64	
Rural	0.82	0.81	0.81	0.80	
<i>Women (25-54): share in urban areas:</i>					
Married	0.85	0.83	0.84	0.85	
Not married	0.93	0.92	0.92	0.93	
Bolivia	2000	2005	2008	2011	2014
<i>Women (25-54): share currently married:</i>					
Urban	0.74	0.74	0.71	0.71	0.70
Rural	0.84	0.83	0.81	0.84	0.83
<i>Women (25-54): share in urban areas:</i>					
Married	0.66	0.67	0.67	0.68	0.70
Not married	0.78	0.78	0.79	0.82	0.83
India	1999	2004	2011		
<i>Women (25-54): share currently married:</i>					
Urban	0.89	0.88	0.88		
Rural	0.91	0.91	0.92		
<i>Women (25-54): share in urban areas:</i>					
Married	0.26	0.26	0.29		
Not married	0.29	0.31	0.40		
Indonesia	2000	2004	2007	2014	
<i>Women (25-54): share currently married:</i>					
Urban	0.81	0.83	0.83	0.85	
Rural	0.87	0.88	0.88	0.89	
<i>Women (25-54): share in urban areas:</i>					
Married	0.42	0.43	0.44	0.50	
Not married	0.53	0.53	0.54	0.59	
Vietnam	2002	2006	2010	2014	
<i>Women (25-54): share currently married:</i>					
Urban	0.79	0.78	0.80	0.82	
Rural	0.85	0.85	0.86	0.87	
<i>Women (25-54): share in urban areas:</i>					
Married	0.25	0.27	0.31	0.32	
Not married	0.34	0.38	0.41	0.41	
Tanzania	2000	2006	2014		
<i>Women (25-54): share currently married:</i>					
Urban	0.67	0.67	0.67		
Rural	0.81	0.78	0.76		
<i>Women (25-54): share in urban areas:</i>					
Married	0.20	0.26	0.37		
Not married	0.34	0.38	0.49		
Jordan	2006	2008	2010	2014	
<i>Women (25-54): share currently married:</i>					
Urban and rural	0.77	0.77	0.79	0.80	

Sources: See Table 2.A.1. Notes: <sup>†</sup>Before 2004, the Brazilian PNAD does not include rural municipalities of the North Region.

TABLE 2.A.19: India and Jordan: decomposition of FLFP

	India		Jordan	
Pr(Labor force) last year	0.182	<i>N</i> = 28,247	0.146	<i>N</i> = 30,458
Pr(Labor force) first year	0.184	<i>N</i> = 33,462	0.117	<i>N</i> = 25,834
Difference	-0.002		0.029	
	At 1999 coeff.	At 2011 coeff.	At 2006 coeff.	At 2014 coeff.
Own education	0.010*** (0.002)	0.004* (0.002)	0.012*** (0.001)	0.013*** (0.001)
Log income	-0.007*** (0.001)	-0.008*** (0.002)	0.001** (0.000)	0.000 (0.000)
Male salaried emp.	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Household head education	-0.012*** (0.002)	-0.009*** (0.001)	0.001 (0.001)	-0.000 (0.001)
Social group	-0.001* (0.001)	-0.001 (0.001)	0.005*** (0.001)	0.003*** (0.000)
Age	0.001* (0.001)	-0.001 (0.001)	-0.003*** (0.001)	-0.004*** (0.001)
Children	0.002** (0.001)	0.002 (0.002)	0.009*** (0.001)	0.006*** (0.001)
Region dummies	0.002*** (0.001)	0.002** (0.001)	0.002*** (0.000)	0.003*** (0.001)
Survey wave dummies			-0.006* (0.003)	0.001 (0.004)
Total covariates	-0.005	-0.011	0.020	0.023
Coefficients & unobservables	0.003	0.009	0.009	0.006
<i>N</i>	61,709	61,709	56,292	56,292

TABLE 2.A.20: Brazil and South Africa: decomposition of FLFP

	Brazil		South Africa	
Pr(Labor force) last year	0.675	<i>N</i> = 38,596	0.681	<i>N</i> = 17,890
Pr(Labor force) first year	0.621	<i>N</i> = 39,193	0.682	<i>N</i> = 11,361
Difference	0.054		-0.002	
	At 2002 coeff.	At 2013 coeff.	At 2001 coeff.	At 2014 coeff.
Own education	0.053*** (0.002)	0.045*** (0.002)	0.029*** (0.002)	0.034*** (0.003)
Log income	-0.002*** (0.001)	0.001 (0.001)	-0.003*** (0.001)	-0.001 (0.001)
Male salaried emp.	-0.001*** (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
Household head education	0.012*** (0.001)	0.001 (0.003)	0.003 (0.002)	-0.003 (0.004)
Social group	0.005*** (0.001)	0.002*** (0.000)	0.009*** (0.001)	0.009*** (0.001)
Age	-0.006*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)	-0.001 (0.001)
Children	0.009*** (0.001)	0.013*** (0.001)	0.007*** (0.001)	0.009*** (0.001)
Region dummies	0.000* (0.000)	0.001*** (0.000)	-0.000 (0.000)	0.005*** (0.001)
Survey wave dummies			0.021*** (0.004)	-0.002 (0.004)
Total covariates	0.072	0.057	0.061	0.051
Coefficients & unobservables	-0.018	-0.003	-0.062	-0.053
<i>N</i>	77,789	77,789	29,251	29,251

TABLE 2.A.21: Indonesia and Bolivia: decomposition of FLFP

	Indonesia		Bolivia	
Pr(Labor force) last year	0.529	<i>N</i> = 50,243	0.645	<i>N</i> = 3,771
Pr(Labor force) first year	0.386	<i>N</i> = 86,076	0.667	<i>N</i> = 1,517
Difference	0.142		-0.022	
	At 2000 coeff.	At 2014 coeff.	At 2000 coeff.	At 2014 coeff.
Own education	0.028*** (0.004)	0.020*** (0.001)	0.014** (0.007)	0.012*** (0.002)
Log income	-0.022*** (0.002)	-0.045*** (0.004)	-0.015*** (0.004)	-0.008*** (0.001)
Male salaried emp.	-0.001*** (0.000)	-0.001*** (0.000)	0.001 (0.001)	0.001* (0.000)
Household head education	-0.005*** (0.001)	-0.001 (0.001)	0.002 (0.003)	0.004** (0.002)
Social group			-0.007*** (0.002)	-0.003* (0.002)
Age	0.001 (0.001)	0.002*** (0.001)	-0.006*** (0.002)	-0.004*** (0.001)
Children	0.001 (0.000)	0.001* (0.001)	0.008 (0.008)	0.006 (0.004)
Region dummies	-0.007*** (0.000)	-0.006*** (0.000)	0.003*** (0.001)	-0.005*** (0.001)
Total covariates	-0.005	-0.030	-0.000	0.002
Coefficients & unobservables	0.147	0.172	-0.022	-0.024
<i>N</i>	136,319	136,319	5,288	5,288

TABLE 2.A.22: Tanzania and Vietnam: decomposition of FLFP

	Tanzania		Vietnam	
Pr(Labor force) last year	0.812		0.874	<i>N</i> = 2,040
Pr(Labor force) first year	0.868		0.858	<i>N</i> = 5,252
Difference	-0.056		0.016	
	At 2000 coeff.	At 2014 coeff.	At 2002 coeff.	At 2014 coeff.
Own education	0.012*** (0.003)	0.013*** (0.003)	0.017*** (0.003)	0.019*** (0.000)
Log income	-0.032*** (0.008)	-0.010** (0.004)	-0.013** (0.006)	-0.005*** (0.000)
Male salaried emp.	-0.002 (0.002)	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)
Household head education	0.006** (0.003)	0.002 (0.003)	-0.003* (0.002)	-0.003*** (0.000)
Social group			-0.000 (0.000)	-0.000*** (0.000)
Age	0.000 (0.001)	0.001 (0.001)	-0.006*** (0.002)	-0.010*** (0.000)
Children	0.004 (0.004)	0.002 (0.007)	-0.002 (0.002)	-0.004*** (0.000)
Region dummies	-0.005** (0.003)	0.022*** (0.002)	0.005*** (0.001)	0.004*** (0.000)
Survey wave dummies	-0.002** (0.001)	0.002*** (0.001)		
Total covariates	-0.020	0.033	-0.003	-0.001
Coefficients & unobservables	-0.037	-0.089	0.019	0.017
<i>N</i>	6,060	6,060	7,292	7,292



**TABLE 2.A.23:** South Africa: decomposition of FLFP, 1995–2014

Pr(Labor force) 2014	0.681	<i>N</i> = 17890
Pr(Labor force) 1995	0.585	<i>N</i> = 7601
Difference	0.096	
	At 1995 coefficients	At 2014 coefficients
Own education	0.047*** (0.004)	0.040*** (0.003)
Log income	0.003*** (0.001)	0.001 (0.001)
Male salaried emp.	-0.004*** (0.001)	-0.002 (0.001)
Max adult married male education	0.002 (0.002)	-0.003 (0.003)
Race	0.017*** (0.001)	0.015*** (0.002)
Age	-0.005*** (0.002)	-0.000 (0.001)
Children	0.011*** (0.001)	0.015*** (0.001)
Province dummies	0.000 (0.000)	0.004*** (0.000)
Survey wave dummies		-0.002 (0.004)
Total covariates	0.071	0.068
Coefficients & unobservables	0.025	0.028
<i>N</i>	25,491	25,491

TABLE 2.A.24: Educational attainment and social group: common variables for all countries

Country	Education			Social group		
	Less than primary	Primary completed	Secondary completed	Tertiary	Reference (low FLFP)	High FLFP
Bolivia	< Basic	Basic, Intermediate	Secondary	Any tertiary	(Language) Only Spanish	Spanish and indigenous
Brazil	< Primary	Elementary (1-4), Elementary (5-8)	High school	Any tertiary	(Race) White, Asian	Black, Mixed, Indigenous
India	Illiterate, Literate	Primary, Middle school	Secondary	Any tertiary	(Caste, religion) Hindus non-SCST, Muslim	Hindus SCST, Other
Indonesia	< Primary	Primary, Junior high school	Senior high school	Any tertiary		
Jordan	< Primary	Primary, Preparatory, Lower Secondary	Secondary	Any tertiary	(Nationality) Iraq, Syria, Egypt, Other Arab	Jordan, Other non-Arab
South Africa	< Primary	Primary, Secondary not completed	Secondary	Any tertiary	(Race) White, Indian / Asian	Black, Colored
Tanzania	Never attended, Primary not completed	Primary, Ordinary secondary (< Form IV)	Ordinary secondary (Form IV), Advanced secondary (Forms V-VI)	Any tertiary		
Vietnam	< Primary	Primary, Lower secondary	High school	Any tertiary	(Ethnicity) Other	Kinh



TABLE 2.A.26: Decomposition of FLFP between countries: last year

[illegible]

### 3 The Roots of Female Emancipation: the Initializing Role of Cool Water\*

**Abstract:** Reviewing the literature on the deep historic roots of gender inequality, we theorize and provide evidence for a trajectory that (1) originates in a climatic configuration called the “Cool Water” (CW-) condition, leading to (2) late female marriages in preindustrial times, which eventually pave the way towards (3) various gender-egalitarian outcomes today. The CW-condition is a specific climatic configuration that combines periodically frosty winters with mildly warm summers under the ubiquitous accessibility of fresh water. It embodies opportunity endowments that significantly reduce fertility pressures. The resulting household formation patterns empowered women and reduced gender inequality.

---

\*This chapter is co-authored with Amy C. Alexander, Stephan Klasen, and Christian Welzel. We would like to thank Laura Barros, Oded Galor, Lisa Höckel, Lennart Kaplan, Susan Steiner, Holger Strulik, Mikołaj Szoltysek, participants at the Gender Governance Link conference in Goettingen, the Quality of Government Institute’s Internal Conference in Budapest, and seminars at the Universities of Goettingen and Freiburg for valuable comments, and Florian Klassen and Alexander Stöcker for research assistance. We gratefully acknowledge financial support from Lower Saxony’s Ministry for Science and Culture for the project *The Gender-Governance Link: Gender Equality and Public Goods Provision* under the initiative “Geschlecht–Macht–Wissen: Genderforschung in Niedersachsen”. Alexander was also supported through the European Union Seventh Framework Research Project PERDEM (Performance of Democracy, grant agreement 339571). The views expressed here are ours and do not represent those of the funders.

### 3.1 Introduction

There is a growing consensus in the literature that gender equality in its various manifestations goes together with a host of beneficial outcomes, from economic productivity to distributional justice, physical security, generalized trust, honest government, effective democracy and other quality of life aspects (e.g., World Bank, 2001, 2011; Fish, 2002; Inglehart and Norris, 2003; Coleman, 2004; Klasen, 2002; Klasen and Lamanna, 2009; Duflo, 2012; Branisa et al., 2013; Alexander and Welzel, 2015). Another strand of literature highlights the importance of female agency during the preindustrial Malthusian era for the timing and pace of the fertility transition and consequent projection into the modern era of sustained economic growth (e.g., Galor and Weil, 1996; Lagerlöf, 2003; Diebolt and Perrin, 2013; Prettnner and Strulik, 2017). These beneficial effects of gender equality have sparked interest in the deep historic roots of women's emancipation from male domination (e.g., Doepke and Tertilt, 2009; Alesina et al., 2013; Hansen et al., 2015; Jayachandran, 2015).

In the footprints of Ester Boserup, scholars locate the roots of gender inequality in societies' agrarian history. While one group of researchers sees the transition to agriculture in and by itself as the origin of greater inequality (Hansen et al., 2015), others are more specific and identify the plow-using type of agriculture as the source (Alesina et al., 2013; Giuliano, 2015). Still others argue that scarcity of arable land strengthens agriculture's contribution to gender inequality (Hazarika et al., 2015).

Building on this scholarship, we theorize and provide evidence for an overlooked trajectory that (1) originates in a climatic configuration called the "Cool Water" (CW-) condition, from where the path leads to (2) late female marriages in preindustrial times, which eventually pave the way towards (3) various gender egalitarian outcomes today.

The CW-condition is a specific climatic configuration that combines periodically frosty winters with mildly warm summers under the ubiquitous accessibility of fresh water. It is most prevalent in Northwestern Europe and some of its former colonial offshoots (Welzel, 2013, 2014). We argue that, over time, household formation patterns of an agrarian economy evolve, in part, as a response to the challenges and opportunities ingrained in its climate. The CW condition embodies opportunity endowments that significantly reduced fertility pressures on women, which favored late female marriages already in the preindustrial era. In CW areas, the relatively favorable disease environment implied that lower fertility was needed to achieve the desired number of children that survive into adulthood. In addition, the ubiquitous availability of water, as a fundamental agricultural input, reduced mobility costs of farming households and defied centralized

control of land, which favored nuclear over extensive families and neo-locality over patri-locality. The resulting family and household patterns placed women into a better position to struggle for more gender equality during the subsequent transitions toward the industrial and post-industrial stages of development. Hence, enduring territorial differences in the CW-condition (henceforth CW) predict differences in preindustrial female marriage ages, which in turn predict differences in gender equality today.<sup>1</sup>

In support of this causal chain, we use the CW index: a combination of absolute latitudes, continuous rain, and mild summers (Welzel, 2013, 2014). We present cross-country evidence showing that CW has a sizable and consistent impact on current female marriage ages as well as the age gap between the spouses: a one standard deviation increase in the CW index is associated with a 10 month reduction on the male-female difference in mean ages at first marriage. Furthermore, we can trace this effect back to the preindustrial epoch using European historical data on marital ages from Dennison and Ogilvie (2014). Both these effects are robust to including the other ‘deep drivers’ discussed in the literature. In line with these results, we show in reduced-form estimates that CW has a large and significant impact on today’s gender gaps.

In the absence of experimental control, the two lurking threats to the validity of any causal interpretation are always reverse causality and omitted variable bias. Reverse causality is of no concern in our setting. The reason is the distinct temporal ordering of our variables along a far-reaching sequence of separate historic epochs, extending from (a) *original* environmental conditions manifest in CW to (b) *preindustrial* marriage patterns to (c) *post-industrial* gender-egalitarian outcomes.

With respect to omitted variable bias, we first note that the CW condition is exogenous to any of the variables we consider on the left-hand side of our causal chain. But there is always the risk that unobserved country characteristics might account for the relationship between CW and its hypothesized outcomes.

In order to address this problem, we go at length in double-checking our results against many possible confounding factors. First, we conduct multiple sensitivity tests to the composition of the sample and find that our conclusions are not driven by certain world regions or subregions. Second, we control for a battery of additional variables proposed in the literature and find that the CW effect persists throughout. Third, following Oster (2017), we estimate how large the effect of unobservables needs to be to explain away the

---

<sup>1</sup>For the link between family systems and gender equality in the preindustrial era see, for example, Hajnal (1965, 1982); Hartman (2004); De Moor and Van Zanden (2010); Szoltysek et al. (2017). For evidence on persistent effects of historic family systems on contemporary socio-economic outcomes see Reher (1998); Giuliano (2007); Duranton et al. (2009); Ebenstein (2014); Dilli et al. (2015); Dilli (2016); Rijpma and Carmichael (2016).

CW effect. For reasonable upper bounds on the full model's  $R^2$ , the degree of selection on unobservables would have to be large relative to selection on the observable variables of our most restricted models. Thus, it seems unlikely that omitted variable bias alone could account for our results.

Another concern relates to the fact that countries did not exist in today's borders invariantly throughout the temporal scope of our analyses. To address this issue, we "ancestry-adjust" the countries' CW-scores using Putterman and Weil's (2010) post-1500 World Migration Matrix. Doing so changes the unit of analysis from today's countries to ancestral populations. Since this exercise fully reproduces all of our major results, it is safe to conclude that they are not an artifact of using countries in today's borders as the unit of analysis.

The remainder of the chapter proceeds in the following steps. Section 3.2 reviews the literature and derives from this discussion our theoretical propositions. Section 3.3 introduces the data and variables used to demonstrate the empirical validity of our propositions. The fourth section describes the empirical strategy while section 3.5 presents the findings. Finally, we conclude with a discussion of our findings' theoretical implications, in section 3.6.

## 3.2 Theoretical discussion

In section 3.2.1, we start by reviewing existing hypotheses that link gender equality to specific features of a society's agrarian past. We then describe in detail how historical household formation patterns varied across settings, highlighting the distinction between societies where women marry at relatively later ages vis-à-vis those where they marry immediately after puberty (section 3.2.2). In late marriage societies, women are relatively more empowered, as argued in section 3.2.3. Finally, in section 3.2.4, we build our main hypothesis that the norm of late marriages is, in part, determined by temperature and water patterns, through their impact on fertility pressures on women.

### 3.2.1 Original sources of gender (in)equality

Three recent studies suggest alternative origins as the decisive historic drivers of gender inequality today. To begin with, Alesina et al. (2013) revive Boserup's (1970) thesis that the participation of women in preindustrial agriculture differed significantly between plow-using and plow-free cultivation systems. The plow constituted a gender-biased technology as it required more upper body strength than alternative tools, such as



the digging stick or the hoe. As a result, in societies that adopted the plow, women reallocated their time away from farming towards domestic activities. This labor division along gender lines became gradually encultured into enduring norms. Alesina et al. show that the fraction of a country's population whose ancestors practiced preindustrial plow agriculture is negatively correlated with contemporary female participation in the labor force, politics, and corporate ownership. Moreover, among children of immigrants in the US and Europe, those with plow-using ancestors hold less egalitarian beliefs about the appropriate role of women.

In contrast, Hansen et al. (2015) argue that the transition from humans' original foraging lifestyle to sedentary agriculture *as such* is a driver of preindustrial disparities in gender roles, no matter what particular cultivation methods have been used. As these authors suggest, the earlier the transition to sedentary agriculture happened, the more intense the cultivation methods became thereafter. Intense methods of cultivation generate a demand for cheap mass labor, which in turn increases the fertility pressures on women. As a consequence, women reallocate their time from fieldwork to raising children and other indoor activities related to caretaking. Societal beliefs about gender roles then incrementally evolved in support of this labor division. Accordingly, societies with longer histories of agriculture had more time to enculture patriarchal values in their moral systems. Indeed, Hansen et al. show that longer histories of agriculture are negatively correlated with female participation in the labor force, politics, education, as well as the sex ratio at birth. The correlation remains significant even after controlling for ancestral plow use, which retains its significance. Not only is this relationship present in a cross-country sample but also for a sample of European regions, and for a sample of second-generation immigrant children in the US.

The third study by Hazarika et al. (2015) argues that historic resource scarcity shaped cultures of gender discrimination. This claim is consistent with patterns of sex-inequality and resource availability in some non-human primate species, gender gaps in prehistoric human skeleton sizes, and contemporary evidence on the relationship between material deprivation and gender bias in intra-household resource allocation. According to this thesis, prehistorical differences in resource scarcity gave rise to a persistent culture of gender discrimination. The authors measure historic resource scarcity by limitations of arable land and show that these limitations are negatively correlated with present-day measures of gender equality across countries and population sex-ratios across districts in India.

While these insights are valuable, we offer a fourth explanation that we believe adds an important element to understanding the deep drivers of gender inequality and also

provides a plausible transmission channel.

### 3.2.2 Historical household formation patterns

The feature in preindustrial household types receiving most attention refers to a bundle of elements that Hajnal (1965, 1982) branded as the “Western family pattern” and described as unique to Northwestern Europe (Todd, 1985, 1987; Hartman, 2004). Based on archival records from the seventeenth and eighteenth centuries, Hajnal (1982) stresses four exceptional features of household formation in Northwestern Europe, which supposedly date back at least to Medieval times<sup>2</sup>: (a) late ages at first marriage for women, resulting in smaller age gaps between husband and wife; (b) a considerable proportion of women (and men) never married, which, together with late marriage, implied for women that they lived under lower pressure to maximally exploit their reproductive potential; (c) neo-local residence: couples move into their own household upon marriage—a feature that reduces obligations to the extended family and, thus, fosters individual autonomy; and (d) a widespread practice among both adolescent men and women to work as contracted servants in non-kin households until marriage.

In sharp contrast, household formation patterns in the then advanced areas of the Middle East, China, and India—which have survived with remarkably small changes until the twentieth century—involve the exact opposite features: (a) much earlier marital ages for women, often resulting in large age gaps between husband and wife; (b) near universal marriage, indicating an intolerance of women just living by themselves (including quick remarriage of widows); (c) patri-local residence: freshly married couples move into the household of the husband’s parents under their headship; and (d) an almost exclusive use of family labor for household production.

We, of course, acknowledge that substantial variation existed around these two stylized household systems (e.g., Szołtysek, 2014). But for reasons of clarity, it is helpful to take this simplification as a point of departure.

From the gender perspective, the most fundamental element of the Northwest European household system is late ages at first marriage for women (Hajnal, 1965, 1982; Smith, 1981; Hartman, 2004). In Northwestern Europe, women would typically marry in their twenties, while in the early marriage systems of the Middle East, India, and China,

<sup>2</sup>See Hartman (2004, chapter 3) for a review of the historical evidence. For example, in his classical study of the Germanic barbarian tribes, *De Germania*, from around 98 AD, Tacitus already describes a Germanic late marriage pattern, in contrast to the early marriage tradition of the Roman Empire: “The [Germanic] young men marry late and their vigor is thereby unimpaired. The girls, too, are not hurried into marriage. As old and full-grown as the men, they match their mates in age and strength, and the children reproduce the might of their parents” (quoted in Herlihy, 1985, p. 73).

brides would marry in their early teens, immediately after puberty. In late marriage societies, women's ages at first marriage were elastic with respect to local economic conditions (Smith, 1981; Carmichael et al., 2016; Cinnirella et al., 2017) and, indeed, for certain societies, at certain points in time, ages at first marriage were very high even by contemporary standards: for example, the singulate mean age at first marriage (henceforth: SMAM) from Norway's 1875 census was 27.3 years for women and 28.5 for men (Ruggles, 2009, Appendix Table). Similarly, it was on average 26 for women and 28 for men in rural Germany in 1740–1860 (Klasen, 1998). As a consequence of late marital ages for women, the age gap between the spouses was substantially lower than in early marriage societies, where the husband would often be ten to fifteen years older than his wife. For example, in Bangladesh, the SMAM in 1974 was 16.4 years for women and 24 for men (UN 2009).

Another important aspect of late marriage societies is the high proportion of never married individuals of both sexes. At times, the never married constituted more than ten percent of the adult population in preindustrial Northwestern Europe, whereas they were rarely more than five percent of the adult population in early marriage societies (Hajnal, 1965, 1982).

In late medieval Northwestern Europe, late marriages were intimately linked with the institution of service through which a large share of young boys and girls would leave the parental household in their early- to mid-teens and circulate as servants on a contractual basis, typically for several years until marriage. The practice of contracted, non-kin service means that households form on the basis of consent instead of lineage. The fact that marriage itself has been an act of agreement among adult non-relatives, instead of being pre-arranged among people belonging to the same family circle, further detached household formation from lineage and strengthened its contractual character. Apparently, the contractual household system reaches as far back in time as mid-fourteenth century in England (Poos, 1991) and even the ninth century around Paris (Herlihy, 1985), if not earlier. Importantly, service did not mean *domestic* service, in fact “[m]ost servants were not primarily engaged in domestic tasks, but were part of the workforce of their master's farm or craft enterprise” (Hajnal, 1982, p. 473). Contractual household service defied a vertical allocation of labor across social strata: “It was not only the poor and landless whose children went into service. Those who operated their own farms and even farmers with large holdings sent their children into service elsewhere, sometimes replacing them with hired servants in their own household” and “[t]here was no assumption that a servant, as a result of being in service, would necessarily be socially inferior to his or her master” (Hajnal, 1982, pp. 471 and 473). Thus defined,

the institution of service fulfilled two roles in late marriage societies: (i) it provided a subsistence base for unmarried young men and women as well as a means to accumulate sufficient savings to then start a family, and (ii) it created a flexible workforce for ageing farm tenants in need of support.

The same logic clarifies why contracted household service was not a widespread phenomenon in early marriage societies: whenever girls married at puberty, there was no chance for service to be part of their premarital life-cycle, and the availability of family labor, including children, in multi-generational households meant that local labor markets were peripheral.

### 3.2.3 Implications for gender equality

The Northwest European system of late marriage for women and life-cycle service for the young had profound implications for gender relations. While Hajnal already outlined the most important of these implications, Hartman (2004) extends the analysis in both breadth and depth.

First, marrying later reduces women's childrearing burden during the most productive years of their life-cycles. In addition, women marrying in their mid-twenties are more mature, experienced, and arguably more confident than women marrying at puberty. Combined with a smaller age gap to their spouses, women in late marriage societies accordingly enjoy more intra-household bargaining power than their early marriage counterparts. Moreover, in late marriage societies, the intervals between generations are inevitably larger than in early marriage societies. Therefore, while it was common for newly married couples to head their own household, it was typical of earlier marriage societies for the wife to join the husband's multi-generational household. As a result, in early marriage household systems, there are often more individuals positioned above the young wife: not only her parents-in-law but also other members of the husband's kin, like his siblings, uncles, and cousins. In such a setting, a freshly married woman has little bargaining power. To sum up, the absolute age at first marriage, the age gap between the spouses, and the amount of household members who are hierarchically superior to the wife are all important determinants of adult women's status in the household.

Second, in late marriage societies, individual consent is a necessary condition for marriage, whereas in early marriage societies parental consent is decisive. The mere fact of marrying later meant that a considerable proportion of the parents of freshly married couples were no longer alive at the time of the marriage. For instance, "as in sixteenth-century Lyon, one-third of teenagers becoming apprentices and *one-half of the*

*young women marrying for the first time* were fatherless; [and] as in seventeenth-century Bordeaux, more than one-third of the apprentices had neither parent alive” (Davis, 1977, p. 87; emphasis added). In contrast, in earlier marriage systems, young women were more likely to have living parents at the time of their marriage and living in-laws throughout a substantial portion of their marital years. Unsurprisingly, marriage norms in these systems were tilted towards parental arrangement rather than individual choice (e.g., Edlund and Lagerlöf, 2006). It is plausible to assume that individual choice lowers the intra-household bargaining costs for both spouses; either because they match with partners that have similar preferences, or because mutual consent brings about a higher degree of altruism to individual preferences.

Third, the institution of contractual farm service is a driver of gender equality norms and behaviors in its own right. As shown in Figure 3.1, contractual service was a common phase in the life-cycle of *both* men and women across Northwestern Europe.<sup>3</sup> Indeed, it is striking how similar the patterns of service were among the life-cycles of both sexes: the prevalence rates are not too different in absolute terms (although, in general, they were higher for men across all age groups) and their age-distributions are remarkably similar. Over the life-cycle, service started roughly at 15–19 years of age for both sexes—an age range at which most women in early marriage societies were already married. Service peaked at 20–24 years of age and then declined as people left service to get married. As a result, the life-cycle of young men and women in late marriage societies is much more similar than in early marriage societies, where young brides are separated early on from their male peers by the experiences of marriage and motherhood, often living in seclusion. The convergence of life experiences between the sexes in Northwestern Europe was thus a force challenging the boundaries between gender identities. By contrast, in the Middle East, China, or India, the divergence of life experiences early on kept gender identities and domains of activity strictly separate (Hartman, 2004).

In Northwestern Europe, the years of service were an ideal opportunity for young people to accumulate savings, skills, and meet potential marriage partners (Hajnal, 1982; De Moor and Van Zanden, 2010). For the young servants, high mobility (De Moor and Van Zanden, 2010) strengthened their maturity and weakened patriarchal norms, to the benefit of women (Reher, 1998; Hartman, 2004). As summarized by Hajnal (1982, pp. 474–475), “[w]hile in service, women were not under the control of any male relative.

---

<sup>3</sup>A similar institution existed in Japan (Cornell, 1987). Among all non-Western agrarian civilizations, Japan is the one whose geo-climatic configuration is by far the closest to Northwestern Europe’s (as defined by the Cool water index, see section 3.3).

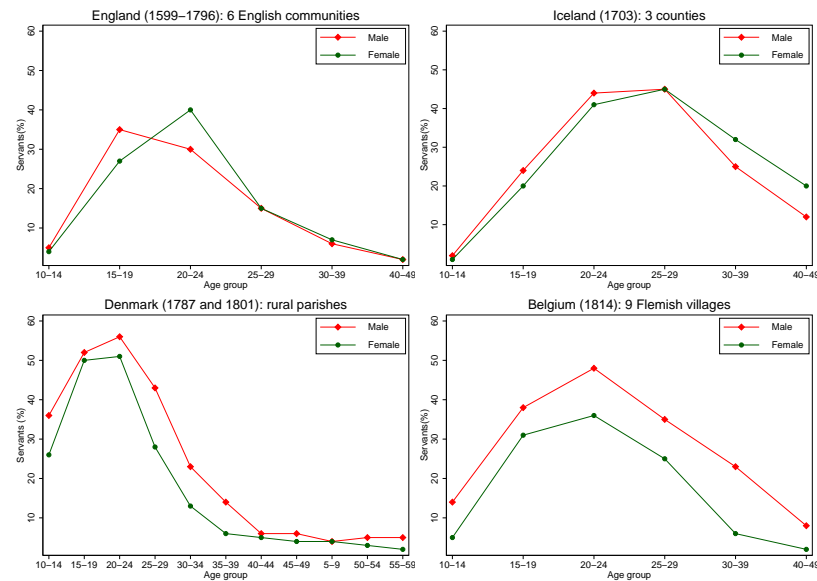


FIG. 3.1: Service prevalence by age and sex for selected Northwest European countries

Source: Hajnal (1982): Tables 13, 14, and 17.

They made independent decisions about where to live and work and for which employer. There was also financial independence even though women servants' wages were lower than men's."<sup>4</sup>

Remarkably, the implications of gender gaps at first marriage were well understood early on, as recorded in legal codes and councils from early medieval Europe (Herlihy, 1985). For example, a late eight century council (796–97) held at Fréjus, in today's Southeastern France, "explicitly states that the groom and bride should be 'not of dissimilar age but of the same age.' Many abuses take place, it warns, when the groom is adult and the bride immature, or *vice versa*. [...] The council strongly advocated that marriage unite mature partners of equal age" (Herlihy, 1985, pp. 75–76). Thus, the implications derived in this section are not anachronistic.

To avoid misunderstandings, we are by no means claiming that late female marriage ages created anything close to a perfect balance in gender value. For example, studies have found gender gaps in survival opportunities in late marriage societies (e.g., Klasen, 1998), even though the magnitude of the gaps tended to be smaller than in current-day South Asia or China (Klasen, 2003). Instead, we suggest that late marriage societies

<sup>4</sup>A recent study on eighteenth-century Amsterdam shows that wages earned by female servants were substantial: "most servant women could save between one-third and half of the amount of money an unskilled man could save in the same period of time" (Boter, 2017, p. 71).

generated opportunities for female individualization that were not available for their early marriage counterparts.<sup>5</sup> While these opportunities did not eradicate women's discrimination, they built the basis for emancipatory struggles against it. The consequences of these struggles are evident in the global variation in gender discrimination across countries today. Indeed, women of late marriage societies were in a better position to turn external events to their own advantage. The experience of service, for instance, could be turned into labor force participation during war periods (Goldin, 1991; Acemoglu et al., 2004). Centuries of high fertility elasticity with respect to environmental factors facilitated the adoption of modern family planning technologies, such as the contraceptive pill (Goldin and Katz, 2002). Moreover, a legacy of consensual marriages, joint decision making, pooling of resources at marriage, and higher female bargaining power meant that women could reap the benefits of new consumer technologies that greatly reduced the burden of household chores (Greenwood et al., 2005; Cavalcanti and Tavares, 2008; Coen-Pirani et al., 2010). In early marriage societies, these opportunities passed by without major changes in the position of women.

### 3.2.4 Origins of household formation patterns

Diamond (1997) reinvigorates the notion that agriculture is forcefully shaped by an environment's natural endowments. Specifically, what kind of crops can be grown and what type of cattle can be bred depends directly on climatic conditions, especially a climate's thermo-hydrological configuration, that is, temperature and water patterns. Consequently, the ways in which farming households form in an agrarian economy should be a response to the challenges and opportunities embodied in this economy's climatic conditions. Should this be an accurate premise, the peculiarity of the Northwest European household pattern should be a response to the peculiarity of its climatic conditions.

This raises the question of what is so particular about the climate in Northwestern Europe. And if there is something particular about it, why does it incentivize a late marriage type of household formation?

In trying to answer this question, Welzel (2013, 2014) summarizes the particular thermo-hydrological constellation of Northwestern Europe's climate as the "Cool Water"

---

<sup>5</sup>Centuries-old household formation systems shape women's self-perceived agency and expectations about, say, partner selection and marital consent. For example, in the Indian Human Development Survey of 2005, 55 percent of married women who felt they played a role in the selection of their husbands first met them close to or on the wedding day (Banerji et al., 2008); a situation hard to conceive of wherever mutual consent, instead of parental consent, is the prevailing norm.

(CW-) condition. The CW-condition combines periodically (albeit not permanently) frosty winters with mildly warm summers under the ubiquitous and permanent accessibility of fresh water sources. This condition is indeed most prevalent in Northwestern Europe and fades gradually away as one moves to Eastern and Southern Europe. Outside Europe, the CW-condition only prevails in the former settler colonies of Northwestern Europe (i.e., parts of North America and Australia and New Zealand) and—at a less pronounced level—in Japan, the Korean peninsula, and adjacent territories in East Asia, as well as the isolated Southern tips of South America and Africa. According to Welzel, the significance of the CW-condition originates in the fact that it bestows on people some very basic existential autonomies that are absent under other conditions. These autonomies incentivize a late marriage type of agrarian household formation. Two types of existential autonomy are particularly noteworthy: *reproductive autonomy* and *water autonomy*.<sup>6</sup>

*Reproductive autonomy* is the degree to which people are exempted from the pressure to maximize fertility. A crucial determinant of fertility pressure in agrarian societies is the infant mortality rate incurred naturally by an environment's pathogen load; the infectiousness of its water sources; the scarcity of fresh water, dairy products, and other foods; and heat stress. Welzel claims that the thermo-hydrological features of the CW-condition reduce these risk factors. As a result, women in CW-areas could afford lower fertility levels to achieve the desired number of children that survive into adulthood.<sup>7</sup> Moreover, the type of combined cereal-livestock farming made possible by the CW-condition requires lower labor inputs than, for instance, irrigation-managed agriculture and rice or tropical crop cultivation. Lower demands for labor inputs, including child labor, further reduce fertility pressures. Hence, women enjoyed more reproductive autonomy with respect to the timing of first marriage, the interval between pregnancies, and a higher recognition of their worth beyond the birth and care of children.

Reproductive autonomy partly originates in climatically induced disease security. The prevalence of non-host pathogens—from Malaria to yellow fever, dengue fever and other tropical diseases—increases with average temperatures, a fact known as the latitudinal gradient (Guernier et al., 2004; Cashdan, 2014). Before 1750, infant mortality was mainly

<sup>6</sup>See Welzel (2013, 2014) for broader implications of the CW-condition for long-run development. Here, the discussion focuses exclusively on household formation patterns.

<sup>7</sup>We speculate that this is the mechanism implicit in Alfred Marshall's observation, back in 1890, that: "The age of marriage varies with the climate. In warm climates where childbearing begin early, it ends early, in colder climates it begins later and ends later; but in every case the longer marriages are postponed beyond the age that is natural to the country, the smaller is the birth-rate; the age of the wife being of course much more important in this respect than that of the husband" (Marshall, 2013, p. 150).



determined by infectious diseases (e.g., Deaton, 2013, pp. 81–87).<sup>8</sup> In CW-areas, both cold temperatures and abundance of fresh water lower the pathogen load, also because fresh water is safer to drink under cold temperatures (Cashdan, 2014). The resulting lower fertility requirement in CW-areas meant that women's time was less constrained by childrearing: young women could therefore postpone marriage in response to labor market incentives (Smith, 1981; Cinnirella et al., 2017).<sup>9</sup>

*Water autonomy* is the ability to access fresh water freely and permanently either for consumption, fishing, agriculture, transportation, or water power. Water autonomy is higher in colder, rainy areas. The availability of water was a major constraint for farming households in traditional agrarian societies. Areas suitable for large-scale irrigation agriculture developed more autocratic institutions and labor-repressive regimes compared to rain-fed agricultural societies (Bentzen et al., 2017). The coordination of large irrigation infrastructure projects and the allocation of scarce water resources among farmers were best achieved by a centralized authority that could enforce water-sharing rules and collect taxes to finance infrastructure investment. The monopoly of scarce water resources then became an instrument through which a small landowning elite confiscated rents from a mass of peasant households. This system of agricultural production was typical of Middle Eastern, Andean and Mexican societies, as well as India and China (Wittfogel, 1957). In contrast, rain-fed agricultural production was highly decentralized. The level of rents extracted by large feudal landowners in Northwest Europe never reached the amount extracted by their irrigation-areas counterparts (e.g., Jones, 1981; Powelson, 1994; Mitterauer, 2010). Moreover, the ubiquitous availability of fresh water makes most available land arable, thus turning valuable land into a ubiquitous asset that defies

---

<sup>8</sup>Of course, Northwest European countries were early adopters of modern public health innovations, such as vaccinations, from the second half of the eighteenth century onwards. Ager et al. (2018), for example, show how the introduction of the smallpox vaccine in Sweden in 1801 reduced infant mortality and, consequently, gross fertility. In this article, we are referring to long-run preindustrial differentials in infant mortality *before* such innovations became available. For example, in 1750–1800, before the introduction of smallpox vaccination, Sweden's infant mortality rate was fluctuating around a stable mean of 211 deaths out of 1000-live births (Ager et al., 2018, Figure 3, Panel A). In 1960, the corresponding figure was 210 in Egypt, and 220 in Nepal; in 1965, it was 225 in Afghanistan, and 263 in Yemen (World Bank, 2016). If we assume that these latter countries experienced some improvement in infant mortality rates in the two hundred year period 1750s–1960s, they must have exhibited historical infant mortality rates that were substantially higher than those of Sweden. The lack of available cross-country data on preindustrial infant mortality rates, particularly for non-European countries, prevents us from testing this mechanism in an econometric framework.

<sup>9</sup>Some authors have proposed a direct link between high pathogen prevalence and collectivist cultures, as societies develop strong distrust of strangers and in-group bias as an evolutionary strategy against infectious diseases (Fincher et al., 2008; Murray and Schaller, 2010). This theory could also explain why, in area with high pathogen prevalence, preferences for extended households based on kinship were relatively stronger.

centralized control. Entry barriers and fixed costs of farming are lower when there is no need to maintain expensive irrigation (Haber, 2012). Finally, the lush pastures typical of CW-areas lend themselves to a combined form of cereal and livestock farming that widens nutritional options and easily feeds a small family while keeping the demand for child labor at the low end. Together with a lower infant mortality, the weak demand for child labor further reduces the fertility pressures on women.

Thus, areas with cool temperatures and rainfall throughout the seasons enhance *water autonomy* and *reproductive autonomy*, thus proliferating the natural basis of egalitarian individualism and its expression in the late-marriage system. Over time, these marriage patterns became embedded in both formal and informal institutions, thereby persisting until today. We note that, in theory, the CW condition could still have a direct impact on present day marriage ages. We believe however that, if present, this direct effect is at most marginal. The mechanisms of *reproductive* and *water* autonomy lost relevance in most modern countries either because family farming is now a residual activity, or because medical innovations and public health measures have weakened the link between geography and infant mortality. Therefore, we posit that any persistence of the CW effect on contemporary marriage patterns runs through the long-term effect of CW on historical marriage patterns, rather than through a contemporary direct effect.

Voigtländer and Voth (2013) highlight another important mechanism. They show for early modern England that a higher share of land devoted to animal husbandry, relative to plow agriculture, increased the ages at first marriage for women and the proportion of young women in service. The reason is that, given their comparative disadvantage in plow agriculture, women have higher employment prospects in pastoral farming. According to Voigtländer and Voth, another push towards the late marriage pattern resulted from the Black Death, which wiped out at least one third of the European population in the fourteen century. Consequently, the land-to-labor ratio increased, which made animal husbandry more attractive because it is a land-intensive and labor-saving sector compared to plow agriculture. Yet, a shift to animal husbandry was an option only in regions where there was sufficient rainfall to sustain year-round grazing fields. In contrast, “[a]gricultural conditions in Mediterranean countries did not favor the pastoral farming of the type common in Northwest Europe. In particular, low rainfall made it impossible to keep large herds of cattle and sheep in the same area year-round” (Voigtländer and Voth, 2013, p. 2250).<sup>10</sup> Thus, if the Black Death was indeed

<sup>10</sup>In regions where animal husbandry depends on great distance traveling, this activity becomes incompatible with childrearing and women lose their comparative advantage.

the event responsible for a further push towards the late marriage pattern<sup>11</sup>, its impact was conditional on local geo-climatic characteristics embodied in the CW-condition.

The timing of the transition to sedentary agriculture is a potential confounding factor of our CW-thesis. Especially in Europe, agriculture spread from the Fertile Crescent towards the European peripheries until it reached Northwestern Europe around 5,000 BCE. Clearly, this was a transition from a warm and arid core towards increasingly colder and wetter regions. Thus, in Europe, the effects we attribute to the CW-condition could be very hard to distinguish from those of being a later adopter of agriculture. However, existing evidence contradicts this idea. In a detailed study, Olsson and Paik (2016) explore how the spread of agriculture throughout European sub-national regions affects contemporary individualistic values, measured from World Value Survey data. They show that, in the period 9,000-5,500 BCE, the diffusion of agriculture across Europe from Southwest Asia is correlated with lower preferences for obedience; i.e., regions with shorter agricultural histories value obedience relatively less. However, this relationship completely disappears after 5,500 BCE, just as agriculture is about to reach Northwestern Europe (Olsson and Paik, 2016, Table 5). Over the two millennia during which agriculture spread throughout Northwestern Europe, there is no correlation between the timing of agricultural adoption and societal preferences for obedience.

Our theory explains this apparent puzzle. It is the CW-condition, not the years of agriculture, that explains societal preferences for obedience. The correlation between obedient orientations and the length of the agrarian legacy exists only as long as the CW-areas remain non-agrarian. In the moment they turn agrarian, the correlation vanishes and it does not matter anymore when the transition took place. What mattered was that, once adopted, agricultural societies in those regions enjoyed higher levels of reproductive and water autonomy than their Southern and Eastern European neighbors. As a result, over the centuries, CW-areas preserved the lower valuation of obedience typical of pre-agrarian societies, independently of the length of their agricultural histories.<sup>12</sup>

From the theoretical discussion above follow four propositions, which the remainder of this article will investigate:

- (1) A society's CW-condition correlates positively with its historical age at first marriage for women. Because of data limitations, we demonstrate this proposition mostly for European countries and a selection of non-European countries, repre-

<sup>11</sup>See Dennison and Ogilvie (2014, p. 673) for a rebuttal of this claim.

<sup>12</sup>A further piece of evidence consistent with our interpretation is that the effect of years of agriculture on preferences for obedience stops being statistically significant as soon as latitude is included as a control variable (Olsson and Paik, 2016, Table 3a).

senting major civilizations in pre-colonial history.

- (2) Because centuries-old household formation patterns become encultured in a society's enduring norms, a society's CW-condition correlates positively with contemporary ages at first marriage for women. We demonstrate this proposition for all countries in the world with available data.
- (3) Historically as well as contemporary late ages of marriage for women are positive determinants of various gender egalitarian outcomes today, mediating the initial impact of the CW-condition.
- (4) The CW-condition's emancipatory impulse is not absorbed by other deep historic drivers championed in the literature, but is a significant driver even when controlling for these other effects.

### 3.3 Data and descriptives

We briefly describe the variables used in the empirical section and present selected descriptive statistics. A full list of all variables used, their sources and summary statistics can be found in Table 3.A.1.

As main contemporary measures of gender equality, we use the average female to male ratio in 1990–2010 of: (i) labor force participation rates for the age group 25–59 from the ILO Laborsta EAPPEP 6th Revision (2011), (ii) life expectancy at birth from the World Bank's World Development Indicators, and (iii) mean years of education of the 25+ years old from Barro and Lee (2013). In addition, we consider the share of firms with some degree of female ownership, the share of parliament seats held by women, and the UNDP's Gender Development Index.

Our contemporary data for ages at first marriage is from the United Nations' *World Marriage Data* (UN 2009), which covers the period 1960–2008. We use the singulate mean age at first marriage (SMAM) for men and women. For each country, we select data from its first available year, which ranges from 1960 to 2006 (see Table 3.1) with the period 1960–1980 accounting for 80 percent of the observations.<sup>13</sup>

We take historical data on European female ages at first marriage from Dennison and Ogilvie (2014). They collect marital ages between 1500 and 1900 from 365 studies

<sup>13</sup>We make only two adjustments. First, for Malta, we do not use data from its first available year, 1967, because it includes nationals living outside the country; instead we use data for the next available year, 1985. Second, for Jamaica, we do not use the first available year, 1970, because it is an extreme outlier from trend, as shown in Figure 3.A.1; instead we use data for the next available year, 1982.

**TABLE 3.1:** Descriptive statistics for selected variables

	Mean	(Std. Dev.)	Min.	Max.	N
<i>Average female-male ratio in 1990–2010:</i>					
Labor force participation	0.68	(0.21)	0.15	1.01	191
Life expectancy	1.07	(0.04)	0.99	1.21	202
Years of education	0.82	(0.22)	0.21	1.41	146
<i>Ages at first marriage:</i>					
<i>Contemporary</i>					
Female	21.96	(2.88)	15.56	32.19	214
Year of obs.	1975.97	(9.26)	1960	2006	214
Male	26.13	(2.28)	21.13	34.49	209
Female/male	0.84	(0.07)	0.64	0.98	209
Male–female	4.12	(1.76)	0.5	9.93	209
<i>Historical (Europe only, ref. = England)</i>					
Female	-2.07	(3.1)	-6.81	2.36	28
<i>Deep determinants:</i>					
Cool water	0.48	(0.15)	0.21	0.83	183
Years of agriculture	4.31	(2.42)	0	10	165
Plow	0.48	(0.48)	0	1	227
Agricultural suitability	0.54	(0.33)	0	0.98	214

Sources: See accompanying text.

on historical demography and harmonize the data by regressing the marital ages on several characteristics of the sources from which they were extracted.<sup>14</sup> We use the country-specific coefficients from that multivariate regression, where England is the omitted country (Dennison and Ogilvie, 2014, Table 2). The fact that female ages at first marriage were rising, on average, throughout the whole 1500–1900 period is not a major issue since Dennison and Ogilvie’s regression also controls for the historical time period covered by each of the demographic studies, thereby purging the country-specific estimates from any trend effects. Denmark’s 2.36 value, for example, “shows that its female age at first marriage was 2.36 years higher than England’s, controlling for time period, unit of observation, settlement size, publication type, and sources and methods used” (Dennison and Ogilvie, 2014, p. 663). Table 3.1 shows that, on average, the 28 European countries for which data are available have female ages at first marriage around 2 years below that of England. Yet, there is substantial variation: Belarusian brides were almost 7 years younger (at first marriage) than English brides, whereas Danish brides were approximately 2 years older. The female age at first marriage for England, the reference country, is 25.26 years.

For non-European historical data, we take Gapminder’s female SMAM for the period 1800–1900. Of the 27 countries with available data, 11 are non-European.<sup>15</sup> We interpret

<sup>14</sup>Carmichael et al. (2016, p. 200) discuss limitations of these data.

<sup>15</sup>Armenia, Azerbaijan, Bangladesh, China, Egypt, Georgia, India, Japan, Kazakhstan, Pakistan, Sri

these data with caution, since they are collected from a variety of sources and, in some cases, supplemented by *ad hoc* qualitative adjustments or backward extrapolations by Gapminder.<sup>16</sup>

Considering the deep historic determinants of gender equality, we use the three variables most prominently discussed in the literature. To begin with, *years of agriculture* is the number of thousands of years from 1500 C.E. since the Neolithic revolution, as provided by Putterman and Trainor (2006). *Plow usage* and *agricultural suitability* are taken from Alesina et al. (2013), with plow usage indicating the proportion of a country's population with ancestors that used the plow in preindustrial agriculture. Agricultural suitability measures the suitability of ancestral land for the cultivation of barley, wheat, sorghum, rye, foxtail millet, or pearl millet. Summary statistics for these variables are shown in Table 3.1.

### The Cool Water index

To measure the CW-condition, we rely on geo-climatic data from Parker (2000) as well as Gallup et al. (2010). These data have no specific time frame. Thus, apart from minor short-term fluctuations—such as the little ice age and interglacial warming—the CW-condition captures territorial differences in thermo-hydrological conditions that have been relatively constant over the past 11,000 years when the last ice age ended.

The CW-condition constitutes a specific thermo-hydrological configuration, namely, the combination of cold winters with mild summers under the ubiquitous availability of fresh water. This condition is prevalent in territories of high latitude in coastal proximity.

To capture the CW-condition's thermal aspect, we take latitude as the starting point, using each country-centroid's latitude in degrees.<sup>17</sup> Higher latitudes get us away from the tropics and into areas with lower seasonal temperatures on average.

However, not all high latitude areas comprise the CW-condition's moderate seasonality, which combines winter cold with mostly mild summer heat. Thus, we need to further qualify latitude for this additional thermal condition. To do so, we take the usual peak temperature (in degrees Celsius) in a country-territory's hottest month of

---

Lanka, and United States.

<sup>16</sup>For more details, see the original documentation of these data at <https://www.gapminder.org/wp-content/uploads/2008/10/gapdoc009.pdf>. For the 10 European countries for which both Gapminder and Dennison and Ogilvie (2014) report data, the correlation coefficient is 0.95. Non-European data are of much lower quality. They are only used for an additional analysis; results are posted in the Appendix.

<sup>17</sup>The highest latitudes are 90 degrees at the poles. No country-centroid comes even close to that. In the Northern hemisphere, the highest latitude is obtained by Norway at about 67 degrees. In the Southern hemisphere it is New Zealand at about 42 degrees. We equate Norway's 67 degrees with 1 and standardize all other country-centroids' latitudinal degrees to this maximum.

the year, which is July or August in the Northern hemisphere and February or March in the South. Interestingly, while average annual temperature correlates strongly and negatively with latitude<sup>18</sup>, summer heat is uncorrelated with latitude.<sup>19</sup> For instance, summer heat peaks are as high in Mongolia as they are in Somalia. This pattern reflects the fact that mild summers mostly prevail in high latitudes but not all high latitudes belong into this category because they are divided into continental climates (with high summer peaks) and maritime climates (with low summer peaks). Thus, latitude is an ideal representative of cold winters, yet not of mild summers. To measure indeed the mildness of summers, we calculate the inverse of peak summer heat.

For the very same reason, high latitudes include most of the countries that possess the hydrological features of the CW-condition, and yet this is only a subset of the countries in high latitudes. Therefore, it is necessary to capture the hydrological features of the CW-condition by an additional, independent measure: continuous rain.

To capture continuous rainfall, average annual precipitation per month would be a misleading yardstick. The reason is that average annual precipitation correlates strongly with tropical climates and negatively with latitude<sup>20</sup> and is, thus, untypical for CW-regions. Most tropics have a monsoon season in which the extreme amount of rain is excessive and water, while abundant, is wasted and harmful (e.g., floods). Misleadingly, these extremes inflate the measure of average annual precipitation per month. Thus, the necessary qualification needed to capture the CW-condition's typical precipitation pattern is a focus, instead, on whether the rainfall in a region's driest month is high. To capture this feature, we use the typical rainfall level (in cubic millimeters) in a country's driest month.<sup>21</sup> To correct a distribution skewed to the top in this measure, we calculate its square root. Doing so moves extreme outliers at the high end of this measure (i.e., Pacific islands) closer to the center of the distribution. We call this measure henceforth *continuous rain*.

The next question is how to combine (1) latitudinal height, (2) summer mildness, and (3) continuous rain. These measures should be combined in a way that best represents in a single indicator the thermo-hydrological configuration typical of the CW-condition. Instead of imposing a theoretical solution on this problem, we subject the three measures to an exploratory factor analysis. As it turns out, the three measures reflect two independent dimensions. Continuous rain and mild summers represent one dimension,

<sup>18</sup>The country-centroids' latitudes correlate with the countries' annual mean temperatures at  $r = -0.89$  ( $N=183$ ;  $p$ -value = 0.001, 2-tailed).

<sup>19</sup> $r = 0.07$  ( $N=183$ ;  $p$ -value = 0.336, 2-tailed).

<sup>20</sup>Average annual rainfall per month and latitude correlate at  $r = -0.53$  ( $N=177$ ;  $p$ -value = 0.00, 2-tailed).

<sup>21</sup>Continuous rain in this definition is literally uncorrelated with average annual rainfall per month.

with factor loadings of 0.80 (mild summers) and 0.77 (continuous rain). High latitudes, by contrast, represent a separate dimension, on which only this measure shows a major loading (i.e., a factor loading of 0.95).<sup>22</sup>

In terms of substance, the first dimension of this factor solution represents *maritime climates*, which capture the *water*-component of the CW-condition. The second dimension, by contrast, represents what is unique to high latitudes irrespective of maritime climates, which is *cold winters*. In other words, the second dimension captures the *coolness*-component of the CW-condition.

These results show that CW is a condition that combines two independent components into a single configuration. Accordingly, the measurement of CW should represent this pattern and proceed as an additive combination of its two independent components. Following this premise, we calculate for each country its factor score on the first dimension (i.e., the water factor) and on the second dimension (i.e., the coolness factor).<sup>23</sup> The latter represents the coldness of winters independent of maritime climates, for which reason country scores on the first dimension (i.e., the water factor) and the second one (i.e., the coolness factor) are uncorrelated. Hence, we can calculate CW as an additive combination of two independent components by taking the arithmetic mean of the water and the coolness factor.<sup>24</sup>

Should the effects of CW's coolness component and its water component indeed add on each other, this additive combination will capture both effects in a single measure. In other words, if it is really the combination of coolness and water that makes the difference, this additive measure will isolate that effect.

In addition, however, as a validity check, we point to evidence that the CW-condition, as measured here, correlates with a similar measure of climatic configuration. In the Koeppen-Geiger classification of climate zones (Peel et al., 2007), the CW-condition correlates strongly with a country's share of land area in what is called the cold-to-temperate zones that lack a dry season.<sup>25</sup>

We also acknowledge that a key concern with our CW-index relates to differences

<sup>22</sup>The factor analysis has been conducted across 183 countries for which all three measures are available. The analysis has been conducted under the "Kaiser-criterion," advising the extraction of as many factors as there are with Eigenvalues above 1. The factor loadings we report are obtained after a "varimax-rotation." The factor solution explains 75 percent of the variance.

<sup>23</sup>We compute regression-based factor scores, i.e., weighted averages of factor loadings, the raw variables of interest, and the inverse of their covariance matrix (Thompson, 1951).

<sup>24</sup>We do this after having normalized the two factor-z-scores for each country into a range from 0 for the lowest scoring observation and 1 for the highest scoring one.

<sup>25</sup>A country's CW-condition correlates with its territorial share in the temperate-rainy climate zone (called "cf") at  $r = 0.66$  and with its territorial share in the cold-rainy zone (called "df") at  $r = 0.49$  ( $N=156$ ;  $p$ -value = 0.00, 2-tailed for both correlations).



in country area size. Indeed, scores on the CW-index might not be comparable across countries with different area sizes when bigger size implies higher within-country variability in the CW-condition. In the Online Appendix, we deal with this issue.<sup>26</sup> Territorial country size is entirely unrelated to within-country variability in the CW-condition. Moreover, only 14 percent of the total variation in the CW-condition is within countries, while 86 percent is between countries. Hence, country-mean differences in the CW-condition are significant and meaningful as they capture by far most of the existing territorial variation in CW-conditions.<sup>27</sup>

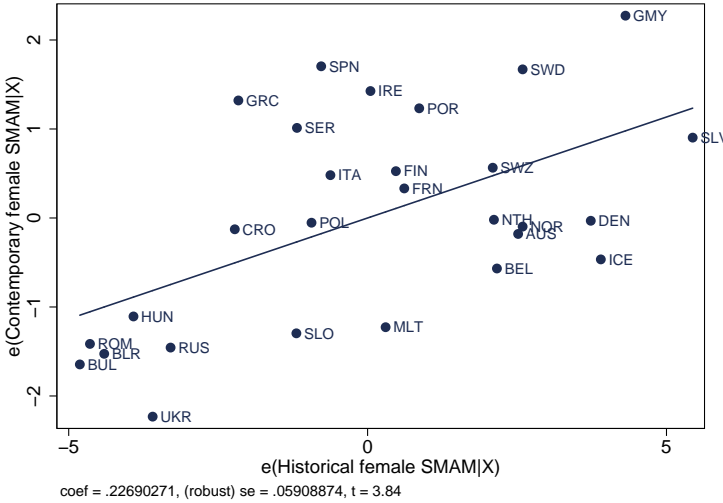
To ensure comparability to the previous literature we take the same set of baseline historical controls used by Alesina et al. (2013): (1) the presence of large domesticated animals, (2) the number of levels in political hierarchies, and (3) the level of economic complexity proxied by the type of settlement patterns (e.g., nomadic vs. complex settlements). The only exception is that we do not include the proportion of ancestral land that is tropical or subtropical in our baseline specification for reasons discussed below. Nevertheless, we include this variable in our set of additional controls for robustness checks. As contemporary controls, we follow the literature in using the natural log of per capita income and its square.<sup>28</sup>

For Europe, we find that ages at first marriage for women have persisted over centuries. Figure 3.2 displays the positive correlation between historical and contemporary marital ages for women. Countries with higher marital ages for women in 1500–1900 have older brides at first marriage in the postwar period. Moreover, as hypothesized, there is a negative correlation between the husband-wife age gap and the CW-condition (Figure 3.3). The relationship looks fairly linear; furthermore, it is not just a “European artifact” but holds for different continents and world regions. To test more rigorously if these descriptive correlations are in fact meaningful, we move to a multivariate regression framework which is outlined in the next section.

<sup>26</sup>A previous version of this index (Welzel, 2014) also included the coastline share of a country’s borders as a proxy for temperate maritime climates. We discuss this change and compare the two versions in the Online Appendix.

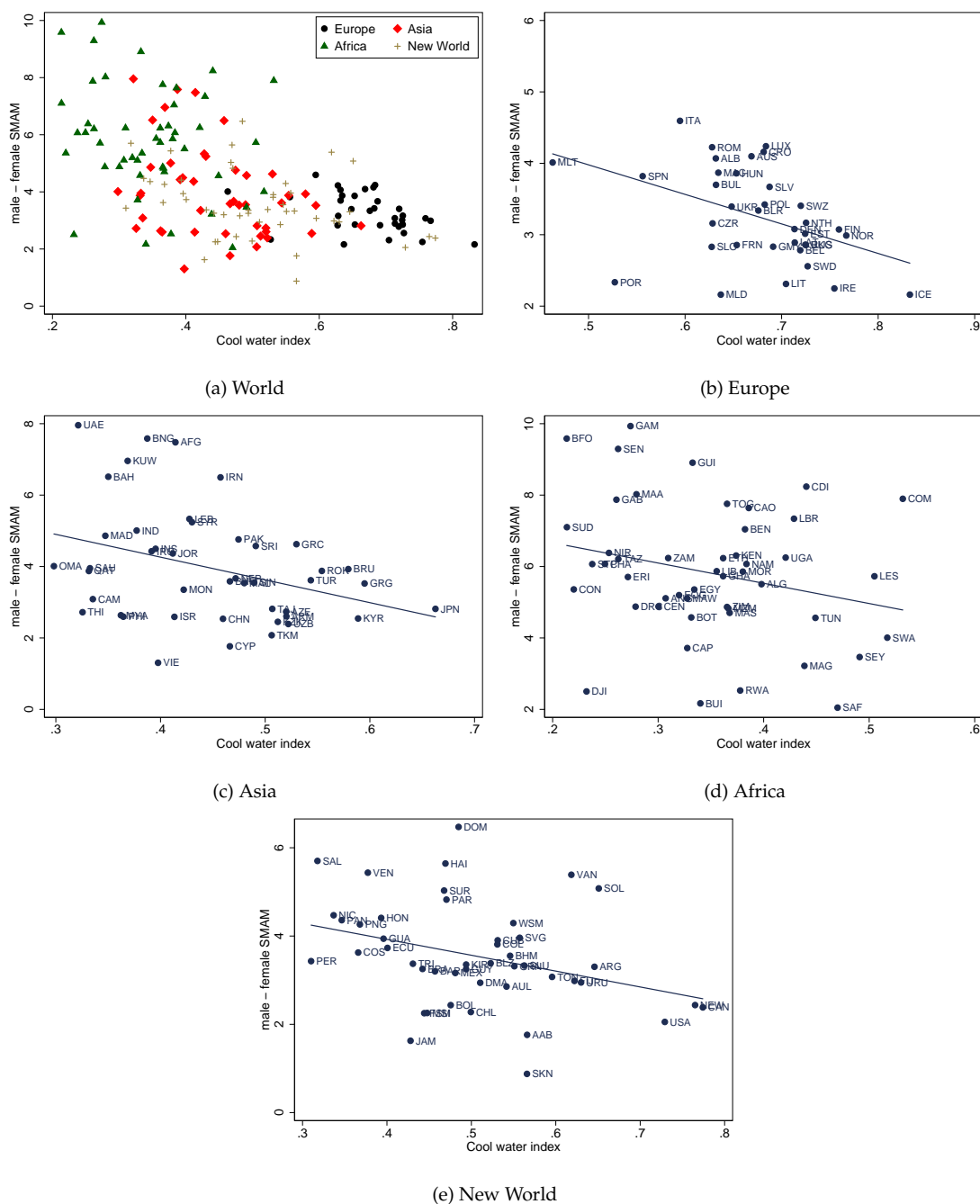
<sup>27</sup>For a world distribution of the CW-index see Figure 3.A.2.

<sup>28</sup>We use a large set of additional control variables in our robustness checks. For convenience, these are introduced in the text whenever necessary. For a full list see Table 3.A.1.



**FIG. 3.2:** Historical and contemporary female SMAM for 28 European countries

*Notes:* Linear correlation controlling for the year of observation of contemporary female SMAM, 1966–2002. Historical data from Dennison and Ogilvie (2014, Table 2); contemporary data from UN (2009). SMAM is singulate mean age at first marriage.



**FIG. 3.3:** CW-condition and male-female difference in SMAM

*Notes:* Period is 1960–2006. For each country first year available is shown. Data from UN (2009). SMAM is singulate mean age at first marriage.

### 3.4 Empirical strategy

We test our hypotheses by estimating regressions of the form:

$$y_i = \alpha + \mathbf{H}_i' \boldsymbol{\beta} + \mathbf{X}_i^H \boldsymbol{\gamma} + \mathbf{X}_i^C \boldsymbol{\delta} + \theta_c + \epsilon_i \quad (3.1)$$

where  $y_i$  is the outcome variable of interest for country  $i$ ;  $\mathbf{H}_i$  is a vector of potential deep determinants—*CW-index*, *years of agriculture*, *plow*, and *agricultural suitability*;  $\mathbf{X}_i^H$  is a vector of historical controls of country  $i$ 's ancestral population;  $\mathbf{X}_i^C$  is a vector of contemporary control variables;  $\theta_c$  is a continent fixed effect. To ensure comparability to the previous literature we take the same set of historical controls used by Alesina et al. (2013). Unlike Alesina et al. (2013), we do not include the proportion of ancestral land that is tropical or subtropical in our baseline regressions because tropical areas are already partially captured in our CW measure: they correlate very strongly with lower latitudes. Accordingly, ancestral tropical climate correlates strongly and negatively with the CW-index ( $r \approx -0.72$ ), while being, in our view, a much less fine-grained geoclimatic measure than the CW-index. Nevertheless, we estimate additional regressions where tropical climate is included to show that the CW-index is not merely capturing different degrees of exposure to tropical climates. As contemporary controls, we also follow previous literature in using the natural log of per capita income and its square, typically referring to the same time period of the outcome variable,  $y$ . The inclusion of income levels on the right hand side of the regression equation raises the issue of endogeneity, to the extent that most of the other historical regressors are thought to partially determine current income levels. However, since we are interested in the persistent effect of deep determinants in contemporary outcomes, it seems natural to condition on contemporary income levels. The empirical exercise then asks the following question: how relevant are deep rooted variables for explaining the share of variation in  $y$  that is left unexplained once current income levels have been taken into account?

The main potential flaw of our empirical strategy is that the effect of the CW-index might be spurious due to omitted variable bias. Given the cross-sectional nature of our data, we cannot remove time-invariant unobservables by employing country fixed effects. We go at great lengths to convince the reader that our relationships of interest are not driven by third factors. First, we explicitly control for several candidates of such omitted factors in additional regressions. Second, we estimate how large selection on unobservables needs to be relative to selection on observable characteristics in order to fully explain away the CW-effect (Altonji et al., 2005; Oster, 2017). Third, we test the

sensitivity of our estimates to subsample selection. In particular, we differentiate between Old versus New World samples to prevent mass migration movements post-1500 from biasing our results (e.g., Olsson and Paik, 2016, p. 205), and exclude, respectively, Europe and Sub-Saharan Africa from the sample because the former is home to the most extreme historical version of a late marriage pattern and the latter is unique in its prevalence of polygynous marriages.

## 3.5 Results

### 3.5.1 Ages at first marriage and gender equality

We start by establishing that contemporary ages at first marriage between the spouses correlate with gender equality today (Table 3.2). We use several alternative measures for contemporary ages at first marriage: the female and male SMAM, the ratio of female-to-male SMAM, and the difference between male and female SMAM.<sup>29</sup> The indicators of gender equality are the average female-to-male ratios in labor force participation rates, life expectancy, and years of education for the period 1990–2010.<sup>30</sup> Due to reverse-causality concerns, we use the earliest available year of SMAM data for each country and exclude countries whenever this year is later than 1990. To account for possible worldwide trends in SMAM over time, we include the year of the SMAM observation as a control variable.<sup>31</sup>

As shown in Table 3.2, countries with older brides and younger grooms have higher female-to-male labor force participation ratios (column 1), higher female-to-male life expectancy ratios (column 4), and higher female-to-male years of education ratios (column 8). These relationships are confirmed by the positive and highly significant effect of the female-to-male SMAM ratio for gender ratios in labor force participation and life expectancy (columns 2 and 5), as well as the negative and highly significant effect of

<sup>29</sup>Note that these three measures have different interpretations. When including both the female and male SMAM as separate regressors, their coefficients estimate the effect on  $y$  of one additional SMAM, on average, for gender  $X$ , holding the SMAM for gender  $Y$  constant. When using the ratio between female and male SMAM, one implicitly weights the age differences between the spouses in the inverse proportion of their age levels. When using the simple difference between male and female SMAM, one weights age differences equally, irrespective of the average age level of the spouses at first marriage.

<sup>30</sup>We average the dependent variables over a 20-year period because a single year might be unrepresentative of the actual cross-sectional differences between countries. The results are robust to using the dependent variables for 2000, instead of averaging between 1990–2010; available upon request.

<sup>31</sup>As a robustness check, we run the analyses taking the earliest data point for the period 1985–1994. By reducing the time window considerably, we can avoid potential problems of inter-temporal comparison. While our sample size is reduced from 119–132 countries to 101–115, the results do not change qualitatively and are available upon request.

TABLE 3.2: Determinants of gender gaps: ages at first marriage

	Average female-male ratio in 1990–2010								
	Labor force participation			Life expectancy			Years of education		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Ages at first marriage:</i>									
Female	0.037*** (0.010)			0.006*** (0.002)			0.017* (0.009)		
Male	-0.032*** (0.012)			-0.006*** (0.002)			-0.005 (0.010)		
Female/male		1.014*** (0.269)			0.164*** (0.043)			0.425* (0.245)	
Male–female			-0.035*** (0.010)			-0.006*** (0.002)			-0.014 (0.009)
Year of obs.	-0.002 (0.003)	-0.002 (0.003)	-0.002 (0.003)	0.001 (0.000)	0.000 (0.000)	0.001 (0.000)	-0.003 (0.003)	-0.003 (0.003)	-0.002 (0.003)
<i>Deep determinants:</i>									
Years of agriculture	-0.041*** (0.010)	-0.040*** (0.010)	-0.041*** (0.010)	-0.004** (0.002)	-0.004** (0.002)	-0.004** (0.002)	-0.024** (0.010)	-0.025** (0.010)	-0.026** (0.011)
Plow	-0.095* (0.052)	-0.095* (0.052)	-0.099* (0.052)	-0.004 (0.009)	-0.002 (0.009)	-0.003 (0.009)	0.023 (0.098)	0.017 (0.096)	0.015 (0.096)
Agricultural suitability	0.115** (0.052)	0.118** (0.053)	0.118** (0.052)	-0.001 (0.010)	-0.001 (0.010)	-0.001 (0.010)	-0.046 (0.071)	-0.044 (0.069)	-0.045 (0.069)
<i>Historical controls:</i>									
Large animals	0.018 (0.081)	0.015 (0.083)	0.023 (0.081)	-0.020 (0.020)	-0.022 (0.020)	-0.020 (0.020)	-0.190* (0.104)	-0.187* (0.106)	-0.186* (0.106)
Political hierarchies	-0.005 (0.020)	-0.006 (0.020)	-0.005 (0.020)	0.001 (0.003)	0.001 (0.003)	0.001 (0.003)	0.006 (0.016)	0.007 (0.016)	0.008 (0.016)
Economic complexity	0.012 (0.013)	0.012 (0.013)	0.013 (0.013)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.011 (0.014)	-0.010 (0.014)	-0.010 (0.014)
Tropical climate	-0.145*** (0.050)	-0.140*** (0.047)	-0.139*** (0.048)	-0.024*** (0.009)	-0.025*** (0.008)	-0.025*** (0.008)	-0.048 (0.048)	-0.036 (0.048)	-0.036 (0.048)
<i>Contemporary controls:</i>									
Income per capita (log)	-0.547*** (0.159)	-0.557*** (0.158)	-0.527*** (0.160)	0.137*** (0.033)	0.130*** (0.030)	0.135*** (0.031)	0.588*** (0.192)	0.618*** (0.185)	0.629*** (0.185)
(Income per capita (log)) <sup>2</sup>	0.028*** (0.009)	0.029*** (0.009)	0.027*** (0.009)	-0.008*** (0.002)	-0.008*** (0.002)	-0.008*** (0.002)	-0.030*** (0.011)	-0.031*** (0.011)	-0.031*** (0.011)
<i>Continent dummies:</i>									
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	131	131	131	132	132	132	119	119	119
R <sup>2</sup>	0.581	0.585	0.580	0.571	0.574	0.571	0.630	0.627	0.625
adj. R <sup>2</sup>	0.518	0.527	0.521	0.507	0.515	0.511	0.568	0.568	0.566

Notes: OLS estimates are reported with robust standard errors in parentheses. "Ages at first marriage" are single-year mean years at first marriage (SMAM) from UN (2009) for the period 1960–1990. For each country, earliest year available is selected. "Years of agriculture" is the number of years (in thousands) since the Neolithic revolution (from 1500) from Puterman and Trainor (2006). "Plow" is the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013). "Agricultural suitability" and *Historical controls* are from Alesina et al. (2013). The natural log of per capita income and its square are measured in the same time period as the dependent variable.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

the male to female difference in SMAM for the same outcomes (columns 3 and 6). For gender education ratios, the coefficients have the expected sign but are either statistically insignificant (column 9) or barely significant (column 8). However, this result is driven by Middle Eastern and North African countries. The MENA region experienced large increases in female education (e.g., in the Gulf States, by allowing women to pursue higher education) without corresponding improvements in labor market participation or ages at first marriage.<sup>32</sup>

Overall, the estimated coefficients of Table 3.2 have economic relevance. They suggest, for example, that a one-year reduction in the average age difference between the groom and the bride is associated, on average, with a 3.5 percentage point increase in the ratio of female-to-male labor force participation and 0.6 percentage point increase in the ratio of female-to-male life expectancy years.

Consistent with Hansen et al. (2015), longer histories of agriculture are negatively and significantly correlated with gender equality in labor force participation, health, and education. Ancestral plow use is also negatively correlated with gender equality in labor force participation, which replicates the findings of Alesina et al. (2013). It is also the case that historical agricultural suitability is positively associated with higher female participation in the labor force, as hypothesized by Hazarika et al. (2015). However, both the plow and agricultural suitability are uncorrelated to gender equality in life expectancy and education. Since, at this stage, we do not use the CW-index as a regressor, we can reproduce the full specification of Alesina et al. (2013), which includes tropical climate in the vector of historical controls. All else being equal, countries with a high share of tropical ancestors perform significantly worse in labor force participation and life expectancy but not significantly worse in education.

The coefficients of income per capita, both the linear and the quadratic, are always highly significant. However, their signs imply different relationships between income levels and gender equality depending on the dimension considered. For labor force participation, there is evidence of a U-shaped relationship consistent with the feminization-U hypothesis (Boserup, 1970; Goldin, 1995).<sup>33</sup> For the health and education dimensions, in contrast, there is an inverted-U shaped relationship between income and gender

<sup>32</sup> If we exclude the MENA countries for the regression in Table 3.2, column 9, the effect of male-to-female SMAM difference increases in absolute magnitude by a factor of 1.7 and becomes highly significant:  $\hat{\beta}_{w/o\ MENA} = -0.024$ , robust s.e. = 0.009, p-value = 0.008. Including the MENA region, but measuring the ratio of female-to-male education in 1990 (instead of the 1990–2010 average) also produces a larger coefficient (in absolute terms) than that of column 9:  $\hat{\beta}_{1990} = -0.018$ , robust s.e. = 0.008, p-value = 0.028.

<sup>33</sup> The turning point (i.e., the minimum of the U curve) implied by the estimates on column 1 is at 15,679 per capita PPP-\$. Whereas the feminization U-hypothesis holds for cross-sections of countries, its empirical relevance has been challenged in panel data analyses (see Gaddis and Klasen, 2014).

equality, suggesting decreasing marginal returns to income.<sup>34</sup>

Furthermore, higher SMAM for women (and lower SMAM for men) positively and significantly correlates with alternative measures of gender equality previously used in the literature such as the share of firms with some degree of female ownership, the share of parliament seats held by women, and the UNDP's Gender Development Index, as shown in Table 3.A.2.

We explore the sensitivity of these results to subsample analyses (Table 3.3). Excluding Sub-Saharan Africa (Panel A), the Americas and Oceania, i.e., the New World (Panel B), or Europe (Panel C) does not affect the overall finding that higher ages at first marriage for women and lower ages at first marriage for men are positive and significant correlates of gender equality in labor force participation and life expectancy, although insignificant for gender equality in years of education. However, when the estimation is done for a sample of Old World countries, the SMAM coefficients become statistically significant for the education outcomes (Panel B: columns 7–9). A relevant finding, from the subsample analyses, is that both the effect of ancestral plow and agricultural suitability on labor force participation vanish once Sub-Saharan Africa is excluded (Panel A: columns 1–3). This is not surprising, given that Sub-Saharan Africa is the signature non-plow region in the world (Baumann, 1928; Boserup, 1970). At the same time, the SMAM effects are larger without Sub-Saharan African countries; for the male-female SMAM difference, for example, the coefficient is almost twice as large as in the full world sample.

### 3.5.2 Cool Water breeds late-marriage societies

Having shown that ages at first marriage are indeed important factors for gender equality, broadly defined, we now test if the CW-condition is a relevant determinant of gender gaps in ages at first marriage. We estimate the baseline specification, with the alternative SMAM variables on the left hand side and include the CW-index as a new explanatory variable.

Table 3.4 shows the results. As hypothesized, the CW-index has a positive and significant effect on female SMAM, the ratio of female-to-male SMAM and a negative effect on the difference between male and female SMAM. These effects do not disappear once the other deep determinants are included. In fact, the CW-impact on the gender

<sup>34</sup>For health, the relationship between income and the gender life expectancy ratio turns negative at a relatively per capita low income level of 3,924 PPP-\$. This is likely related to the fact that, in high income countries, the relative survival advantage of females falls, due to falling relative importance of a biological survival advantage of females in infancy and old age, and also due to greater similarity in economic activity and behavior. For years of education, the turning point occurs at a much higher per capita level of 19,765 PPP-\$ (using estimates from columns 4 and 7, respectively).



TABLE 3.3: Determinants of gender gaps: ages at first marriage; subsample analysis

	Average female-male ratio in 1990–2010								
	Labor force participation			Life expectancy			Years of education		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: without Sub-Saharan Africa									
SMAM:									
Female	0.069*** (0.016)			0.006* (0.003)			0.002 (0.013)		
Male	-0.064*** (0.018)			-0.008** (0.003)			0.011 (0.015)		
Female/male		1.751*** (0.413)			0.159** (0.076)			0.138 (0.334)	
Male–female			-0.068*** (0.016)			-0.006** (0.003)			-0.000 (0.013)
Deep determinants:									
Years of agriculture	-0.037*** (0.010)	-0.038*** (0.011)	-0.038*** (0.010)	-0.006*** (0.002)	-0.005*** (0.002)	-0.005*** (0.002)	-0.014* (0.008)	-0.015* (0.008)	-0.015* (0.008)
Plow	0.068 (0.050)	0.065 (0.048)	0.061 (0.049)	0.000 (0.013)	0.005 (0.013)	0.004 (0.012)	0.016 (0.064)	0.003 (0.062)	-0.001 (0.062)
Agricultural suitability	0.042 (0.059)	0.041 (0.056)	0.050 (0.056)	0.013 (0.014)	0.008 (0.013)	0.009 (0.013)	-0.055 (0.057)	-0.047 (0.057)	-0.045 (0.057)
N	97	97	97	98	98	98	90	90	90
R <sup>2</sup>	0.699	0.698	0.698	0.559	0.550	0.551	0.605	0.593	0.592
Panel B: Old World									
SMAM:									
Female	0.041*** (0.012)			0.007*** (0.002)			0.024** (0.011)		
Male	-0.035** (0.014)			-0.006** (0.003)			-0.002 (0.013)		
Female/male		1.131*** (0.300)			0.195*** (0.046)			0.554** (0.268)	
Male–female			-0.039*** (0.011)			-0.007*** (0.002)			-0.017* (0.010)
Deep determinants:									
Years of agriculture	-0.035*** (0.011)	-0.035*** (0.011)	-0.036*** (0.011)	-0.004** (0.002)	-0.004** (0.002)	-0.004** (0.002)	-0.019 (0.012)	-0.020 (0.013)	-0.021 (0.013)
Plow	-0.122* (0.070)	-0.117* (0.069)	-0.121* (0.069)	-0.013 (0.011)	-0.012 (0.011)	-0.013 (0.011)	0.013 (0.155)	0.019 (0.157)	0.018 (0.158)
Agricultural suitability	0.097 (0.066)	0.102 (0.066)	0.096 (0.065)	-0.009 (0.012)	-0.008 (0.011)	-0.009 (0.012)	-0.044 (0.106)	-0.055 (0.103)	-0.059 (0.104)
N	105	105	105	105	105	105	94	94	94
R <sup>2</sup>	0.593	0.600	0.593	0.635	0.638	0.634	0.622	0.616	0.612
Panel C: without Europe									
SMAM:									
Female	0.033*** (0.010)			0.007*** (0.002)			0.014 (0.010)		
Male	-0.028** (0.012)			-0.005** (0.002)			-0.002 (0.011)		
Female/male		0.904*** (0.277)			0.188*** (0.043)			0.349 (0.264)	
Male–female			-0.031*** (0.010)			-0.006*** (0.002)			-0.011 (0.009)
Deep determinants:									
Years of agriculture	-0.045*** (0.012)	-0.045*** (0.012)	-0.046*** (0.012)	-0.003 (0.002)	-0.003 (0.002)	-0.003 (0.002)	-0.026** (0.012)	-0.028** (0.012)	-0.029** (0.013)
Plow	-0.107* (0.055)	-0.106* (0.054)	-0.110** (0.054)	-0.003 (0.009)	-0.003 (0.009)	-0.004 (0.009)	0.011 (0.102)	0.003 (0.099)	0.001 (0.099)
Agricultural suitability	0.133** (0.053)	0.136** (0.053)	0.136** (0.052)	-0.004 (0.010)	-0.003 (0.009)	-0.003 (0.010)	-0.043 (0.074)	-0.041 (0.072)	-0.042 (0.071)
N	101	101	101	102	102	102	90	90	90
R <sup>2</sup>	0.589	0.592	0.587	0.512	0.522	0.509	0.620	0.616	0.614

Notes: OLS estimates are reported with robust standard errors in parentheses. All regressions include the same set of historical and contemporary controls as in Table 3.2. *Historical controls* are: ancestral domestication of large animals, ancestral settlement patterns, ancestral political complexity, and fraction of ancestral land that was tropical or subtropical from Alesina et al. (2013). *Contemporary controls* are the natural log of per capita income and its square, measured in the same time period as the dependent variable, and the year of the SMAM observation. Continent dummies are included. *Panel B: Old World* shows results of regressions in which all countries from the Americas and Oceania (i.e., the *New World*) are excluded. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

differences becomes *larger* in absolute terms. The estimates in column 8, for example, imply that a one standard deviation increase in the CW-index is associated with a 10 months reduction on the average gap between male and female ages at first marriage.<sup>35</sup> This is a sizable effect: 10 months corresponds to roughly 20 percent of the world's average gender gap in ages at first marriage, which is 4.12 years in the period considered (see Table 3.1). In contrast, none of the other deep determinants—whether years of agriculture, historical plow use, or historical agricultural suitability—are significant at the 5 percent level, once the CW-index is included. Interestingly, the ages at first marriage for women and men increase with per capita income (averaged over the period 1960–1980), following an inverted U-shaped function that peaks around 13,000–13,500 PPP-\$. But there is no evidence that women's SMAM approaches men's as countries get richer, since the income coefficients for the female-male age ratio or the male-female age difference are not statistically significant.<sup>36</sup> These results suggest that while marital ages do respond to economic development and follow global trends they do so similarly for *both* sexes, thus leaving ratios and differences untouched. Over time, as income levels rise, there is no evidence of convergence between female and male ages at first marriage, which supports the view that persistent, deep-rooted patterns dominate this relationship.<sup>37</sup>

The results hold for the usual three subsamples: without Sub-Saharan Africa, without the New World, and without Europe, as shown in Table 3.5.

We have thus far a negative correlation between the CW condition and gender gaps in ages at first marriage in the postwar period. Our interpretation is however that this present day correlation emerges from the effect of geo-climatic conditions on *preindustrial* marriage patterns. Over time, these marriage patterns became embedded in both formal and informal institutions, thereby persisting until today. Although we cannot empirically exclude that the CW condition has a direct contemporary effect of ages at first marriage, we can ask whether the CW effect is mainly tied to a geographical area *as such* or to its ancestral inhabitants (Putterman and Weil, 2010). Hansen et al. (2015) show that the negative effect of longer histories of agriculture on gender equality becomes

<sup>35</sup>  $\hat{\beta}_{cw} * \sigma_{cw} = -5.902 * 0.145 \approx -0.856$  years ( $\approx -10.27$  months).

<sup>36</sup> The coefficients for the year of the SMAM observation tell a similar story: while ages at first marriage have increased over time for both women and men (columns 2 and 4), the time trends for the ratio or differences between the sexes are not statistically significant.

<sup>37</sup> In Table 3.A.3, we “unpack” the CW index by replacing it in the regressions with its raw variables. Absolute latitude and summer mildness correlate negative and significantly with the male-female age gap at first marriage. The coefficient for continuous rain is also negative but not significant at conventional levels. In column 7, we show that both factor scores, *coolness* and *water*, have large, negative, significant effects. Moreover, there is no evidence of additional interaction effects between the two factors (column 8).

TABLE 3.4: Determinants of ages at first marriage

	Singulate mean age at first marriage							
	Female		Male		Female/male		Male–female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Deep determinants:</i>								
Cool water	6.727*** (2.546)	4.316* (2.494)	1.949 (2.423)	-1.555 (2.402)	0.195*** (0.047)	0.221*** (0.075)	-4.672*** (1.302)	-5.902*** (2.078)
Years of agriculture		-0.312** (0.147)		-0.159 (0.112)		-0.007* (0.004)		0.154* (0.092)
Plow		-1.105 (1.025)		-0.253 (0.860)		-0.034 (0.026)		0.890 (0.686)
Agricultural suitability		0.310 (0.906)		0.728 (0.824)		-0.019 (0.024)		0.523 (0.646)
<i>Historical controls:</i>								
Large animals	1.074 (1.121)	0.731 (1.363)	0.694 (1.035)	0.336 (1.096)	0.020 (0.023)	0.019 (0.035)	-0.361 (0.575)	-0.395 (0.877)
Political hierarchies	0.658*** (0.245)	0.753** (0.295)	0.329 (0.203)	0.254 (0.259)	0.013** (0.006)	0.020** (0.008)	-0.303* (0.163)	-0.482** (0.218)
Economic complexity	-0.242 (0.156)	-0.061 (0.121)	-0.030 (0.106)	0.064 (0.110)	-0.007 (0.005)	-0.003 (0.004)	0.190 (0.128)	0.099 (0.112)
<i>Contemporary controls:</i>								
Income p.c. (log)	5.726** (2.369)	6.092*** (2.113)	6.471*** (1.951)	6.632*** (1.807)	-0.002 (0.062)	0.010 (0.061)	0.918 (1.625)	0.759 (1.631)
(Income p.c. (log)) <sup>2</sup>	-0.307** (0.137)	-0.321*** (0.119)	-0.350*** (0.115)	-0.349*** (0.102)	0.000 (0.003)	-0.001 (0.003)	-0.052 (0.092)	-0.039 (0.092)
Year of SMAM obs.	0.080*** (0.022)	0.059*** (0.022)	0.061*** (0.020)	0.050** (0.021)	0.001** (0.001)	0.001 (0.001)	-0.025 (0.016)	-0.015 (0.016)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	138	125	134	121	134	121	134	121
R <sup>2</sup>	0.459	0.433	0.312	0.253	0.495	0.517	0.468	0.498
adj. R <sup>2</sup>	0.407	0.355	0.244	0.146	0.445	0.448	0.415	0.426

Notes: OLS estimates are reported with robust standard errors in parentheses. "Singulate mean age at first marriage" data are from UN(2009) for the period 1960-2006. For each country, earliest year available is selected and controlled for with variable "Year of SMAM obs." "Cool water" is the cool water index described in section 3.3. "Years of agriculture" is the number of years (in thousands) since the Neolithic revolution (from 1500) from Putterman and Trainor (2006). "Plow" is the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013). "Agricultural suitability" and Historical controls are from Alesina et al. (2013). The natural log of per capita income and its square are averaged over the period 1960-1980. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

stronger after weighing their variable on the timing of the Neolithic revolution with post-1500 migration flows from Putterman and Weil (2010). In the same spirit, we create an alternative version of the CW index by weighing it with Putterman and Weil's World Migration Matrix data. This "ancestry-adjustment" strengthens the CW effect (see Table 3.A.4 for point estimates), suggesting that the historical CW condition of a population matters more than the CW condition of its present-day place of residency. This is consistent with our hypothesis that the results reveal the long-run persistence of a preindustrial relationship, rather than a contemporary one, between a society's geo-climatic configuration and its household formation patterns. However, adjusting the CW index in this way is problematic if the migration flows are endogenous to marriage patterns, as in the case where areas with favorable CW conditions were to

TABLE 3.5: Determinants of ages at first marriage: subsample analysis

	Singulate mean age at first marriage							
	Female		Male		Female/male		Male–female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: without Sub-Saharan Africa								
<i>Deep determinants:</i>								
Cool water	5.489*	3.786	1.277	-1.140	0.170***	0.193**	-4.208***	-5.038**
	(3.123)	(3.121)	(2.934)	(2.633)	(0.045)	(0.081)	(1.170)	(2.082)
Years of agriculture		-0.263		-0.167		-0.004		0.100
		(0.164)		(0.135)		(0.004)		(0.095)
Plow		-3.671**		-2.571*		-0.056*		1.265
		(1.398)		(1.329)		(0.030)		(0.816)
Agricultural suitability		1.683		2.023*		-0.005		0.303
		(1.416)		(1.201)		(0.027)		(0.634)
<i>N</i>	96	85	95	84	95	84	95	84
<i>R</i> <sup>2</sup>	0.418	0.406	0.370	0.338	0.344	0.374	0.308	0.355
adj. <i>R</i> <sup>2</sup>	0.333	0.277	0.278	0.192	0.248	0.236	0.206	0.213
Panel B: Old World								
<i>Deep determinants:</i>								
Cool water	7.409***	4.348*	2.200	-1.308	0.213***	0.206**	-5.092***	-5.639**
	(2.458)	(2.535)	(2.279)	(2.262)	(0.069)	(0.093)	(1.920)	(2.574)
Years of agriculture		-0.304**		-0.128		-0.007*		0.172*
		(0.127)		(0.091)		(0.004)		(0.100)
Plow		0.900		1.652**		-0.023		0.771
		(0.956)		(0.720)		(0.032)		(0.832)
Agricultural suitability		-0.143		0.121		-0.019		0.411
		(0.867)		(0.713)		(0.029)		(0.784)
<i>N</i>	104	98	100	94	100	94	100	94
<i>R</i> <sup>2</sup>	0.376	0.458	0.256	0.344	0.438	0.491	0.427	0.478
adj. <i>R</i> <sup>2</sup>	0.316	0.382	0.181	0.247	0.382	0.416	0.370	0.401
Panel C: without Europe								
<i>Deep determinants:</i>								
Cool water	7.553**	5.341*	2.929	-1.215	0.195***	0.253***	-4.521***	-6.664***
	(3.085)	(2.803)	(2.931)	(2.832)	(0.055)	(0.084)	(1.491)	(2.364)
Years of agriculture		-0.356**		-0.204*		-0.007		0.149
		(0.159)		(0.120)		(0.004)		(0.099)
Plow		-1.304		-0.321		-0.040		1.021
		(1.051)		(0.896)		(0.027)		(0.706)
Agricultural suitability		0.199		0.773		-0.026		0.708
		(0.979)		(0.907)		(0.026)		(0.697)
<i>N</i>	115	103	111	99	111	99	111	99
<i>R</i> <sup>2</sup>	0.461	0.432	0.315	0.264	0.466	0.488	0.436	0.467
adj. <i>R</i> <sup>2</sup>	0.403	0.342	0.239	0.141	0.407	0.402	0.373	0.378

Notes: OLS estimates are reported with robust standard errors in parentheses. All regressions include the same set of historical and contemporary controls as in Table 3.4. *Historical controls* are: ancestral domestication of large animals, ancestral settlement patterns, ancestral political complexity, and fraction of ancestral land that was tropical or subtropical from Alesina et al. (2013). *Contemporary controls* are the natural log of per capita income and its square averaged over the period 1960–1980, and the year of the SMAM observation. Continent dummies are included. *Panel B: Old World* shows results of regressions in which all countries from the Americas and Oceania (i.e., the *New World*) are excluded. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

attract immigrants with strong preferences for late marriage patterns of household formation. Indeed, Northwestern Europeans have largely settled in the regions of the New World with the highest score of the CW index.<sup>38</sup> The unadjusted CW measure is free from this specific source of endogeneity bias. Thus, even though the estimated effects are stronger with the ancestry-adjustment, we decide, as a matter of caution, to present the remaining results without this adjustment.

It is important to note that if, by construction, the CW index would uniquely fit the geo-climatic features of Northwest Europe and its New World colonies, then it would necessarily be also correlated with all the unobservable factors that might explain its (potentially) unique preindustrial late marriage pattern. If this were the case, our results would be completely spurious. It is therefore essential to refute this possibility. First, we use a data-driven approach (factor analysis) for the construction of the CW index in order to minimize the concern that the world distribution of this variable results from *ad hoc* measurement assumptions. Second, we show that our results are not dependent on Northwest Europe or Western offshoots. Both including dummy variables for these groups of countries or excluding them altogether from the estimation sample does not affect the main result: a highly significant and negative effect of the CW index, only slightly weaker in magnitude (Table 3.A.5).<sup>39</sup> In other words, the association of the CW index with ages at first marriage is not a spurious idiosyncrasy of Northwest Europe and its offshoots, but a broader relationship that holds for the rest of the globe.

The inclusion of further control variables does not affect the relationship between the CW-index and the average age gap between groom and bride. We start by including tropical climate in column 2 of Table 3.6, since it could be that the relevant variation captured by the CW-index is that between tropical and non-tropical countries. The results show otherwise; the coefficient for CW remains negative and highly significant, whereas the tropical climate variable is statistically indistinguishable from zero. This demonstrates that the CW-index is more than just an inverse measure of tropical temperatures. What distinguishes it from such an inverse measure is that it gives a premium not just on high latitude but more specifically on high latitude with minimized seasonal extremity and continuous rain.

Another possible source of error is that the CW-index captures the fact that European

<sup>38</sup>The correlation coefficient of the CW indexes adjusting or not for post-1500 migration is 0.96. For comparison, the correlation between the adjusted and unadjusted years of agriculture variable used by Hansen et al. (2015) is 0.85.

<sup>39</sup>Northwestern Europe includes: Belgium, Denmark, France, Germany, Iceland, Ireland, Netherlands, Norway, Sweden, and the United Kingdom. Excluding France from this list has no impact on the results. Western offshoots are Australia, Canada, New Zealand, and the United States.

colonizers settled by and large in all the major CW-areas outside Europe, with the exception of Japan and the Korean peninsula. Thus, it could simply be that settlers from late marriage European societies “exported” the late marriage pattern to their overseas offshoots. Even though the subsample analyses of Tables 3.5 and 3.A.5 do not support this argument, we provide a more rigorous test by including the weighted genetic distance between each country and the United Kingdom from Spolaore and Wacziarg (2009).<sup>40</sup> If the CW-index was indeed a mere proxy for areas of European settlement around the world, we would expect the relationship between CW and ages at first marriage to vanish once the genetic distance from Western Europe is held constant. However, as seen in column 3, the coefficient of CW remains negative, statistically significant, and, if anything, the effect becomes *stronger*. On column 4, we go beyond controlling for contemporary income difference across countries and also control for preindustrial differences in the level of development. Following the literature, we use population density in 1500 as a proxy for development in the Malthusian era (e.g., Spolaore and Wacziarg, 2013, footnote 3). But the CW-effect does not change.

Alternatively, it could be the case that per capita income levels do not reasonably proxy other developmental dimensions that might be driving the correlation between the CW-index and ages at first marriage. In particular, education levels and formal institutions are plausible candidates for such omitted factors. More educated individuals marry later, and better formal institutions could be stronger at enforcing minimum-marital-ages legislation, or recognizing individual consent as the basis for a lawful marriage. In columns 5–6 we include, respectively, the mean years of education in 1950 for the total population and by gender; in columns 7–8, we include the *polity2* score in 1980 as a measure of democracy and the World Bank’s rule of law variable in 2000 (Kaufmann et al., 2011) as a measure of institutional quality. The CW coefficient remains negative and significant, and of comparable magnitude throughout.

Religion poses a particular challenge. While certain authors regard religion as a crucial determinant of gender inequality (e.g., Fish, 2002; Inglehart and Norris, 2003; Lagerlöf, 2003; Carmichael, 2011), others have argued that religion is endogenous to pre-existing factors. As such, religion would be a “bad control” to include since it would shut down important transmission mechanisms from the deep determinants to the outcome of interest. Boserup (1970) recognized this problem in her original plow vs. shifting agriculture argument. In particular, she claims that the use of the veil or the burqa was a direct consequence of female domestic seclusion due to plow agriculture;

<sup>40</sup>The weighted genetic distance is the expected value of the genetic distance between two randomly picked individuals for each pair of countries. See more details in Spolaore and Wacziarg (2009).

only afterwards did it become incorporated in the religious practices of those societies. Being a consequence of the plow, religion would mistakenly absorb much of the plow's gender-inegalitarian effect in any regression of gendered outcomes. Alesina et al. (2013) find indeed that including religion reduces the effect of the plow by 20 percent.

In general, this strand of argument claims that emerging religions absorbed, incorporated and codified many pre-existing local practices and beliefs, rather than having introduced them. Hartman (2004) argues that, in medieval Europe, the Catholic doctrine of individual consent being a sufficient condition for the validity of marriage was widely followed in the Northwestern societies but rarely so in (deeply Catholic) Italy, Spain, and Southern France. Similarly, Hansen et al. (2015, p. 378) and Hazarika et al. (2015, pp. 19–20) discuss how pre-existing gender norms influenced early Islamic doctrine.

Despite the controversy on whether religion is a “bad control”, we include the population shares of Catholics, Protestants, and Muslims in 1980 (with other religions as the reference group). The CW effect is still negative and statistically significant (column 9) but, as expected, about 32 percent smaller in magnitude. Relative to other religious groups, higher shares of Catholics and Protestants are associated with smaller age gaps between the spouses, whereas a larger share of Muslims in the population correlates with a wider age gap between husband and wife at first marriage.

Finally, we add a country's per capita oil production in 2000 to capture Ross's (2008) argument that high oil endowments crowd-out women from the labor force. The low employment prospects and, consequently, low returns to education for women could incentivize early female ages at first marriage. Indeed, per capita oil production is positively correlated with larger age gaps between the groom and the bride, but its coefficient is insignificant. In any case, the CW effect remains unchanged by the inclusion of this additional control (column 10). Even when we include all the additional controls simultaneously in column 11, the negative correlation between the CW-index and the male-to-female difference in age at first marriage remains significant at the 5 percent level.

In addition to the regressions of Table 3.6, we perform further robustness checks and present the results in Table 3.A.6 of the Online Appendix. As additional historical controls, we include a measure for the preindustrial intensity of agriculture and the proportion of ancestral subsistence provided by animal husbandry (taken from Alesina et al., 2013)<sup>41</sup> to account for the possibility that animal husbandry delays ages at marriage

---

<sup>41</sup>Unless otherwise noted, all additional variables included are from Alesina et al. (2013). For the original source and construction method see their Online Appendix. We would like thank the authors for making their dataset publicly available.

TABLE 3.6: Determinants of ages at first marriage: additional controls

	Singulate mean ages at first marriage: male–female										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Cool water	-5.902*** (2.078)	-6.054*** (2.315)	-6.041*** (2.087)	-5.921*** (2.051)	-5.291** (2.395)	-5.335** (2.404)	-6.450*** (2.372)	-5.781*** (2.056)	-4.016** (1.787)	-5.464** (2.173)	-4.870** (2.353)
Years of agriculture	0.154* (0.092)	0.155* (0.093)	0.068 (0.109)	0.133 (0.094)	0.095 (0.099)	0.090 (0.105)	0.141 (0.100)	0.116 (0.097)	-0.071 (0.098)	0.149* (0.090)	-0.130 (0.100)
Plow	0.890 (0.686)	0.892 (0.689)	0.687 (0.734)	1.032 (0.715)	0.867 (0.766)	0.881 (0.786)	0.999 (0.743)	1.064 (0.671)	0.431 (0.669)	0.877 (0.679)	1.297* (0.754)
Agricultural suitability	0.523 (0.646)	0.523 (0.649)	0.556 (0.643)	0.362 (0.667)	0.606 (0.716)	0.621 (0.723)	0.551 (0.656)	0.699 (0.628)	0.685 (0.572)	0.574 (0.650)	0.754 (0.620)
Tropical climate		-0.091 (0.443)									-0.882 (0.603)
F <sub>ST</sub> from U.K. (weighted)			-5.727 (3.534)								1.599 (4.385)
Population density in 1500				0.018 (0.014)							0.023* (0.012)
Total years of schooling in 1950 (log)					-0.327 (0.250)						-0.036 (0.230)
Male years of schooling in 1950 (log)						-0.147 (0.352)					
Female years of schooling in 1950 (log)						-0.154 (0.248)					
Polity2 in 1980							-0.011 (0.024)				0.045* (0.026)
Rule of law in 2000								-0.360 (0.239)			-0.598* (0.303)
Catholic shares in 1980									-0.009* (0.005)		-0.006 (0.007)
Protestant shares in 1980									-0.017** (0.007)		-0.013 (0.008)
Muslim shares in 1980									0.022*** (0.006)		0.024*** (0.008)
Oil production (per capita)										1.057 (1.038)	1.687 (1.030)
Historical & contemporary controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	121	121	120	119	109	109	113	121	116	121	99
R <sup>2</sup>	0.498	0.498	0.514	0.516	0.508	0.509	0.492	0.511	0.630	0.503	0.679

Notes: OLS estimates are reported with robust standard errors in parentheses. *Additional controls*: “Tropical climate” is from Alesina et al. (2013), “F<sub>ST</sub> weighted genetic distance to the U.K.” is from Spolaore and Wacziarg (2009), “Population density in 1500” is from Klein Goldewijk et al. (2010), years of education in 1950 are from Barro and Lee (2013), “Polity in 1980” is the polity2 score from the Center for Systemic Peace, “Rule of law in 2000” is the rule of law indicator from the World Bank’s World Governance Indicators (Kaufmann et al. 2011), “Religious shares in 1980” are the shares of the population of different religions from La Porta et al. (1999), “Oil production (per capita)” is the number of barrels produced per person per day in 2000 from Alesina et al. (2013). All regressions include the same set of historical and contemporary controls as in Table 3.4. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



for women (Voigtländer and Voth, 2013). We also add the proportion of ancestral subsistence provided by hunting since hunter-gatherer societies display higher levels of gender equality (Dyble et al., 2015). To test the idea that male dominance over women derives from the emergence of private property (Engels, 1902), we also include the share of ancestors from ethnicities where land inheritance rules were absent. Furthermore, as additional contemporary controls, we include two warfare variables: both the number of years of civil and inter-state conflict for each country from 1816 until 2007 and the terrain ruggedness index from Nunn and Puga (2012). The latter is included because flatter regions are easier to invade but also easier to irrigate and plow than rugged terrain, and may also experience less rainfall than mountainous regions. War could either be detrimental for women if it reinforces gender violence and patriarchy in society, but it could also have positive effects if women are called to replace men in the labor force, thus postponing marriage (Whyte, 1978). A more direct effect of war is to reduce the supply of young men in the marriage market leading to later marital ages or higher proportions of never married women. To complement the genetic distance variable and the World Migration Matrix in measuring post-1500 global migratory flows, we add the share of a country's population (in 2000) that is of Western European descent. We also include a communist dummy since communist regimes had explicit policies to promote gender equality and, in some cases, fought traditional marriage practices such as arranged marriages, or child marriages. Finally, we add the share of GDP accruing to agriculture, manufacturing, or services in 2000 since labor demand in female-dominated sectors will likely impact female marital ages (e.g., Ross, 2008).

Controlling for these additional variables, both in a stepwise manner or simultaneously, does not affect our main result: the CW coefficient is always statistically significant at least at the 5 percent level and its size ranges from -5.206 to -6.296 (Table 3.A.6).<sup>42</sup> Overall, after controlling for a myriad of additional variables, we confirm our baseline finding that countries with a stronger CW-condition have systematically lower age gaps at first marriage between spouses.

Finally, we estimate how large the ratio ( $\delta$ ) of selection on unobservables relative to selection on observables needs to be in order to explain away the CW effect (Altonji et al., 2005; Oster, 2017). Table 3.7 presents the estimates for  $\delta$ , using the method described in Oster (2017). In our restricted regression, the male-female difference in SMAM is regressed on the CW-index and continent dummies. The controlled regressions

---

<sup>42</sup>Reestimating the specifications of Tables 3.5 and 3.6 using ancestry-adjusted CW index and years of agriculture produces even stronger CW effects, which are always highly statistically significant. Results available upon request.

**TABLE 3.7:** Assessing unobservable selection: estimates of  $\delta$ 

Controls in		$\beta$ [ $R^2$ ]		$\delta$ for $\beta = 0$ given $R_{max}$		
Restricted set	Full set	Restricted (1)	Full (2)	$R_{max} = 1$ (3)	$R_{max} = 1.3\bar{R}$ (4)	$R_{max} = 1.25\bar{R}$ (5)
Continent dummies	as in Table 3.6, column 11	-5.239 [0.428]	-4.870 [0.679]	0.416	0.636 [ $R_{max} = 0.882$ ]	0.749 [ $R_{max} = 0.848$ ]
Continent dummies	as in Table 3.A.6, column 10	-5.326 [0.433]	-5.669 [0.617]	0.451	0.864 [ $R_{max} = 0.802$ ]	1.009 [ $R_{max} = 0.771$ ]

Notes: The dependent variable in all regressions is the male-female SMAM difference from UN (2009).  $R_{max}$  is the assumed  $R^2$  from the hypothetical full regression, i.e., with both observable and unobservable variables included.  $\bar{R}$  is the  $R^2$  from the regression with the full set of observable controls. See Oster (2017) for more details on how to estimate  $\delta$ . Calculations were done using her Stata package *psacalc*.

are column 11 of Table 3.6 and column 10 of Table 3.A.6. We estimate  $\delta$  for different values of  $R_{max}$ , which is the assumed  $R^2$  from a hypothetical regression that includes all the observable and unobservable variables simultaneously. When  $R_{max}$  has the maximum possible value of 1, the degree of selection on unobservables would have to be 42–45 percent that of selection on the included observables for the CW effect to be zero. However, Oster (2017) argues that  $R_{max}$  is likely below 1, for example, due to measurement error in the outcome variable. She proposes, as a reasonable choice, a  $R_{max}$  that is 1.3 times greater than the  $R^2$  of the regression with the full set of observables.<sup>43</sup> When implementing this 1.3 factor,  $\delta$  is never below 0.6 (column 4); nor below 0.75, if we reduce the factor to 1.25 (column 5). Altonji et al. (2005) propose  $\delta = 1$  as an *ad hoc* cutoff for robustness. But given that the regressions with the full set of controls include most of the relevant variables proposed in the literature—from preindustrial and contemporary per capita income, economic structure (both historical and contemporary), to education, formal institutions, geography, warfare, migration, and religion—, we argue that it is unlikely that the degree of selection on unobservables is still at least 0.6 times as large as the degree of selection on all these observables.

### 3.5.3 Cool Water and historic late marriages

We now turn to Europe to show that the relationship between the CW-condition and ages at first marriage is deep rooted in history. In so doing we focus on a much more homogeneous sample: both marriage age as well as the CW-condition are more similar within Europe than between Europe and other parts of the world. Given the much lower variation and smaller sample size, it will be *harder* to find significant relationships within this more homogeneous group. The CW-index is a positive determinant of female marital ages (Table 3.8). Despite the much reduced variability, the CW coefficient continues to be

<sup>43</sup>Oster (2017) calculates this multiplicative factor of 1.3 as the cutoff for which  $\delta$  is less than unity in 90 percent of the results in a random sample of top-published randomized control trials in economics.

TABLE 3.8: Europe: historical female ages at first marriage

	Historical female age at first marriage, 1500–1900							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Deep determinants:</i>								
Cool water	14.424** (5.175)	20.872*** (6.894)	13.733 (10.357)	18.809 (11.132)	12.814** (5.708)	19.033** (7.291)	12.873** (5.443)	18.894*** (6.492)
Years of agriculture			0.105 (0.703)	-0.005 (0.643)				
Plow					-1.668 (1.168)	-2.028 (1.206)		
Agricultural suitability							-2.988* (1.728)	-4.361** (1.613)
<i>preindustrial development:</i>								
Population density in 1500		0.091*** (0.032)		0.093*** (0.032)		0.093*** (0.033)		0.098*** (0.031)
<i>N</i>	27	26	26	25	27	26	27	26
<i>R</i> <sup>2</sup>	0.152	0.299	0.105	0.267	0.161	0.313	0.187	0.374
adj. <i>R</i> <sup>2</sup>	0.118	0.238	0.027	0.163	0.091	0.219	0.119	0.289

Notes: OLS estimates are reported with robust standard errors in parentheses. “Historical female age at first marriage” data are from Dennison and Ogilvie (2014), see more details in section 3.3. Countries included: Austria, Belarus, Belgium, Bulgaria, Croatia, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Malta, Netherlands, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, and Ukraine. “Cool water” is the cool water index described in section 3.3. “Years of agriculture” is the number of years (in thousands) since the Neolithic revolution (from 1500) from Putterman and Trainor (2006). “Plow” is the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013). “Agricultural suitability” is from Alesina et al. (2013). “Population density in 1500” is from Klein Goldewijk et al. (2010). \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

statistically significant in all specifications, except when years of agriculture are included in columns 3 and 4. In those specifications, the standard errors of the CW-index jump upward due to multicollinearity between the CW-index and the timing of the Neolithic revolution ( $r \approx -0.70$ ) for this subsample of European countries. Vice versa, and in accordance with this finding, years of agriculture as well is insignificant under control of CW: within Europe, the two variables become almost indistinguishable and, hence, absorb each other. However, the size of the CW effect is not much affected, only its standard error. In columns 5 and 6, the plow variable is insignificant which, given that all countries included (except Iceland) were traditional plow societies, is not surprising.<sup>44</sup> Agricultural suitability is highly significant in column 8 but its negative sign runs against the positive effect of historical resource abundance on gender egalitarian norms posited by Hazarika et al. (2015).

Importantly, once again the results are not a statistical artifact driven by more developed areas being located in CW regions. The CW effect is robust to the inclusion of population density in 1500 as a proxy for preindustrial development. While it is true that societies with higher population densities had, on average, older brides at

<sup>44</sup>In other words, this sample of European countries, for which historical marital ages are available, removes any potential plow effect by virtue of almost exclusively containing plow societies.

first marriage, controlling for this actually increases the estimates of the CW variable. The reason is that most societies with extremely high CW scores (e.g., Iceland, Sweden, Norway, and Denmark) had relatively lower population densities and were *not* among the most wealthy and developed nations of Europe in this period (see also Dennison and Ogilvie, 2014).

Moreover, the positive and significant relationship between the CW-index and historical female ages at first marriage holds for a nineteenth century sample of both European and non-European countries, using data from Gapminder (Table 3.A.7). In this setting, the CW-effect retains significance, even taking the timing of the Neolithic Revolution into account.<sup>45</sup>

### 3.5.4 Cool Water and contemporary gender equality

We have shown that CW-index is associated with smaller male-to-female differences in ages at first marriage which, in turn, are positively correlated with contemporary female-to-male ratios in labor force participation and life expectancy.

Now, we estimate the reduced-form impact of the CW-index on those present day measures of gender equality. The reduced-form coefficient of the CW-index will be a composite of the effect of CW operating through reduced sex differences in marital ages *plus* all the other potential transmission channels that are not controlled for in our regression setup.

The results, displayed in Table 3.9, suggest that the reduced-form effect of the CW-index on the female-male labor force participation ratio (Panel A) is positive and robust in terms of statistical significance to the inclusion of other deep determinants. One standard deviation increase in the CW index is associated with a 0.34 standard deviations increase in the female to male labor force participation ratio.

The CW-index is also a significant positive correlate of contemporary female-male ratio in life expectancy (Table 3.9, Panel B). One standard deviation increase in the CW index is associated with a 0.26 standard deviations increase in the female to male life expectancy ratio.

Finally, consistent with the lack of correlation between ages at first marriage and the gender ratio in years of education, the reduced-form coefficient of the CW-index is small and statistically insignificant (Table 3.9, Panel C). The only robust negative deep determinant of gender equality in education is years of agriculture.

---

<sup>45</sup>Once again, the results are robust to ancestry-adjusting the CW index and years of agriculture.

TABLE 3.9: Determinants of gender gaps: reduced form estimates

	Average female-male ratio in 1990–2010							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Labor force participation								
<i>Deep determinants:</i>								
Cool water	0.626*** (0.178)	0.447*** (0.165)	0.715*** (0.195)	0.480** (0.185)	0.612*** (0.175)	0.310* (0.171)	0.578*** (0.199)	0.492*** (0.182)
Years of agriculture		-0.055*** (0.011)			-0.048*** (0.011)	-0.054*** (0.011)		-0.048*** (0.011)
Plow			-0.190*** (0.062)		-0.188*** (0.056)		-0.181*** (0.055)	-0.170*** (0.054)
Agricultural suitability				0.149** (0.067)		0.136** (0.062)	0.136** (0.059)	0.104* (0.055)
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemporary controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	156	146	156	156	146	146	156	146
<i>R</i> <sup>2</sup>	0.282	0.452	0.332	0.309	0.493	0.471	0.354	0.504
Panel B: Life expectancy								
<i>Deep determinants:</i>								
Cool water	0.073*** (0.025)	0.060** (0.025)	0.078*** (0.026)	0.061** (0.027)	0.073*** (0.025)	0.052* (0.028)	0.066** (0.029)	0.068** (0.028)
Years of agriculture		-0.005*** (0.002)			-0.005** (0.002)	-0.005*** (0.002)		-0.005** (0.002)
Plow			-0.011 (0.011)		-0.015 (0.011)		-0.011 (0.010)	-0.015 (0.011)
Agricultural suitability				0.012 (0.010)		0.007 (0.012)	0.011 (0.010)	0.004 (0.011)
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemporary controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	158	147	158	158	147	147	158	147
<i>R</i> <sup>2</sup>	0.408	0.466	0.413	0.413	0.474	0.468	0.418	0.475
Panel C: Years of education								
<i>Deep determinants:</i>								
Cool water	0.245 (0.205)	0.114 (0.200)	0.255 (0.190)	0.254 (0.198)	0.087 (0.189)	0.119 (0.195)	0.265 (0.179)	0.085 (0.175)
Years of agriculture		-0.029*** (0.008)			-0.030*** (0.009)	-0.029*** (0.008)		-0.030*** (0.009)
Plow			-0.017 (0.073)		0.029 (0.085)		-0.017 (0.074)	0.029 (0.087)
Agricultural suitability				-0.009 (0.057)		-0.004 (0.057)	-0.010 (0.056)	0.001 (0.059)
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemporary controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	130	127	130	130	127	127	130	127
<i>R</i> <sup>2</sup>	0.581	0.618	0.582	0.581	0.619	0.618	0.582	0.619

Notes: OLS estimates are reported with robust standard errors in parentheses. “Cool water” is the cool water index described in section 3.3. “Years of agriculture” is the number of years (in thousands) since the Neolithic revolution (from 1500) from Putterman and Trainor (2006). “Plow” is the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013). *Historical controls* are: ancestral domestication of large animals, ancestral settlement patterns, and ancestral political complexity from Alesina et al. (2013). *Contemporary controls* are the natural log of per capita income and its square, measured in the same time period as the dependent variable. Continent dummies are included. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

### 3.6 Conclusion

Reviewing the burgeoning literature on the remote historic drivers of gender inequality, we presented evidence for an overlooked trajectory that (1) originates in the CW-condition, from where the path leads to (2) late female marriages in preindustrial times, which eventually pave the way towards (3) various gender egalitarian outcomes today.

In theorizing this evidence, we argue that the CW-condition embodies opportunity endowments that significantly reduced fertility pressures on women, which favored late female marriages in the preindustrial era. The resulting family and household patterns placed women into a better position to struggle for more gender equality during subsequent economic transitions toward the industrial and post-industrial stages of development. Hence, enduring territorial differences in the CW-condition predict preindustrial female marriage ages, which in turn predict gender equality today.

Our theory is compatible with and actually integrates several separate theories on the historic origins of gender (in)equality. First, the argument that scarcity in arable land favored historic gender inequality is incorporated, because the CW-condition explains the absence of such scarcity. Second, the argument that irrigation dependence favored historic gender inequality is incorporated, because the CW-condition explains the absence of such dependence. Third, the argument that disease prevalence favored historic gender inequality is incorporated, because the CW-condition explains the absence of such prevalence. Fourth, the fact that a longer lasting agrarian legacy explains preferences for obedience only until a certain temporal threshold is explained by the theory, because this threshold is located at the time when the CW-areas in Northwestern Europe adopted agriculture. Fifth, the argument that European descent favored historic gender equality is incorporated, because European descent is linked to historic gender equality only in CW-areas but not outside them. In conclusion, we suggest that our theory of female emancipation provides a credible umbrella in unifying previous theories of gender equality.

### 3.A Appendix

#### Data

For a list of the variables used in this chapter, some descriptive statistics, a short description, and original sources see Table 3.A.1.

Here, we extend the discussion on the CW-index from section 3.3. A previous version of this index (Welzel, 2014) also included the coastline share of a country's borders as a proxy for temperate maritime climates. However, this measure neglects the orientation of the coast, which due to prevailing winds, is determinant for the existence of a temperate climate at non-tropical latitudes. Second, for the purposes of this article, access to sea proxies for many other effects unrelated to cool water (e.g., trade access, fisheries) and could therefore confound the interpretation of the CW effect. In any case, the correlation coefficients between the two versions of the index are 0.96 (ancestry-unadjusted) and 0.98 (ancestry-adjusted). Figure 3.A.3 plots the current version of the CW index against the previous one. All the econometric results are robust to the inclusion of the coastline variable—coastline as a share of a country's borders—as an additional control variable.

A key concern with our CW-index relates to differences in country area size. Indeed, scores on the CW-index might not be comparable across countries with different area sizes when bigger size implies higher within-country variability in the CW-condition. For instance, both Australia and Italy have similar CW-scores: 0.54 and 0.59, respectively. But in the case of Australia, the score refers to a country that is about 25-times larger than Italy. Accordingly, one would assume that the same CW-score glosses over a much bigger within-country CW-variation in Australia than in Italy. If so, the two CW-scores would appear to be inequivalent, despite the fact that they are numerically similar.

To examine this issue, we use a dataset from the Peace Research Institute (PRIO) in Oslo whose observational units are spatial “grid cells” (Tollefsen et al., 2012). The size of these grid cells approximates 55 by 55 kilometers at the equator. The inhabited grid cells of today's country-territories amount to 64,818 in number. We can roughly replicate our CW-index by average temperature measures and indications of the occurrence of droughts on the grid cell level. The measure of the CW-condition is less detailed than the one we use at the country level. If we nevertheless find that the two measures correlate strongly, we have assurance of the original measure's validity. This is indeed what we find: there is an almost 70 percent match between our original measure of the CW-condition and aggregations of the CW-condition from grid cell data.

The grid cell data allow us to estimate within-country variation in the CW-condition.

To do so, we examine the standard deviations around given country averages and the coefficients of variance, which express the ratio of the standard deviation to the mean. Doing so yields surprising findings.

First, only 14 percent of the variance in the CW-condition across the globe's roughly 65,000 inhabited grid cells represents differences within countries. By the same token, fully 86 percent of the CW-variance derives from differences between countries. Thus, country averages in the CW-condition are significant and meaningful because they depict by far most of the territorial variation in the CW-condition.

Second, territorial country-size has no influence whatsoever on within-country CW-variation. Hence, the suspicion that the 0.54 CW-score of Australia is incomparable to Italy's 0.59 CW-score because Australia's score supposedly hides much more CW-variation than Italy's is mistaken. Indeed, the coefficient of variance for Italy's CW-condition is 0.09, which is even marginally larger than Australia's 0.08. Additional examples illustrate the point: variation of the CW-condition in Canada is not larger than in Slovakia (both at 0.04); likewise, variation of the CW-condition in China is not bigger than in Panama (both at 0.09); most strikingly, variation of the CW-condition in Russia is not bigger than in Jordan (both at 0.04). In conclusion, the concern that differences in country area size make CW-scores incomparable across countries dissolves.

### Additional Tables and Figures

TABLE 3.A.1: Description of variables used and their source

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
<i>Average female-male ratio in 1990–2010:</i>							
Labor force participation	0.68	(0.21)	0.15	1.01	191	Female / male: % of ages 25-59 in labor force.	ILO Laborsta EAP6 <sup>th</sup> Revision (2011)
Life expectancy	1.07	(0.04)	0.99	1.21	202	Female / male: life expectancy at birth.	World Development Indicators
Years of education	0.82	(0.22)	0.21	1.41	146	Female / male: mean years of schooling, ages 25+.	Barro and Lee (2013)
Female SMAM	21.96	(2.88)	15.56	32.19	214	Female singulate mean age at first marriage.	UN (2009)
Male SMAM	26.13	(2.28)	21.13	34.49	209	Male singulate mean age at first marriage.	UN (2009)

*Continued on next page*



TABLE 3.A.1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
Year of obs.	1975.97	(9.26)	1960	2006	214	Year of earliest data point of female SMAM for each country.	UN (2009)
Historical female SMAM (Europe)	-2.07	(3.1)	-6.81	2.36	28	Female age at first marriage; country-specific coefficient from regression adjusting for data source characteristics. Europe only; reference country is England; data period is 1500-1900.	Dennison and Ogilvie (2014, Table 2)
Historical female SMAM (World)	21.52	(4.57)	12.6	28	27	Female singulate mean age at first marriage; earliest data point for the period 1800-1900.	Gapminder
Cool water	0.48	(0.15)	0.21	0.83	183	Cool water index; see section 3.3 for details.	Welzel (2013, 2014)
Cool water, ancestry-adjusted	0.47	(0.15)	0.01	0.77	165	Cool water index, ancestry-adjusted using the post-1500 migration matrix from Putterman and Weil (2010).	
Years of agriculture	4.31	(2.42)	0	10	165	Thousands of years from 1500 C.E. since the Neolithic revolution.	Putterman and Trainor (2006)
Years of agriculture, ancestry-adjusted	4.79	(2.23)	0.06	9.9	165	Years of agriculture, ancestry-adjusted using the post-1500 migration matrix from Putterman and Weil (2010).	
Plow	0.48	(0.48)	0	1	227	Share of a country's population with ancestors that practiced plow agriculture.	Alesina et al. (2013)

*Continued on next page*

TABLE 3.A.1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
Agricultural suitability	0.54	(0.33)	0	0.98	214	Share of ancestral land suitable for growing barley, wheat, sorghum, rye, foxtail millet, or pearl millet.	Alesina et al. (2013)
Large animals	0.93	(0.21)	0	1	227	Share of a country's population with ancestral domestication of large animals.	Alesina et al. (2013)
Political hierarchies	3.3	(1.04)	1	5	227	Ancestral number of political jurisdictional hierarchies (1-5) beyond the local community.	Alesina et al. (2013)
Economic complexity	6.38	(1.38)	1	8	227	Ancestral economic development based on 8 settlement patterns: from nomadic or fully migratory to complex settlements.	Alesina et al. (2013)
Tropical climate	0.74	(0.42)	0	1	211	Share of ancestral land that was tropical or subtropical.	Alesina et al. (2013)
$F_{ST}$ from U.K. (weighted)	0.09	(0.07)	0	0.23	179	Expected genetic distance between a randomly chosen individual from a given country and a randomly chosen individual from the U.K., using the genetic distances of their respective ancestor populations.	Spolaore and Wacziarg (2009)
Population density in 1500	9.09	(14.41)	0	100.67	186	Estimated population per squared kilometer in 1500.	Klein Goldewijk et al. (2010)
<i>Years of schooling in 1950 (log):</i>							
Total	0.5	(1.2)	-4.44	2.19	146	Log of mean years of schooling, ages 25+.	Barro and Lee (2013)

*Continued on next page*

TABLE 3.A.1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
Male	0.73	(1.09)	-4.4	2.21	146	Log of male mean years of schooling, ages 25+.	Barro and Lee (2013)
Female	0.06	(1.6)	-5.24	2.18	146	Log of female mean years of schooling, ages 25+.	Barro and Lee (2013)
Polity2 in 1980	-1.86	(7.51)	-10	10	142	Democracy score on a 10 point scale: from -10 (hereditary monarchy) to +10 (consolidated democracy).	Center for Systemic Peace
Rule of law in 2000	-0.03	(1)	-2.31	1.94	201	Rule of law perceptions, measured in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.	Kaufmann et al. (2011)
<i>Religious shares in 1980:</i>							
Catholic	34	(37.03)	0	99.10	152	% of Catholics in total population.	La Porta et al. (1999)
Protestant	12.59	(20.9)	0	97.8	152	% of Protestants in total population.	La Porta et al. (1999)
Muslim	22.87	(35.77)	0	99.90	152	% of Muslims in total population.	La Porta et al. (1999)
Oil production (per capita)	0.04	(0.16)	0	1.36	186	Barrels produced per person per day in 2000	Alesina et al. (2013)
Female ownership	35.74	(16.4)	2.8	86.8	131	% of firms in the World Bank Enterprise Surveys with some female ownership. The surveys were conducted between 2003 and 2010, depending on the country.	Alesina et al. (2013)
Women in politics	11.96	(9.02)	0	43	156	% of seats in parliament held by women in 2000.	Alesina et al. (2013)
GDI in 2014	0.93	(0.07)	0.6	1.03	161	Gender Development Index in 2014	UNDP (2015)

*Continued on next page*

TABLE 3.A.1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
Intensity agriculture	0.52	(0.46)	0	1	227	Share of country's population with ancestors practicing intensive agriculture or intensive irrigated agriculture.	Alesina et al. (2013)
Subsistence share from husbandry	0.24	(0.16)	0.03	0.92	227	Herding or large animals as a proportion of all ancestral subsistence activities.	Alesina et al. (2013)
Subsistence share from hunting	0.05	(0.05)	0.03	0.31	227	Hunting as a proportion of all ancestral subsistence activities.	Alesina et al. (2013)
Absence of land inheritance rules	0.1	(0.25)	0	1	215	Share of country's population with ancestral absence of inheritance rights of land.	Alesina et al. (2013)
Years of civil conflict (1816-2007)	6.98	(13.67)	0	105	192	Number of years country was involved in civil conflict from 1816-2007. Original source: <i>Correlates of War Database</i> version 4.	Alesina et al. (2013)
Years of interstate conflict (1816-2007)	3.86	(7.69)	0	41	192	Number of years country was involved in interstate conflict from 1816-2007. Original source: <i>Correlates of War Database</i> version 4.	Alesina et al. (2013)
Fraction of European descent	32.56	(41.81)	0	100	164	% of country's population in 2000 with European ancestry.	Alesina et al. (2013)
Rugged	1.38	(1.38)	0	7.81	227	Terrain Ruggedness Index from Nunn and Puga (2012)	Alesina et al. (2013)
Communist dummy	0.24	(0.43)	0	1	199	= 1 if country was formerly communist, and 0 otherwise.	Alesina et al. (2013)

*Continued on next page*

TABLE 3.A.1 – Continued from previous page

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
<i>Share of GDP in 2000:</i>							
Agriculture	16.46	(14.98)	0.11	72.01	170	Measured in %. Originally from the World Bank's World Development Indicators.	Alesina et al. (2013)
Manufacturing	14.29	(7.94)	0.91	39.5	167	Measured in %. Originally from the World Bank's World Development Indicators.	Alesina et al. (2013)
Services	53.45	(14.99)	4.25	81.10	169	Measured in %. Originally from the World Bank's World Development Indicators.	Alesina et al. (2013)

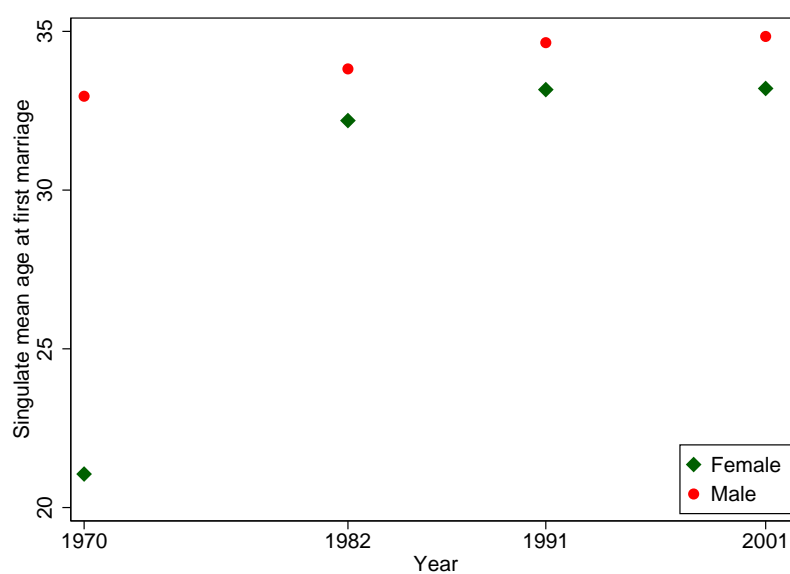


FIG. 3.A.1: Jamaica: ages at first marriage. 1970 is an outlier

Sources: UN (2009).

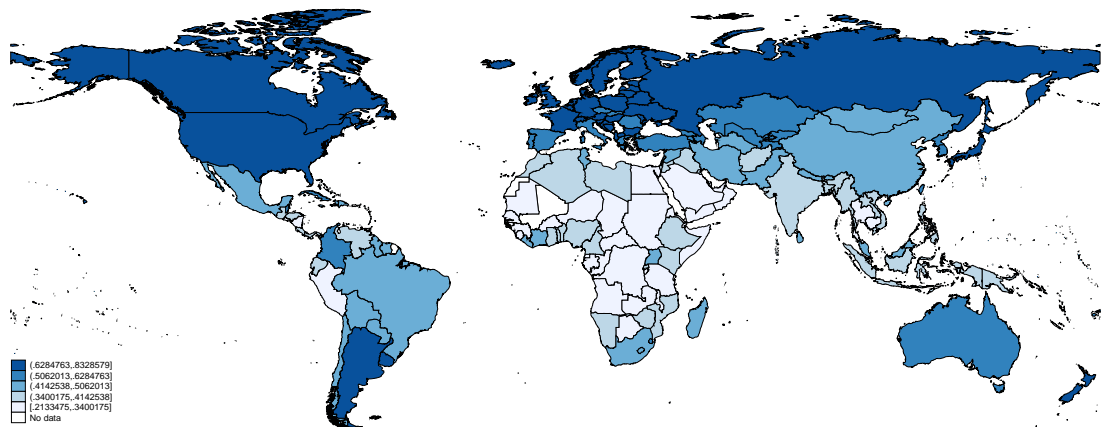


FIG. 3.A.2: World distribution of the CW-condition

TABLE 3.A.2: Additional gender equality outcomes and ages at first marriage

	Share of firms with female ownership, 2003–2010			Share of political positions held by women in 2000			Gender Development Index in 2014		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>SMAM:</i>									
Female	2.765*** (1.024)			1.034* (0.524)			0.009*** (0.003)		
Male	-1.510 (1.182)			-1.284** (0.564)			-0.006* (0.004)		
Female/male		74.609*** (26.338)			30.177** (13.835)			0.263*** (0.081)	
Male–female			-2.334** (0.982)			-1.111** (0.493)			-0.009*** (0.003)
Year of obs.	-0.283 (0.201)	-0.206 (0.179)	-0.161 (0.179)	-0.045 (0.124)	-0.083 (0.111)	-0.068 (0.114)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)
<i>Deep determinants:</i>									
Years of agriculture	-1.011 (1.088)	-0.996 (1.074)	-1.078 (1.088)	-0.801 (0.556)	-0.737 (0.553)	-0.758 (0.549)	-0.012*** (0.002)	-0.012*** (0.002)	-0.012*** (0.003)
Plow	-11.415** (5.262)	-11.767** (5.179)	-12.579** (5.142)	-4.124* (2.281)	-3.785 (2.300)	-3.939* (2.291)	-0.009 (0.013)	-0.010 (0.013)	-0.012 (0.013)
Agricultural suitability	-3.289 (6.086)	-2.373 (6.060)	-2.350 (6.220)	2.217 (2.853)	2.110 (2.871)	2.115 (2.866)	0.022 (0.015)	0.024* (0.014)	0.024* (0.014)
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemporary controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	106	106	106	125	125	125	136	136	136
<i>R</i> <sup>2</sup>	0.290	0.291	0.275	0.382	0.381	0.381	0.688	0.691	0.684
adj. <i>R</i> <sup>2</sup>	0.162	0.173	0.154	0.284	0.289	0.289	0.643	0.650	0.642

Notes: OLS estimates are reported with robust standard errors in parentheses. “Share of firms with female ownership” (in %) is taken from Alesina et al. (2013); originally from the World Bank Enterprise Surveys. “Share of political position held by women” is the percentage of women-held parliament seats, also taken from Alesina et al. (2013). “Gender Development Index” is from UNDP (2015). “Ages at first marriage” are singulate mean years at first marriage (SMAM) from UN (2009) for the period 1960–2000. For each country, the earliest year available is selected. “Years of agriculture” is the number of years (in thousands) since the Neolithic revolution (from 1500) from Putterman and Trainor (2006). “Plow” is the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013). “Agricultural suitability” and *Historical controls* are from Alesina et al. (2013). The natural log of per capita income and its square are measured in 2000. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE 3.A.3: Determinants of ages at first marriage: geo-climatic variables

	Singulate mean ages at marriage: male – female								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Abs. latitude	-2.233*** (0.773)			-3.156*** (0.889)					
Mild summers		-2.542*** (0.825)		-2.868*** (0.815)					
Continuous rain			-0.256 (0.734)	-0.861 (0.790)					
Coolness-component					-0.362 (0.930)		-2.267** (1.067)	-2.551 (2.024)	
Water-component						-2.376** (0.971)	-3.272*** (1.154)	-3.472* (1.821)	
Coolness $\times$ Water								0.535 (3.150)	
Cool water									-5.902*** (2.078)
Years of agriculture	0.269*** (0.082)	0.216** (0.085)	0.228** (0.092)	0.212** (0.092)	0.241*** (0.084)	0.167* (0.091)	0.143 (0.094)	0.144 (0.094)	0.154* (0.092)
Plow	0.612 (0.658)	-0.081 (0.594)	0.064 (0.645)	0.834 (0.643)	0.127 (0.691)	0.093 (0.601)	0.717 (0.689)	0.716 (0.692)	0.890 (0.686)
Agricultural suitability	0.512 (0.604)	-0.166 (0.536)	-0.082 (0.573)	0.664 (0.597)	0.002 (0.610)	-0.147 (0.541)	0.355 (0.634)	0.349 (0.639)	0.523 (0.646)
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemporary controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	121	121	121	121	121	121	121	121	121
R <sup>2</sup>	0.472	0.497	0.450	0.537	0.450	0.485	0.502	0.502	0.498
adj. R <sup>2</sup>	0.396	0.426	0.371	0.460	0.371	0.411	0.425	0.420	0.426

Notes: OLS estimates are reported with robust standard errors in parentheses. "Male – female SMAM" is the gender gap in SMAM data from UN (2009). "Abs. latitude" is the absolute latitude (in degrees) at a country's centroid. "Mild summers" is the inverse of the usual peak temperature (in degrees Celsius) in a country's hottest month of the year. "Continuous rain" is the square root of the typical rainfall (in cubic millimeters) in a country's driest month. The three previous variables are normalized to [0,1]. "Coolness and Water" are the two dimensions extracted from a factor analysis of the previous three variables; see section 3.3 for more details. "Cool water" is the cool water index described in section 3.3. "Years of agriculture" is the number of years (in thousands) since the Neolithic revolution (from 1500) from Putterman and Trainor (2006). "Plow" is the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013). *Historical controls* are: ancestral domestication of large animals, ancestral settlement patterns, and ancestral political complexity from Alesina et al. (2013). *Contemporary controls* are the natural log of per capita income and its square, measured in the same time period as the dependent variable. Continent dummies are included. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE 3.A.4: Determinants of ages at first marriage: ancestry-adjustment

	Singulate mean age at first marriage							
	Female		Male		Female/male		Male–female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Deep determinants:</i>								
Cool water, ancestry-adjusted	5.706** (2.729)	6.133** (2.964)	-2.081 (2.519)	-2.570 (2.744)	0.289*** (0.058)	0.331*** (0.074)	-7.837*** (1.634)	-8.938*** (2.049)
Years of agriculture, ancestry-adjusted		-0.142 (0.152)		-0.071 (0.117)		-0.003 (0.003)		0.066 (0.088)
Plow		-0.635 (0.911)		0.214 (0.704)		-0.034 (0.025)		0.917 (0.653)
Agricultural suitability		-0.221 (0.805)		0.166 (0.711)		-0.022 (0.021)		0.522 (0.560)
<i>Historical controls:</i>								
Large animals	0.278 (0.811)	0.260 (0.862)	-0.404 (0.769)	-0.367 (0.812)	0.024 (0.023)	0.020 (0.024)	-0.666 (0.628)	-0.575 (0.650)
Political hierarchies	0.442* (0.248)	0.575* (0.308)	0.155 (0.190)	0.110 (0.256)	0.010 (0.007)	0.018** (0.009)	-0.248 (0.178)	-0.450* (0.228)
Economic complexity	0.026 (0.111)	0.019 (0.112)	0.169 (0.105)	0.163 (0.105)	-0.003 (0.004)	-0.003 (0.004)	0.115 (0.111)	0.113 (0.103)
<i>Contemporary controls:</i>								
Income p.c. (log)	4.628** (2.306)	5.023** (2.412)	5.399*** (1.990)	5.567*** (2.108)	-0.010 (0.059)	-0.001 (0.059)	0.968 (1.602)	0.766 (1.600)
(Income p.c. (log)) <sup>2</sup>	-0.244* (0.132)	-0.264* (0.137)	-0.284** (0.114)	-0.292** (0.120)	0.001 (0.003)	0.000 (0.003)	-0.051 (0.092)	-0.040 (0.092)
Year of SMAM obs.	0.066*** (0.022)	0.067*** (0.022)	0.048** (0.021)	0.048** (0.021)	0.001** (0.001)	0.001** (0.001)	-0.027* (0.016)	-0.029* (0.016)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	125	125	121	121	121	121	121	121
R <sup>2</sup>	0.375	0.389	0.210	0.214	0.524	0.545	0.514	0.534
adj. R <sup>2</sup>	0.309	0.305	0.122	0.101	0.472	0.480	0.461	0.467

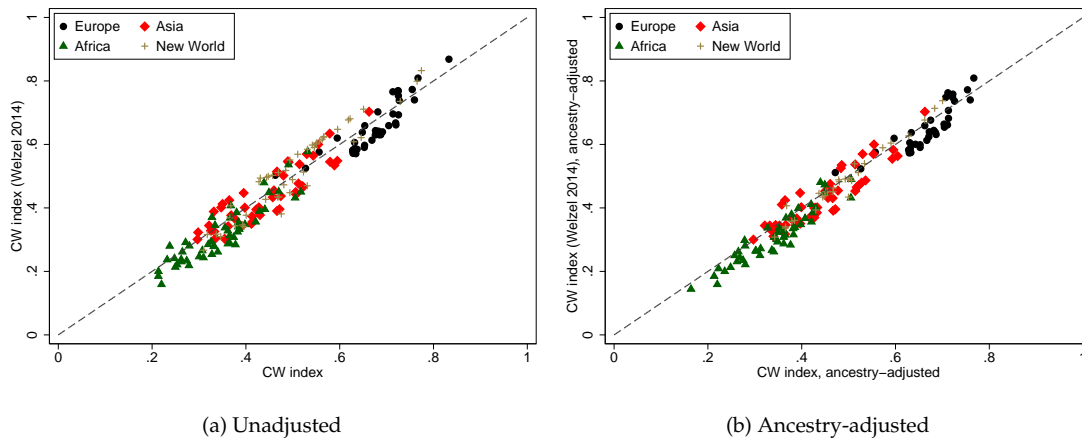
Notes: OLS estimates are reported with robust standard errors in parentheses. "Singulate mean age at first marriage" data are from UN(2009) for the period 1960–2006. For each country, earliest year available is selected and controlled for with variable "Year of SMAM obs." "Cool water, ancestry-adjusted" is the cool water index described in section 3.3 and "Years of agriculture, ancestry-adjusted" is the number of years (in thousands) since the Neolithic revolution (from 1500) from Putterman and Trainor (2006). Both variables are adjusted for a country's ancestral population using the post-1500 migration matrix of Putterman and Weil (2010). "Plow" is the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013). "Agricultural suitability" and *Historical controls* are from Alesina et al. (2013). The natural log of per capita income and its square are averaged over the period 1960–1980. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



**TABLE 3.A.5:** Determinants of ages at first marriage: excluding Northwest Europe and Western offshoots

	Singulate mean ages at first marriage: male–female				
	Full sample		W/o NW Europe	Full sample	W/o NW Europe and offshoots
	(1)	(2)	(3)	(4)	(5)
<i>Deep determinants:</i>					
Cool water	-5.902*** (2.078)	-5.640*** (2.148)	-5.672** (2.167)	-5.372** (2.225)	-5.534** (2.296)
Years of agriculture	0.154* (0.092)	0.153* (0.092)	0.157* (0.094)	0.155* (0.092)	0.161* (0.096)
Plow	0.890 (0.686)	0.852 (0.698)	0.847 (0.702)	0.902 (0.721)	0.916 (0.735)
Agricultural suitability	0.523 (0.646)	0.506 (0.652)	0.516 (0.656)	0.513 (0.656)	0.532 (0.656)
Northwest Europe		-0.421 (0.267)		-0.453* (0.268)	
Western offshoots				-0.568 (0.831)	
Historical controls	Yes	Yes	Yes	Yes	Yes
Contemporary controls	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes
<i>N</i>	121	121	112	121	108
<i>R</i> <sup>2</sup>	0.498	0.500	0.466	0.501	0.438
adj. <i>R</i> <sup>2</sup>	0.426	0.424	0.382	0.419	0.353

Notes: OLS estimates are reported with robust standard errors in parentheses. “Northwest Europe” is a dummy taking value 1 for Belgium, Denmark, France, Germany, Iceland, Ireland, Netherlands, Norway, Sweden, and the United Kingdom. “Western offshoots” is a dummy taking value 1 for Australia, Canada, New Zealand, and the United States. *Baseline historical controls* are: ancestral domestication of large animals, ancestral settlement patterns, and ancestral political complexity from Alesina et al. (2013). *Baseline contemporary controls* are the natural log of per capita income and its square averaged over the period 1960–1980, and the year of the SMAM observation. Continent dummies are included. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

**FIG. 3.A.3:** CW index comparison: with and without coastal borders

Notes: Comparison of CW index, as described in section 3.3, with its previous version from Welzel (2014), which included coastal borders as an additional variable. In the figures, the dashed line is the 45-degree line.

TABLE 3.A.6: Determinants of ages at first marriage: robustness to inclusion of additional controls

	Singular mean ages at first marriage: male–female									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cool water	-5.902*** (2.078)	-5.830*** (2.067)	-5.925*** (2.107)	-5.974*** (2.205)	-5.206** (2.057)	-5.301*** (1.959)	-6.074*** (2.188)	-5.912*** (2.040)	-6.296** (2.462)	-5.669** (2.498)
Years of agriculture	0.154* (0.092)	0.172* (0.101)	0.146 (0.092)	0.162 (0.091)	0.212** (0.099)	0.160* (0.091)	0.160* (0.092)	0.148 (0.091)	0.152 (0.093)	0.231** (0.113)
Plow	0.890 (0.686)	1.298* (0.686)	0.836 (0.715)	0.967 (0.686)	0.760 (0.691)	0.931 (0.681)	1.024 (0.710)	0.966 (0.692)	1.247* (0.733)	1.650** (0.777)
Agricultural suitability	0.523 (0.646)	0.355 (0.617)	0.548 (0.644)	0.636 (0.681)	0.586 (0.653)	0.536 (0.632)	0.616 (0.670)	0.608 (0.632)	0.797 (0.637)	1.024 (0.648)
<i>Historical controls:</i>										
Intensive agriculture		-1.223** (0.602)								-1.129* (0.630)
Subsistence share from husbandry		1.694 (1.689)								1.266 (2.072)
Subsistence share from hunting			-5.290 (5.064)							-4.287 (6.632)
Absence of land inheritance rules				0.813 (0.550)						1.423* (0.744)
<i>Contemporary controls:</i>										
Years of civil conflicts (1816-2007)					-0.013 (0.008)					-0.009 (0.009)
Years of interstate conflicts (1816-2007)					-0.016 (0.014)					-0.011 (0.017)
Terrain ruggedness index						-0.181 (0.113)				-0.175* (0.100)
Fraction of European descent							-0.004 (0.008)			0.005 (0.007)
Communist dummy								-0.759** (0.343)		-0.758** (0.362)
Agriculture share of GDP in 2000									0.021 (0.014)	0.024* (0.014)
Manufacturing share of GDP in 2000									-0.010 (0.020)	0.002 (0.024)
Services share of GDP in 2000									-0.002 (0.013)	-0.006 (0.014)
Baseline historical & contemporary controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	121	121	121	120	121	121	117	120	115	110
R <sup>2</sup>	0.498	0.522	0.505	0.504	0.518	0.510	0.501	0.519	0.537	0.617

Notes: OLS estimates are reported with robust standard errors in parentheses. "Male–female SMAVR" is the gender gap in SMAVR data from UN (2009). For a description of the additional control variables see Table 3.A.1. *Baseline historical & contemporary controls* are as in Table 3.4. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

TABLE 3.A.7: Gapminder: historical female ages at first marriage

	Historical female age at first marriage, 1801–1900							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Deep determinants:</i>								
Cool water	29.490*** (3.670)	28.621*** (4.202)	31.887*** (5.543)	30.966*** (6.484)	29.496*** (3.703)	28.378*** (4.149)	29.252*** (3.283)	28.391*** (3.760)
Years of agriculture			0.280 (0.321)	0.260 (0.346)				
Plow					1.402 (1.061)	1.634 (1.203)		
Agricultural suitability							0.280 (1.611)	0.272 (1.547)
<i>preindustrial development:</i>								
Population density in 1500		-0.019 (0.028)		-0.016 (0.028)		-0.024 (0.028)		-0.019 (0.029)
<i>N</i>	27	27	26	26	27	27	27	27
<i>R</i> <sup>2</sup>	0.761	0.765	0.747	0.750	0.768	0.774	0.761	0.765
adj. <i>R</i> <sup>2</sup>	0.752	0.745	0.725	0.716	0.749	0.745	0.741	0.734

Notes: OLS estimates are reported with robust standard errors in parentheses. "Historical female age at first marriage" data are country averages for the period 1801–1900 from Gapminder. Countries included: Armenia, Azerbaijan, Bangladesh, Belarus, China, Egypt, Estonia, Finland, Georgia, Germany, Iceland, India, Japan, Kazakhstan, Lithuania, Moldova, Netherlands, Norway, Pakistan, Russia, Spain, Sri Lanka, Sweden, Ukraine, United Kingdom, and United States of America. "Cool water" is the cool water index described in section 3.3. "Years of agriculture" is the number of years (in thousands) since the Neolithic revolution (from 1500) from Putterman and Trainor (2006). "Plow" is the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013). "Agricultural suitability" is from Alesina et al. (2013). "Population density in 1500" is from Klein Goldewijk et al. (2010).  
 \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



## 4 Can Parental Migration Reduce Petty Corruption in Education?\*

**Abstract:** The income generated from parental migration can increase funds available for children's education. In countries where informal payments to teachers are common, migration could therefore increase petty corruption in education. To test this hypothesis, we investigate the effect of migration on educational inputs. We use an instrumental variables approach on survey data and matched administrative records from the World Bank's Open Budget Initiative (BOOST) from Moldova, one of the countries with the highest emigration rates. Contrary to the positive income effect, we find that the strongest migration-related response in private education expenditure is a substantial decrease in informal payments to public school teachers. Any positive income effect due to migration must hence be overcompensated by some payment-reducing effects. We discuss a number of potential explanations at the family level, school level, or community level. We furthermore rule out several of these explanations and highlight possible interpretations for future research.

---

\*This chapter is co-authored with Lisa Höckel and Tobias Stöhr. For this volume, I made minor edits and updates on a previous version which is published as: Höckel, L., M. Santos Silva and T. Stöhr (2018), Can Parental Migration Reduce Petty Corruption in Education? *World Bank Economic Review*, 32(1):109–126. We are very grateful to the editors of that journal, three anonymous referees, Inga Afanasieva, Toman Barsbai, Julia Bredtmann, Elena Denisova-Schmidt, Iulian Gramatki, Artjoms Ivlevs, Stephan Klasen, Miquel Pellicer, Rainer Thiele, some unnamed experts, as well as participants at seminars at the University of Goettingen, the IOS Regensburg, the 2015 PEGNet conference, the 2016 AEL conference, and the 2016 NOVAFRICA Ph.D. Workshop for valuable comments.

## 4.1 Introduction

Emigration has long been considered detrimental to origin countries' human capital due to the loss of skilled workers. However, positive effects are possible either through the brain gain mechanism (Mountford, 1997) or due to a positive income effect increasing households' inputs in education.<sup>1</sup> That positive income effect could in theory also increase spending on a particularly corrosive education input—informal payments to teachers. Such payments are common in many developing countries and have also become widespread in post-Soviet countries after the collapse of the USSR as real wages for teachers declined abruptly. This paper shows that the positive income effect can be overcompensated by other channels leading to an overall decrease in informal payments to teachers due to parental migration.

These informal payments are problematic for two main reasons. First, they impose a “tax” on education that may reduce the incentives to human capital accumulation. Second, they distort performance incentives for teachers, parents, and students—for example, by motivating teachers to provide exam results to students, instead of teaching them in line with the curriculum. Thus, informal payments are understood to contribute to a less functional and less egalitarian public education system.<sup>2</sup> Often, they are raised by informal parental committees on a per capita basis and tend to be regressive. While some of the raised funds are spent on maintenance of the school, a large part will supplement wages of teachers. These payments have many organizational similarities to weakly enforced per capita taxes; a fact that can help tailor responses to them. The second and even more problematic form of payments to teachers is competition for higher grades or better treatment of individual students. Here, migrants can be expected to spend more money per child due to an income effect. These bribes are especially common in higher education (ESP/NEPC, 2010).

We study the effect of migration on informal payments and other forms of private educational expenditure and control for self-selection into migration by employing an instrumental variable approach. Our instrument is a network-based pull-effect at the local level, which is constructed using past migrant shares and destination-specific economic growth over time. The identifying assumption is that this network-growth interaction provides exogenous variation in the *ex-ante* costs and returns to migration, but does not otherwise affect the household's educational investment decision.

---

<sup>1</sup>E.g., Yang (2008); Bansak and Chezum (2009); Calero et al. (2009); McKenzie and Rapoport (2010); Antman (2011, 2012); Cortes (2015).

<sup>2</sup>E.g., Heyneman et al. (2008); Osipian (2009); ESP/NEPC (2010)

Our paper is, to our knowledge, the first to document a negative causal effect of parental migration on such informal payments to teachers. We show that the reduction in petty corruption occurs despite migrant households being, on average, wealthier than their non-migrant counterparts. This suggests that the income effect is overcompensated by other channels. School-level variation indicates strong spillovers within schools, which could partly be due to social remittances—i.e., migrants affecting the opinions of those left behind—, and partly due to migrant families' behavior leading to a breakdown of the social norm of taking part in petty corruption.

The results are neither explained by differences in public school funding nor by differences in the share of migrant children across schools. The money saved on informal payments to teachers does not translate into higher spending on out-of-school tutoring (henceforth: tutoring), which is an alternative way of teachers to make up for the loss of informal wage supplements. Rather, we find some evidence that, in migrant households, the main caregiver allocates more time to educational and school-related activities.

The reduction in informal payments may be explained by access to information or value change due to migration. There is a growing literature showing how the migration experience can alter migrants' and their left behind families' political values, social norms, and behavior in general.<sup>3</sup> Since the underlying preferences and beliefs about the spread of corruption are unobserved, this remains a tentative hypothesis. We are however able to rule out several alternative explanations: income-effects, the valuation of education, non-parental caregivers, and several supply side factors, which we measure using matched school budget data, community-level data as well as additional parts of the survey.

The remainder of the chapter proceeds as follows. Section 4.2 anchors the paper in the literature. Section 4.3 provides information on Moldova and corruption in education. Section 4.4 describes the data used, and section 4.5 presents our empirical strategy. The main results are discussed in section 4.6. Section 4.7 tests alternative explanations and the robustness of the main results. Section 4.8 concludes.

## 4.2 Related literature

Especially in developing countries, individual migration can be beneficial for children's education by raising and diversifying overall household income and alleviating credit constraints (Adams and Page, 2005; Calero et al., 2009). However, parental migration

---

<sup>3</sup>See, among others, the contributions of Spilimbergo (2009); Batista and Vicente (2011); Beine et al. (2013); Cameron et al. (2015); Tuccio and Wahba (2015); Barsbai et al. (2017); Ivlevs and King (2017).

can prove detrimental to children's educational achievement. First, parental absence can cause emotional distress jeopardizing school outcomes of children, especially if mothers or both parents are absent (e.g., Zhang et al., 2014; Cortes, 2015). Second, children could substitute for the absent migrant in household chores or even paid work (McKenzie and Rapoport, 2010; Antman, 2011). Third, parental migration might drastically reduce the educational input of the migrant's time.

Crucially, parents could try to make up for such negative effects by paying teachers informally to give their children extra attention, or even bribe them for better grades. In addition, we expect caregivers' time allocation to adjust, when family members migrate. In theory, the income effect could also decrease time allocated to children by remaining adults. However, parents often cite improving the lives of their children as the most important motive for migration. Therefore, we expect them to treat time spent with their child for educational activities as a normal or even luxury good. In this case, parents would invest more time if remittances allow them to work less. Hence, instead of consuming more leisure, we expect the remaining caregiver to increase time inputs in education.

But migration can affect households' educational investment more fundamentally. The preferences and views of immigrants are known to change through acculturation, personal experience, and the exposure to new ideas, knowledge, and institutions (Berry, 1997; Careja and Emmenegger, 2012). For example, the values of immigrants living in Western societies are found to converge to those of the host population over time. Such changed values can have a lasting effect when migrants return to their country of origin (Spilimbergo, 2009; Batista and Vicente, 2011).<sup>4</sup> These effects are not confined to return migration but can also be transmitted through communication with family or friends. Chauvet and Mercier (2014) find spillover effects from the migrant to the non-migrant population in terms of electoral participation and competitiveness. Barsbai et al. (2017) provide evidence that emigration from Moldova, the country studied in this chapter, changed political attitudes and may have resulted in the incumbent Communist government losing the 2009 elections. As the authors discuss, Moldova had very little exposure to the outside world before migration recently took off. In such settings, where information is scarce, diffusion processes are likely to be influential. As petty corruption often depends on the societal belief that it is widespread (Dong et al., 2012; Corbacho et al., 2016), migration may shatter this belief if migrants observe school systems that function without informal payments. In particular, payment schemes that depend on public-good-style contributions may dissolve even if only a few individuals cease

<sup>4</sup>See Docquier and Rapoport (2012) for an excellent discussion of the literature.



contributing (Fehr and Gächter, 2000).

### 4.3 Moldova and corruption in education

Moldova is the poorest country in Europe with an estimated GDP per capita (PPP-adjusted) of \$4,521 (World Bank, 2014).<sup>5</sup> The potential effects of migration and societal spill-overs are particularly visible in Moldova. It is the country with the third highest remittance to GDP ratio (24.9 percent), only surpassed by the Kyrgyz Republic and Nepal (World Bank, 2014). In comparison, other commonly studied economies like Mexico (remittances to GDP ratio of 2 percent) or the Philippines (9.8 percent) are considerably less dependent on remittances. Another advantage is that migration has been a relatively recent phenomenon. After the dissolution of the Soviet Union in 1991, some Moldovans continued working in what is now Ukraine and Russia and were thus suddenly called international migrants. Mass migration, however, only started after the Russian financial crisis of 1998, which increased unemployment and poverty considerably in Moldova. In 2011, emigrants made up 17 percent of the total population (MPC, 2013), meaning that 30–40 percent of children, depending on the sample, are affected by emigration of at least one parent.<sup>6</sup>

As a former member of the Soviet Union, Moldova's public educational system has good coverage (even in rural areas), with enrollment rates of nearly 100 percent for primary and lower secondary schooling, and 87 percent for upper secondary schooling (Table 4.A.1 in the chapter's appendix). Attendance is formally free of charge from first grade up to high school completion<sup>7</sup>, and below tertiary education there are few private schools.

There is a steep socio-economic gradient in educational achievement (Walker, 2011), which some worry might increase due to migration, not least due to widespread informal (and often illegal) payments to schoolteachers and other officials. The institutional causes of these are twofold: teachers' wages are low and often delayed and, socially, there is public tolerance of corruption and insufficient critical input of mass media. According to the 2013 Global Corruption Barometer, 37 percent of households in Moldova that came

<sup>5</sup>In 2013, countries with a comparable per capita GDP (in 2011 \$-PPP) were, for example, Pakistan (\$4,454), Nicaragua (\$4,493), and Lao (\$4,667).

<sup>6</sup>The most common emigration destination for circular migrants is Russia. While migration to Russia is usually characterized by short-term stays and manual labor, emigration to the West is more permanent, service-sector oriented, and feminized (60 percent women). Italy and Romania are particularly important destination countries due to linguistic proximity.

<sup>7</sup>Moldova has compulsory schooling until the end of lower secondary schooling (roughly age 15).

into contact with education authorities paid bribes in the 12 months before the survey and 58 percent of respondents perceived the education system to be corrupt or highly corrupt (Transparency International, 2013). Similarly, in the 2011 Citizen Report Card study, corruption is cited to be the most common difficulty when requiring services from public educational institutions and paying bribes is the second most common way of solving problems after insistence, joint with using personal contacts. Another form of corruption in the education system is the acquisition of unnecessary tutoring from a child's teacher (Carasciuc, 2001). This means that tutoring is often in a gray area between a productive investment in students' cognitive achievement and paying teachers informally. Besides seeking individual gains for one's own child, there is an important social component to making illicit payments to teachers resulting from the interaction of parents, teachers, and school principals (ESP/NEPC, 2010).<sup>8</sup>

The less frowned upon kind of these payments are monetary transfers or in-kind "gifts" that are often collected by informal parental committees. Typically, they either supplement teachers' wages or finance maintenance spending in schools. These expenditures face some of the organizational issues of public goods, including committees dissolving and payments stopping once the number of parents who are willing to contribute declines. There are only relatively blunt mechanisms to enforce payment—for example, parents being excluded from the committee, and teachers ignoring children in class. While payments can be seen as necessary to motivate teachers, there are widespread detrimental consequences, such as especially motivated teachers providing solutions to (standardized) exams—a practice that clearly undermines the education system.<sup>9</sup> Furthermore, monetary transfers that are imposed on a per capita basis might also affect poor households disproportionately since they have to pay a higher share of their income (Emran et al., 2013).

The form of corruption in schools that is locally perceived as most problematic is direct bribing with the purpose of increasing the attention or grades a teacher gives to an individual student at the expense of others. Bribes can also be necessary to gain access to the best public high schools and to universities.<sup>10</sup>

---

<sup>8</sup>ESP/NEPC (2010) describes results from in-depth interviews on informal payments in 7 ex-communist countries. In that study, a majority of Moldovan parents reports being pressured by both teachers and other parents to comply with informal payments.

<sup>9</sup>This problem was so widespread that some time after our survey took place the education minister introduced video surveillance during the final high school exam, a move that led to a spike in failure rates. Something similar has recently been studied in Romania; see Borcan et al. (2017).

<sup>10</sup>Heyneman et al. (2008), for example, discuss survey data which indicate that about 80 percent of university students in Moldova, Bulgaria and Serbia were aware of illegal bribe paying in university admission.

In sum, while payments to teachers are in part motivated by grade-buying or seeking better treatment for the child, a larger share seems to operate as a per capita tax. In the latter case, the extent and magnitude of informal payments is more likely to be determined by local norms, the preferences and the bargaining power of teachers, parents, and school officials, and less by the pursuit of inflated grades or preferential treatment for the child. Both kinds of petty corruption, however, can be expected to affect incentives negatively, increase the socio-economic gradient in educational outcomes, and contribute to a social climate where corruption is an everyday experience.

## 4.4 Data and descriptives

In this section, we discuss the data and present key descriptive statistics of our sample.

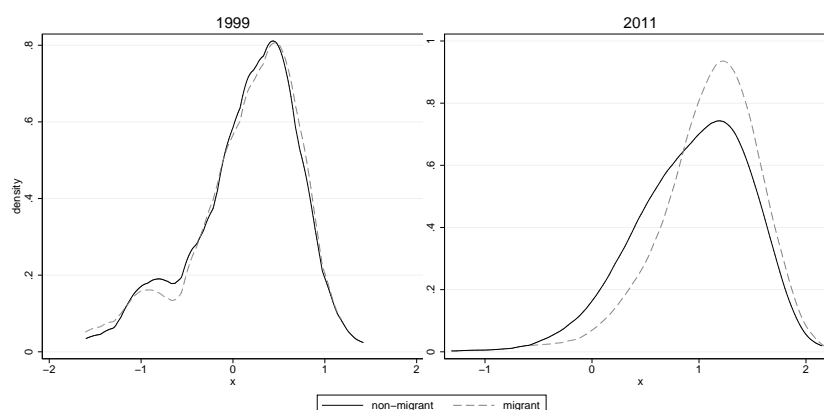
### Data

We use data from a nationally representative household survey conducted in Moldova in 2011–12 (henceforth abbreviated CELB 2012) which was specifically designed to investigate the effects of migration on children and elderly left behind. The survey includes 3,568 household with 12,333 individuals, of which 2,501 are children of age 6–18.<sup>11</sup> In addition to socio-economic characteristics of household members, detailed information on private financial and non-financial inputs into children's education was collected by identifying and interviewing each child's main caregiver.<sup>12</sup> Financial expenditures include payments and other "gifts" to schoolteachers, tutoring expenditures, and transportation expenditures that we will use as different dependent variables in the analysis.<sup>13</sup> Non-financial inputs include how often the main caregiver helps the child with homework and other school activities in the month prior to the survey interview on a six point scale ranging from "never" to "every day". In addition to the household survey, community questionnaires were filled out by local officials, typically in the mayoral office. Finally, we match data from the World Bank's open budget initiative (BOOST) to provide school-level data on public education expenses in the respective communities and schools (see appendix 4.A for more details). Our baseline sample consists of 2,148 children from 1,448 households.

<sup>11</sup>The response rate was above 80 percent. For detailed information on the survey see Böhme and Stöhr (2014); Böhme et al. (2015).

<sup>12</sup>The main caregiver is the person responsible for nutrition, health, and schooling of a child at the time of the survey.

<sup>13</sup>In addition, there is a residual category of "other expenditure" for which we find statistically insignificant effects.



**FIG. 4.1:** Kernel density plots of the household asset index in 1999 and 2011

*Notes:* Authors' calculations based on CELB 2012.

### Descriptive statistics

A migrant household is defined by the existence of at least one adult who, in the 12 months prior to the survey, has spent a minimum of three months living abroad. In our sample, 29 percent of children live in a migrant household (Table 4.1).<sup>14</sup> The average student from migrant households is 12.6 years old, 5 months more than her non-migrant peer. Before accounting for selection into migration, the average grade (GPA) is 0.06 points higher for children in migrant households. Migrant families are slightly larger on average and more likely to come from rural areas. Despite this, their average total income and average per capita income are significantly higher than those of non-migrants.<sup>15</sup> Figure 4.1 also reflects the underlying effect of migration, showing no difference in assets in 1999 but significantly higher assets for migrant families in 2011.<sup>16</sup>

Households in our sample report positive payments to teachers for about 37 percent of all school-age children.<sup>17</sup> Payments to teachers typically vary from 5 to 40 USD per child

<sup>14</sup>Our dataset does not allow us to compare the differences between migrant households with and without children. Comparisons to other representative data (details on request) reveal that in households with children, female migration is on average less common. The education level and gender composition do not differ markedly.

<sup>15</sup>In reality, the difference could be even wider, since migrant households systematically under-report their received remittances and other sources of income (Akee and Kapur, 2012).

<sup>16</sup>The asset indexes were constructed by a weighted-sum of the following items: number of cars, motorcycles, bicycles, washing machines, refrigerators, radios, TVs, computers, and cell phones; existence of working phone landline and internet access; and number of rooms in the house. For 1999, the last three items were excluded due to a large number of missing values. The weights for the index were obtained from a principal component analysis of the asset list. Dividing the divisible assets by the squared root of household size as an equivalent scaling rule does not change Figure 4.1 in any qualitative way.

<sup>17</sup>This figure is remarkably similar to the one reported in the 2013 Global Corruption Barometer: 37 percent of households in Moldova that came into contact with education authorities paid bribes in the 12

TABLE 4.1: Selected summary statistics

	Non-migrant Households			Migrant Households			Mean diff
	N	Mean	(SD)	N	Mean	(SD)	(t-test)
<i>Child characteristics</i>							
Age	1,783	12.28	(3.73)	718	12.68	(3.79)	**
Male	1,783	0.51		718	0.51		
GPA (0-10)	1,355	8.04	(1.07)	555	8.10	(0.93)	*
Serious illness (past year)	1,783	0.29		718	0.26		
Distance to school (min)	1,659	20.76	(18.39)	668	19.92	(17.53)	
<i>Household characteristics</i>							
Total income	1,783	33,819.11	(36,592.44)	718	48,901.40	(49,005.71)	***
Household size	1,783	4.70	(1.39)	718	5.13	(1.75)	***
Mean years education	1,782	10.74	(2.40)	718	10.68	(1.93)	
Urban	1,783	0.24		718	0.15		***
Older siblings	1,783	0.59		718	0.58		
Parents divorced	1,783	0.12		718	0.10		
<i>Private education inputs</i>							
Caregiver time	1,565	3.78	(1.94)	640	3.62	(1.97)	*
Payments to teachers	1,552	89.09	(275.56)	635	65.62	(163.85)	**
Out-of-school tutoring	1,572	192.70	(1,179.89)	642	86.57	(376.58)	***
Transportation	1,565	202.85	(775.61)	644	209.29	(902.19)	
<i>Network-Growth IV</i>							
	Communities	Mean	(SD)	Min	Max		
	129	277.66	(140.22)	2.33	691.70		

Notes: Authors' calculations based on CELB 2012. All monetary values are expressed in Moldovan Lei. GPA stands for Grade Point Average. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively.

per year, which is substantial given that public expenditure for teaching materials per pupil is about 30 USD per year, and wage bills per pupil are about 300 USD per year (c.f. appendix 4.A and Table 4.A.2). In contrast, households only report tutoring expenses for approximately 10 percent of children (c.f. Figure 4.A.1). Despite higher income, both per child informal payments to teachers and tutoring expenses are significantly lower in migrant households compared to non-migrant ones. For transportation expenditure, there is no such difference. The differences in informal payments and tutoring are mostly driven by more migrant households reporting zero payments (not refusals or “don’t know” answers), rather than by smaller positive expenses. This is not only evident at the individual level, but also results in a strong negative correlation at the community level between the share of migrant households and the share of respondents reporting payments to teachers (Table 4.2: Panel A, column 1).<sup>18</sup> The slope of the regression line is approximately  $-0.4$ , a very high value that is statistically and economically significant.

months before the survey (Transparency International, 2013). We focus on the likelihood of paying informal fees rather than the values paid since we assume the decision to participate in the informal fee scheme to be the most affected by a change in preferences. Note that we added 1 LCU to each private expenditure to ensure that the log exists.

<sup>18</sup>See Figure 4.A.2 for an illustration.

Note, though, that our data are designed to be representative at the national but not at the community level. The negative correlation also holds at the individual level (Table 4.2: Panel A, columns 2–5).

## 4.5 Empirical strategy

To analyze whether this strong negative correlation between migration and petty corruption at the community and individual level is indeed closely tied to migration, we estimate the stylized model:

$$y_{ihcs} = \alpha + \delta \text{Mig}_{hc} + X'_{ihcs} \beta + \epsilon_{ihcs} \quad (4.1)$$

where  $y_{ihcs}$  are private inputs to the education of child  $i$  in household  $h$  from community  $c$  and school  $s$ . We consider three financial inputs (informal payments to teachers, tutoring, and transport expenditures) and two non-financial inputs (whether the child is enrolled in school and the frequency with which the caregiver spends time supporting the child in educational activities). The main explanatory variable of interest,  $\text{Mig}_{hc}$  is a household-level dummy variable taking the value one if the child lives in a migrant household and zero otherwise;  $X_{ihcs}$  is a vector of child- and household-level control variables;  $\epsilon_{ihcs}$  is the error term.

Clearly, migrants are not a random population group but rather self-select into migration. Thus, it can be expected that they systematically exhibit distinct unobservable characteristics relative to non-migrants that might bias OLS estimates of equation (4.1). To overcome this problem, we estimate an instrumental variable approach by two-stage least squares (2SLS).<sup>19</sup> Our instrument for migration status is the interaction between preexisting migration networks at the local level and destination-specific economic conditions. Formally, we use the growth rate of per capita GDP for each destination country between 2004–2010 and weight it with the share of migrants that, by 2004, had migrated from the community to that destination.<sup>20</sup> The data for the migrant-destination share at

<sup>19</sup>The most common approach in the literature are instrumental variable strategies exploiting exogenous aggregate factors at the origin or destination: past migration rates (McKenzie and Rapoport, 2010; Antman, 2011; Zhang et al., 2014), financial infrastructure (Calero et al., 2009), and political unrest (Bansak and Chezum, 2009) at the origin level; employment conditions (Antman, 2011; Cortes, 2015) and exchange rate crises (Yang, 2008) at the destination level.

<sup>20</sup>Analytically:

$$\text{Network-Growth}_c = \sum_{j=1}^J \left( \frac{\text{migrants}_{c,j,2004}}{\text{population}_{c,2004}} \sum_{t=1}^T \left( \frac{\text{GDP}_{j,t+1} - \text{GDP}_{j,t}}{\text{GDP}_{j,t}} \right) \right)$$

where  $c$  is the Moldovan community,  $j = 1, 2, 3, \dots, J$  is the migration destination country, and  $t =$

the community level are derived from the 2004 Moldovan Census.<sup>21</sup> The variable has already been employed as an instrument for migration in other studies of the Moldovan context (e.g. Lücke et al., 2012; Böhme et al., 2015).

The rationale behind the use of *Network-Growth* is twofold. First, migrant networks are known to be very important in facilitating *current* migration. The network can provide *ex ante* information and assistance and *ex post* support for the migrant upon arrival (e.g., short-term accommodation, job-searching expertise, paperwork). Thus, pre-existent migrant networks effectively reduce the costs of migration (e.g., McKenzie and Rapoport, 2010). Secondly, the growth of GDP per capita at the destination is a proxy for the country's economic performance and, more importantly, employment conditions that are exogenous to potential migrants in Moldova. An expanding job market is highly attractive for potential migrants and hence a pull factor to this destination (e.g., Antman, 2011).<sup>22</sup>

As a whole, our instrument captures the exogenous variation of migrant networks at the community level, which lowers migration costs, and economic conditions at the destination country, which increases the expected returns of migration. Exploiting variation at the community level, our instrument does not allow exogenizing household-level choices regarding migration, such as the identity of the migrant or the duration of the stay abroad. We can only successfully predict the probability of at least one household member becoming a migrant. As a result, the household's migration status is the main variable of interest in our analysis. Our results should thus be interpreted as the average effects across all migrants and migratory spells.

The validity of the instrument depends on the exclusion restriction that *Network-Growth* must only affect the provision of private educational inputs through migration status. This seems self-evident for the growth of GDP per capita at the destination. It is hard to conceive of a different relationship (i.e., other than migration) through which the changes in per capita growth rates in a set of foreign countries would affect the education investment decisions of a Moldovan household differentially between

---

2004, 2005, ..., 2010 the year.

<sup>21</sup> An advantage of our setting is that migration has been a relatively recent phenomenon in Moldova and, thus, there is little scope for the non-migrant population to be influenced over time due to spillovers and long-term confounding developments. As a robustness check, we exclude from the analysis the migrant households which already had a migrant in 2004 or before, as they might be included in the Census migration rates. The main results do not change qualitatively (available upon request).

<sup>22</sup> To better capture the individual gains from migration, rather than the rise in opportunities, we alternatively use the change in GDP per capita. The results are comparable in magnitude and significance. The Kleibergen-Paap weak identification statistic is however smaller than when using GDP growth for the IV.

communities. For the migration network, we assume that past migration rates are predictors of current migration rates only via network effects and, otherwise, have no influence on the household's education spending. As in Böhme et al. (2015), we include the 2004 share of the community's population who is a migrant to Italy, Romania, Russia, and Ukraine as additional controls in the 2SLS setup. These variables account for proximity to the border and any systematic differences in development that may have emerge due to migration to any of these important destinations between the take-off of migration, in 1999, and the Census, in 2004.<sup>23</sup> Moreover, the IV is not systematically correlated with school expenditures, local economic conditions, as proxied by night lights (Henderson et al., 2012), and local infrastructure or public goods, as reported in the community questionnaire. Further, communities with IV values above and below the median are distributed evenly across the country (Figure 4.A.3). Summary statistics for the IV variable can be found at the bottom of Table 4.1.

## 4.6 Main results

The dependent variables of our empirical analysis are the child's school enrollment status, the three categories of private education spending—payments to teachers, tutoring expenses, transportation expenses—, and the time spent by the caregiver. The reduced form estimates are reported in Table 4.2, Panel B. The lack of a selection correction results in a statistically significant correlation between the instrument and school enrollment<sup>24</sup>, which indicates better migration options for those who leave school after the end of compulsory schooling. Correlations between the instrument and payments to teachers, as well as tutoring expenses, are negative and statistically significant.<sup>25</sup>

<sup>23</sup>Alternatively using only one control for all migrant shares does not yield different results but we prefer keeping to the more conservative ability to control also for different border effects as in that earlier paper.

<sup>24</sup>A one standard deviation increase in the instrument implies a 2.5 percentage point reduction in enrollment.

<sup>25</sup>See Tables 4.A.3, 4.A.4, and 4.A.5 for the point estimates of the control variables for Panels B and C.



**TABLE 4.2:** The effect of migration on private education inputs

**PANEL A**

*Basic OLS results*

	Payment to teachers				
	Community level	Individual level			
		(2)	(3)	(4)	(5)
Age group	(1) all	all	10+	15+	18+
Migration	-0.415*** (0.119)	-0.046 (0.028)	-0.055* (0.031)	-0.069* (0.040)	-0.131** (0.058)
<i>N</i>	129	2,287	1,764	898	330

**PANEL B**

*Reduced form estimates (OLS)*

	School enrollment	Payments to teachers		Out-of-school tutoring		Transportation expenditure		Caregiver time
	(1)	(2) log	(3) $D(Y > 0)$	(4) log	(5) $D(Y > 0)$	(6) log	(7) $D(Y > 0)$	(8)
Network-Growth	-0.000** (0.000)	-0.005** (0.002)	-0.001** (0.000)	-0.002*** (0.001)	-0.000*** (0.000)	0.000 (0.002)	0.000 (0.000)	0.003*** (0.001)
<i>Child characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Household characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Main migration destinations</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	2,223	2,148	2,148	2,170	2,170	2,168	2,168	2,162
$R^2$	0.064	0.068	0.042	0.095	0.084	0.173	0.168	0.265
F-stat	5.72	7.85	5.86	4.11	4.43	13.3	14.0	40.5

*Continued on next page*

TABLE 4.2 – Continued from previous page

**PANEL C***First stage IV regressions*

	Migration							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Instrument</i>								
Network-Growth	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
<i>Child characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Household characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Main migration destinations</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	2,223	2,148	2,148	2,170	2,170	2,168	2,168	2,162
<i>F Statistic</i>	9.6	10.6	10.6	11.0	11.0	10.7	10.7	10.6
<i>Second stage IV regressions</i>	School enrollment	Payments to teachers		Out-of-school tutoring		Transportation expenditure		Caregiver time
	(1)	(2) log	(3) $D(Y > 0)$	(4) log	(5) $D(Y > 0)$	(6) log	(7) $D(Y > 0)$	(8)
Migration	-0.129 (0.097)	-4.430*** (1.717)	-0.829** (0.332)	-1.987** (0.869)	-0.274** (0.126)	0.247 (1.905)	0.091 (0.304)	2.667* (1.449)
<i>Child characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Household characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Main migration destinations</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	2,223	2,148	2,148	2,170	2,170	2,168	2,168	2,162
K-P weakid	9.6	10.6	10.6	11	11	10.7	10.7	10.6
95% CLR confidence set	[-0.4, 0.06]	[-9.25, -1.88]	[-1.76, -0.32]	[-4.99, -0.01]	[-0.71, 0.02]	[-2.34, 2.82]	[-0.28, 0.49]	[0.99, 5.57]
CLR test p-value	0.17	0.00	0.00	0.05	0.07	0.83	0.60	0.00
Cluster-robust 95% AR conf.	[-0.51, 0.02]	[-9.12, -0.83]	[-1.68, -0.08]	[-5.12, -0.65]	[-0.72, -0.07]	[-3.45, 5.91]	[-0.47, 1.02]	[0.66, 8.35]
Cluster-robust AR p-value	0.11	0.03	0.04	0.01	0.01	0.90	0.76	0.01

Notes: Authors' calculations based on CELB 2012. Standard errors in parentheses. Panel A uses heteroskedasticity-robust standard errors throughout. Panels B and C use heteroskedasticity-robust standard errors that cluster at the community level. All models include a constant. *Child characteristics*: age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics*: mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations*: 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. In Panel A, column 1: migration indicates the share of migrant households in the community; the dependent variable is the community's share of respondents reporting positive informal payments to schoolteachers. Interpreting Panel A column 1, please note that the survey was not designed to be representative at the community level. Panel B reports the reduced form where the outcome of the second stage is regressed on the instrument (*Network Growth*) and the endogenous variable (migration) is excluded. Note that interpreting the size of the instrumental variable is not easy, because it is a sum of Network-Growth Interactions. Differences in missing values for the dependent variables explain the different number of observations across columns. Panel C shows the first and second stage regressions. Migration is instrumented using a network-growth interaction IV. K-P weakid is the Kleibergen-Paap weak identification statistic. The CLR test refers to confidence region and the test statistic using the "condivreg" package by Mikusheva and Poi (2006). The cluster-robust AR 95% confidence set is calculated using the "rivtest" package by Finlay and Magnusson (2009). Coefficients for all the control variables shown in supplemental appendix: reduced form (Panel B) in Table 4.A.3, first stage (Panel C) in Table 4.A.4, column 1, and second stage (Panel C) in Table 4.A.5.

The first stage IV estimates are reported in Panel C of Table 4.2. The *Network-Growth*-IV is a positive and highly significant predictor of the household's migration status. The instrument's estimated coefficient implies that a one standard deviation increase in *Network-Growth* increases the likelihood of (at least one) household adult member migrating by approximately 14 percentage points. The Kleibergen-Paap rank test rejects underidentification at least at the 5 percent significance level in all the 2SLS regressions.

The second stage indicates no statistically significant effect of migration on the enrollment probability (column 1) as a result of parental migration. Instead, the results indicate a strong reduction in the likelihood to pay teachers conditional on individual characteristics. The effect is even more pronounced than the negative correlation in panel A (column 3). For tutoring we find a similar negative effect, whereas transport expenditure remains unchanged (columns 5 and 7). Interestingly, the determinants of tutoring are similar to those of paying bribes, supporting the view that tutoring offers a "cleaner" way of making informal payments to teachers. There is some evidence that, in migrant households, caregivers spend more time on the education of their children (column 8). In order to account for potentially inflated point estimates due to weak IVs, we provide the conditional likelihood ratio (CLR) confidence region and cluster robust confidence sets for the respective migration effect at the bottom of the table (Mikusheva and Poi, 2006; Finlay and Magnusson, 2009; Moreira, 2009). Both methods show that the effect of migration on informal payments is bounded away from zero, even when accounting for weak IVs.<sup>26</sup> The results point to a statistically as well as economically significant negative effect of migration on informal payments.<sup>27</sup>

The very strong negative correlation, even after rigorously accounting for self-selection, cannot be the consequence of a mere income effect. At the same time, children's or parents' socio-economic characteristics do not predict petty corruption at the extensive margin very well. While there is more reporting of payments for older students, girls, and by more educated parents—one of the core predictors of income—the other controls are statistically insignificant. Additional analyses yield no evidence of heterogeneous treatment effects by age; yet, this is partly due to imprecise estimates in smaller subsamples (not shown).

<sup>26</sup>In addition, alternative estimates obtained from an IV probit estimation can be found in Table 4.A.6 for comparison.

<sup>27</sup>Table 4.A.7 presents OLS estimates for the same set of covariates. Due to the exclusion of a selection correction, covariates such as household size that are predictive of migration, but not of informal payments, pick up the correlation between migration and informal payments to teachers. The lack of a selection correction also results in statistically significant positive effects on transport expenditure, which are explained by higher available income in migrants households, as shown by additional results (available upon request).

**TABLE 4.3:** The effect of migration on private education inputs: controlling for household assets

<i>Second stage IV regressions</i>	<i>D(Payments to teachers) &gt; 0</i>		<i>Caregiver time</i>	
	(1)	(2)	(3)	(4)
Migration	-0.964** (0.434)	-0.636** (0.259)	3.887* (1.996)	2.665*** (1.003)
Household asset index (log)	0.212*** (0.080)		-0.417 (0.364)	
Household asset index 1999 (log)		0.031 (0.037)		0.151 (0.098)
<i>Child characteristics</i>	Yes	Yes	Yes	Yes
<i>Household characteristics</i>	Yes	Yes	Yes	Yes
<i>Main migration destinations</i>	Yes	Yes	Yes	Yes
N	2,186	1,708	2,354	1,824
K-P weakid	6.601	16.857	6.935	22.776

*Notes:* Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. See endnote 16 for a list of assets included in the asset index. *Child characteristics:* age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics:* mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations:* 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. K-P weakid refers to the Kleibergen-Paap weak identification statistic. Migration is instrumented using a network-growth interaction IV.

Our main results are not explained by differences in household wealth (proxied by a household asset index, Table 4.3). Contemporaneous assets are endogenous to migration and, in fact, constitute one of the main expected transmission channels for the effect of migration on education inputs (columns 1 and 3). Pre-migration differences in wealth across households (columns 2 and 4) should not, and do not, have any impact on the second stage migration coefficient. To sum up, our finding on bribes can neither be explained by wealth differences across migrant and non-migrant households nor by the income effect of remittances.

Regardless of this, the income effect of migration matters by improving families' ability to keep children in school. Whereas over 50 percent of non-migrant parents report barriers that will prevent the child from achieving the caregiver's desired level of education, this is the case for only 35 percent of migrant parents (Table 4.A.8: Panel A). The modal reason, a lack of finances, is cited by over 80 percent of caregivers in either group. Migration reduces barriers, in general, and financial barriers, in particular (Table 4.A.8: Panel B). The income effect in education is thus strong and positive, in stark contrast with its effect on petty corruption.

As a supporting *ad hoc* assessment of the mechanism, log remittances received by the household can be used in place of the migration dummy as the endogenous variable (results available on request). In this case, no more significant correlation between the

endogenous variable and informal payment is found in the second stage, which may be taken as tentative evidence that variation from the instrument does not affect bribe paying through the remittance channel. Even though one has to be careful interpreting such evidence, because it is no longer a valid IV approach, this may be interpreted as suggesting that, instead of remittances, other aspects of migration are likely to be the source of the bribe reducing effect. In line with other research, one might hypothesize that the negative coefficient of migration is explained by a lower willingness to bribe officials in the education system. This could be due to former migrants' own likelihood of bribing teachers or through social remittances (c.f. Ivlevs and King, 2017; Barsbai et al., 2017).

Irrespectively of whether it is the migrants themselves or their families who decrease bribe paying, our finding is promising from a normative point of view. From an economic standpoint, the money not given to teachers as informal "service fees" or "presents", i.e. for rent-seeking, could be used more productively on other household expenses and would stop distorting incentives for teachers and students. The emerging picture is thus a reduction in bribes and a simultaneous increase in the frequency of parental involvement in children's education due to migration. In the next section, possible transmission channels are discussed in more detail.

## 4.7 Transmission channels and robustness

According to the community leaders interviewed in the survey, the most widely perceived constraint to school quality is not a scarcity of staff but of other inputs, such as teaching materials or utilities (Table 4.A.9). Parental education inputs could be affected by the public funding situation of local schools, causing omitted variable bias.<sup>28</sup> Thus, we match our household data with administrative school-level expenditure data from an open budget initiative of the World Bank (BOOST) to ensure that the instrument is not picking up community-level variation in the supply of public education. The matching both datasets is imperfect, because the availability of the budget data was not anticipated at the time of the household survey (see appendix 4.A for a detailed description of the data and matching procedure). We include school expenditures (per pupil) in several categories as additional explanatory variables.<sup>29</sup> The strong negative effect on bribes remains even after adding the additional controls, which approximately

<sup>28</sup>Private educational spending responds to public funding; see, for example, Houtenville and Conway (2008).

<sup>29</sup>We do not find evidence that they are systematically correlated with migration.

halves the sample size.<sup>30</sup> Schools' wage bills, which closely correspond to the ratio of teachers-per-pupil (c.f. Figure 4.A.4), teaching material, and maintenance funds are not significantly correlated with household educational expenditures (Table 4.A.10: columns 1–6). In contrast, schools' expenditures on utilities and transports, for which community leaders often report lacking funds, exhibit signs of substitutability between private and public expenditure. There is also some tentative evidence of substitution between the parental investment of time and the time teachers could allocate to individual children (column 7).<sup>31</sup>

The strong correlation between migration and informal payments to teachers is also robust when controlling for an index of infrastructure quality of the school (Table 4.4: column 1). We furthermore tested whether the migration-induced reduction in informal payments is weaker in worse funded schools, where informal payments may be less controversial, but did not find any robust differences (results available upon request). Attending schools that have funding for buses, or that are more distant from the student's home—both of which proxy secondary and advanced secondary schools that cover larger areas—, correlates positively (although statistically insignificantly) with informal payments. When we add school fixed-effects to compare students within schools, we find that better-off parents pay more to teachers and buy more tutoring services (Table 4.4: columns 3 and 4).<sup>32</sup> This underscores the importance of the income effect. The migration coefficient is negative but insignificant, suggesting that much of the variation associated with migration occurs at the school level. This fits well our discussions with Moldovan experts, who stated that the payments to teachers that are collected by informal parental committees can quickly break down once a few parents refuse to pay them—an effect that often occurs in public good settings when punishment is weak (Fehr and Gächter, 2000). In line with our expert discussions, we thus interpret the migration effects as quickly spilling over within schools.<sup>33</sup>

To ensure that our results are not driven by local heterogeneity across communities rather than migration, we add a within-community dimension to the community-level variation of the original IV. We interact the *Network-Growth-IV* with the mean years

<sup>30</sup>The first-stage estimates are reported in column 2 of Table 4.A.4.

<sup>31</sup>Table 4.A.11 provides OLS results when the sample is split by migration status. The negative coefficient on the teacher-pupil ratio (proxied by wages per pupil) is similar for both migrant and non-migrant households, although statistically insignificant for the former.

<sup>32</sup>Note that migration as a major source of income inequality is not exogenized here due to a lack of a valid within-community IV.

<sup>33</sup>There is no statistically significant correlation between any school budget variable and the migration share of pupils in the household survey. Also, if migrant parents were planning to send their children abroad and therefore stopped paying local teachers, there should be strong differences within schools.

TABLE 4.4: Detailed school funding, infrastructure controls and school fixed effects

Estimator	(1) log(Payments to teachers) OLS	(2) log(Out-of-school tutoring) OLS	(3) log(Payments to teachers) FE	(4) log(Out-of-school tutoring) FE
Migration	-0.449** (0.210)	-0.062 (0.170)	-0.099 (0.214)	-0.074 (0.195)
Log(household income)	0.107 (0.087)	0.186*** (0.055)	0.142 (0.086)	0.123** (0.052)
<i>School infrastructure and budget</i>				
School asset index	-0.000 (0.000)	0.000 (0.000)		
Wages (log)	0.111 (0.457)	-0.019 (0.332)		
Teaching materials (log)	-0.062 (0.257)	-0.145 (0.201)		
Utilities (log)	-0.244 (0.176)	-0.106 (0.118)		
Transports (log)	0.138 (0.101)	0.041 (0.056)		
Maintenance (log)	-0.034 (0.079)	-0.031 (0.067)		
<i>Number of school fixed effects</i>			148	148
<i>Child characteristics</i>	Yes	Yes	Yes	Yes
<i>Household characteristics</i>	Yes	Yes	Yes	Yes
<i>N</i>	873	880	875	882
<i>R<sup>2</sup> / R<sup>2</sup> within</i>	0.116	0.141	0.026	0.027

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the school level in parentheses. The school asset index is based on dummies for working water supply, working hot water, working sewage, working heating capacity, the existence of a library and of a medical office, the number of classrooms, auditorium seats, and the gym and sports field size. This information as well as the school budgets come from BOOST data. School budget variables are per student amounts. School fixed effects estimates in columns 3 and 4. Sample size decreased due to inability to match all schools, c.f. Table 4.A.16 in the online appendix. *Child characteristics*: age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics*: mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations*: 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively.

of education of adult household members, since more educated households can be expected to be better able to respond to the growth-pull mechanism approximated by our IV. The new variable is positively related with migration and statistically significant at the 1 percent level. The estimated effect of migration on payments to teachers is almost identical to our main estimates (Table 4.A.12).

Two motivations for ceasing bribe-paying are plausible. First, migrant parents may be generally less tolerant of corruption due to their experience abroad. Second, migrant parents may demand actual cognitive achievement instead of good grades, because they have witnessed the unimportance of Moldovan certificates relative to actual skills for success abroad. A full 96 percent of caregivers replied that education was important to be successful abroad. Yet, there is no significant reduction in the perceived quality of children's individual schools (Table 4.A.13).<sup>34</sup> This gives us confidence that our results

<sup>34</sup> Alternatively, the main effects of migration on the provision of educational inputs remain unchanged

are not driven by differences in the cost-benefit analysis of the Moldovan school system between non-migrant and migrant households.

In order to provide some evidence of robustness as well as external validity of our study, we draw on another, less detailed dataset to show that a similar negative correlation between migration and bribe paying exists also in data independent of ours. We use the so-called *Barometer of Public Opinion* of Moldova's Institute for Public Policy. This is a well-regarded biannual survey that collects individual opinions on a wide range of topics regarding politics, values, and related issues. Informal payments to authorities and migration status were covered in the April 2013 survey.<sup>35</sup> Those individuals with migration experience to the West were more likely to have had contact with the justice system and were more likely to have been asked for bribes for the solution of their problem. Conditional on reporting *not* paying a bribe, people with any experience of migrating and especially the typically more wealthy migrants to the West were more likely to have been asked to pay informal fees than those without migration experience (odds ratio: 3.6 times). Individuals with migration experience thus seem to be less likely to pay bribes under a given level of pressure to do so.<sup>36</sup>

Still, our results could be driven by the migration-induced change in the identity of the child's caregiver, for example, reflecting that non-parental caregivers (e.g., grandparents, siblings, aunts or uncles) have less involvement in (or knowledge of) the education system and are, therefore, less likely to bribe teachers. They may also have lower opportunity costs of time and may therefore spend more time on the child's education. To rule out this mechanism we re-estimate the main results excluding all children whose caregiver is not a biological parent (Table 4.A.14). The slightly (but not significantly) larger coefficients of migration provide strong evidence that our results are not driven by caregiver change. Our results are furthermore robust to alternative but similar definitions of the migration dummy (e.g., who migrates or how long migration spells have to be). We also find no evidence that our effect is driven by caregivers who are return migrants.<sup>37</sup> More generally, including a dummy variable for return migrant households (i.e., those households with at least one return migrant but *no* current migrants) does not affect the

---

after including the perceived school quality variable as an additional control (available upon request).

<sup>35</sup>The sample contains 1,100 individuals from 76 communities and is nationally representative of the adult population. All results are available upon request.

<sup>36</sup>Our instrumental variable strategy does not allow us to identify destination specific effects. Therefore, our results are the average migration effect across all destinations, not just Western countries. If the effect is entirely driven by migration to the West, where corruption is far less common than in Moldova, then our 2SLS estimates are a lower bound for the true Western migration effect.

<sup>37</sup>We define a return migrant as an adult that spent more than 3 months abroad in one single spell since 1999 but is no longer a migrant at the time of the survey.



migration coefficient in our educational input IV regressions. Return migration itself has a negative coefficient which is smaller in absolute magnitude than the (current) migration estimate, but statistically insignificant (available upon request). Note that correcting for self-selection into return migration lies beyond the scope of this chapter.

Despite controlling for the mean years of education of adult household members in all regressions, it could still be possible that households were sorted on unobserved ability within Moldova. In that case, the size of the 2004 network could be correlated with families' unobservable skills. In the Moldovan context, this hypothesis is very unlikely. In Soviet times, internal migration was highly restricted and centralized. High skilled individuals were not only concentrated in the main cities, where tertiary education was available, but were often deployed as state bureaucrats to agricultural or industrial projects all over the country, especially the countryside. After the collapse of the Soviet Union, there has not been much internal migration. To corroborate our arguments, we re-run our main specifications excluding children living in the two cities, Chişinău and Bălţi, that exert the main pull effect internally. Our results remain fully robust throughout (available upon request).

As seen above, the main results are robust to a host of alternative explanations. If not paying bribes, however, had dire consequences for the children's educational performance, lower corruption might not be in their best interest. We therefore estimate the effects of migration on students' grade point average (GPA) (Table 4.A.15). Throughout the different specifications payments to teachers remain insignificant. In addition, and in line with the literature, we find a negative correlation between migration and the GPA that is partly compensated by household wealth. This underlines that most of the informal payments may not be directly meant to improve grades relative to classmates but rather operate as illicit user fees or per capita taxes. If payment ceases, students do not suffer worse grades, on average. However, students who receive extra attention from teachers due to tutoring (which are partly mere bribes), do better grade-wise. Also, many Moldovans suggest that bribing of teachers for grades is not effective anyway, because students study less hard if they expect to receive higher scores. Another possibility is that deviating from the common situation of paying bribes has no adverse effects, especially as standardized tests are increasingly used in the most important exams with the deliberate aim of fighting corruption in education.

## 4.8 Conclusion

In this chapter we analyze the effect of emigration on petty corruption in education, in particular on informal payments to teachers. Such payments are typically understood to have a dual motivation: fund raising for maintenance of schools as well as supplementing teachers' wages to increase their motivation and/or to focus their attention on individual children. We use the interaction between migrant networks and economic growth at the destination as an instrumental variable for the household's migration status in order to control for selection into migration. Using this IV approach, we document a reduction in informal payments to teachers. This aggregate migration effect consists, among others, of a non-negative income effect that is counteracted by other factors. By excluding alternative explanations, and in line with an emergent literature, we speculate that the widening of migrants' horizon (i.e., additional information or value change) may be the main driver of the reduction in petty corruption.

Incorporating school-level budget data in our analysis, we show that there is no strong correlation between public school funding and petty corruption. Thus, the most socially accepted justification for informal payments to teachers—the need for school maintenance and wage supplements—is not a good predictor of differences between schools. Within schools, additional analysis suggests that reductions in payments to teachers quickly spill over to non-migrants. This is in line with qualitative evidence according to which per capita payments to teachers cease once a few parents in a class refuse to pay, due to only weak enforcement devices in the hands of teachers or other paying parents. Our results fit with novel research that shows how participation in corruption often depends on people perceiving it as widespread. This is a prevalent phenomenon in low and middle income countries. In such a setting simply increasing teachers' salaries and school resources might decrease the perceived legitimacy of informal payments. If budget constraints make this impossible, and these payments continue to exist, structures such as teacher-parent-committees should formalize them as donations. The available funds should then be focused on making teachers wages sufficient while stepping up enforcement of laws against individual corruption. This way, transparency and accountability would be improved while providing solutions for underfunding that do not distort incentives. Both the opportunity to siphon off part of the payments for private use and the necessity to do so would thus decrease.

For bribes used to get the own child ahead of the competition, other measures are likely to be more effective. In a bold move, the Moldovan government recently introduced video-taping of the most important high school exam to put an end to teachers, motivated

by informal payments, telling answers to their classes or, worse, individual students. As such laudable reforms reduce the scope for corruption, they may also make it easier for both migrants and non-migrants to resist corruption. Focusing reform efforts on increasing awareness that petty corruption in education is a problem, stoking demand for educational achievement rather than for good grades, and creating incentives to deviate from the social norm of participating in petty corruption hold promise. This paper thus provides evidence of a petty corruption channel through which the all too often forgotten positive effects of emigration on origin countries can arise. Future work should seek to more clearly disentangle how such effects occur and what role the institutional and social contexts play.

## 4.A Appendix

### Detailed description of school-level data

The data on school-level public expenditures are derived from the World Bank's Open Budget Initiative (or BOOST).<sup>38</sup> The Moldovan Ministry of Finance provides all budgets of public organisms at a very disaggregated level and on a yearly basis, going back to 2005. Each item is classified according to source, function, and expenditure type. In Moldova, the financing of public schools is highly decentralized and typically determined at the municipality (or *rayon*) level. We collect all school-level budgets that were executed during the year 2010 and aggregate expenditures in five categories: (1) staff wages, (2) teaching materials (also includes food and office supplies), (3) utilities, (4) transportation, and (5) maintenance (includes small-scale purchases and repairs of physical capital). We drop all schools which do not have positive executed expenditures on categories (1), (2), and (3), since they are likely to suffer from severe missing data problems. However, we allow for zero executed totals on categories (4) and (5), since these are arguably not always necessary for the core activities of schools.

Finally, we obtain the total number of students for each school from administrative data of the Moldovan Ministry of Education. In summary, we have complete survey data for a total of 2,168 school-age children (6–18 years old) from 1,463 households. School names from the survey and the official records were first matched automatically. In a second step, we matched strings by hand, thus correcting minor errors, such as typos. Wherever we could certainly establish a link, we then manually entered the school code for the respective child. In many cases the string variable covering the school name did not point to a particular school with certainty. Whenever we were less than 100 percent sure about the correctness of a match we did not match the respective child's record. After matching the survey data with the school-level budgets and number of students, we have complete data for a sample of 1,158 children from 853 households. Most of the losses in sample size resulted from not reporting or misreporting the school name in the household survey and missing executed budget data at the school level. To a smaller extent, we could not unambiguously match some school names as reported in the household survey with their counterparts in the BOOST dataset—for example, if parents gave the school name as “liceu <municipality >”, but there were several schools of the respective type in that municipality.

Table 4.A.16 presents summary statistics of the child-observations successfully matched

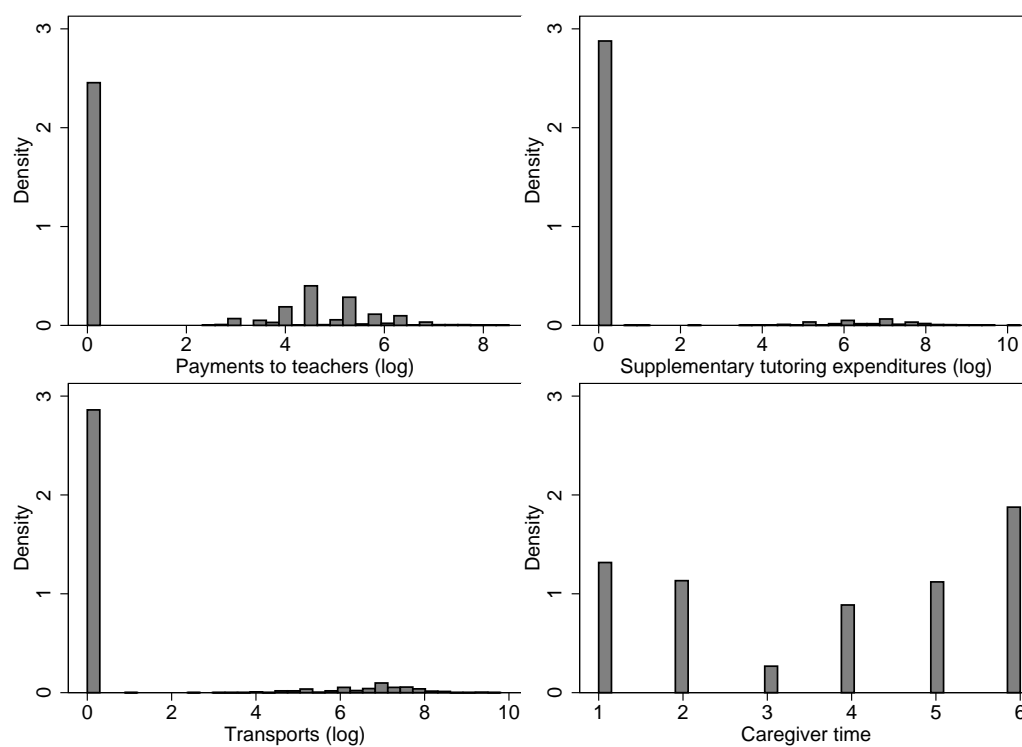
<sup>38</sup>The data are freely available at <http://wbi.worldbank.org/boost/country/moldova>.

across all data sources and of those for which the matching failed. Failure to match is to some extent random but tends to happen more often in urban areas, where, for example, a particular part of town has more than one school of a specific kind. As a consequence, 16–18 years old children who attend upper secondary schooling are also disproportionately missing from the matched sample. The reason is that, at higher education levels, teenagers tend to move away from smaller communities to attend school in more populous towns, where the chances of ambiguous matches across data sources are higher.<sup>39</sup> This pattern also explains why the average distance to school and transportation expenditures are significantly higher for the unmatched sample, while average caregiver time is lower.

For a clearer interpretation of the regression results, Figure 4.A.4 depicts that the school-level variation on budgets for staff wages is almost completely explained by variation in the number of teachers employed. The graph plots the values of school expenditures on wages against the predicted values of a regression of wage spending on the number of teachers. The red dashed line is the identity line (i.e.,  $y = x$ ). The regression's  $R^2$  is approximately 98 percent. Therefore, school budgets for staff wages can be thought of as a representation of the quantity of schoolteachers.

---

<sup>39</sup>Recall that enrollment in upper secondary schooling is no longer compulsory in Moldova.



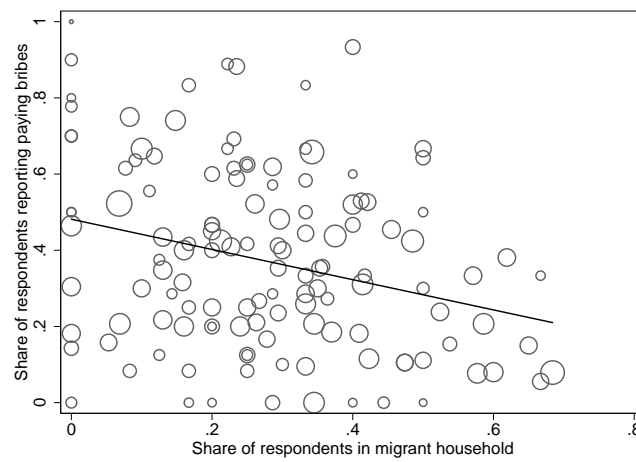
**FIG. 4.A.1:** Histogram of private education inputs

*Notes:* Authors' calculations based on CELB 2012.

**TABLE 4.A.1:** School enrollment rates in Moldova (2010)

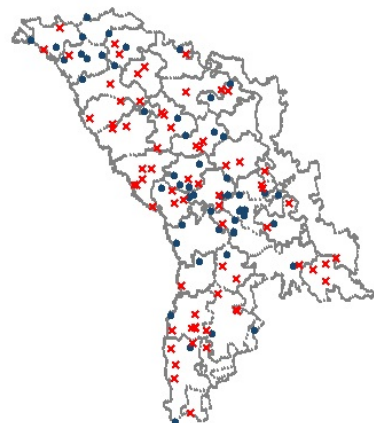
Age	N	Mean <sup>a</sup> (%)	(Std Error) <sup>b</sup>
7–10	696	99	(0.006)
11–15	973	99	(0.004)
16–18	666	87	(0.018)
All	2,335	96	(0.006)

*Source:* Authors' calculations based on CELB 2012. <sup>a</sup> Nationally representative weighted mean. <sup>b</sup> Standards errors clustered at the community level.



**FIG. 4.A.2:** Correlation of bribe payments to teachers and migration rate

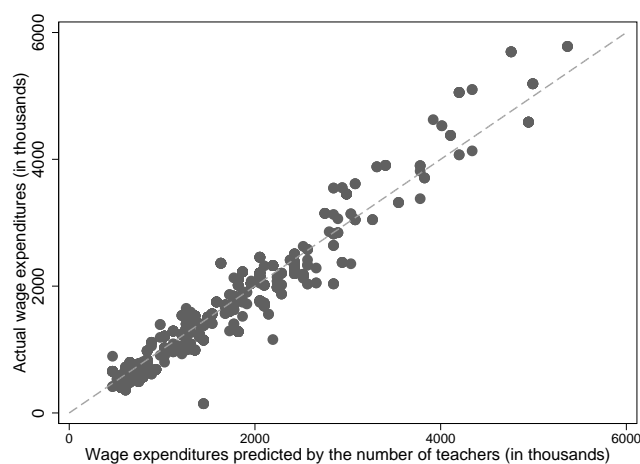
*Notes:* Authors' calculations based on CELB 2012. Each circle represents one community. The size corresponds to the number of children per community.



• Low values    × High values

**FIG. 4.A.3:** Map of communities with above and below median values of the network-growth IV

*Notes:* Graph is taken from Böhme et al. (2015).



**FIG. 4.A.4:** Schools' wage bills in BOOST proxy the number of teachers

*Notes:* Authors' calculations based on CELB 2012 and BOOST data.

**TABLE 4.A.2:** School budgets: descriptive statistics

PANEL A: Descriptives at the student level				
School budgets	Mean	(SD)	Min.	Max.
Total	5,522.73	(1,425.79)	847.57	19,497.42
Categories:				
Wages	4,114.28	(978.93)	372.13	9,646.96
Teaching materials	436.26	(222.29)	11.97	3,265.34
Utilities	409.91	(300.1)	43.16	2,622.54
Transports	159.76	(237.17)	0	1,914.40
Maintenance	349.05	(586.01)	0	9,178.03
<hr/>				
N	1,177			

PANEL B: Descriptives at the school level				
School budgets	Mean	(SD)	Min.	Max.
Total	5,638.62	(1,960.69)	847.57	19,497.42
Categories:				
Wages	4,214.31	(1,199.81)	372.13	9,646.96
Teaching materials	455.83	(300.07)	11.97	3,265.34
Utilities	399	(362.73)	43.16	2,622.54
Transports	122.23	(238.79)	0	1,914.40
Maintenance	388.96	(829.44)	0	9,178.03
Number of students	460.27	(321.75)	53	1,438
<hr/>				
N	182			

*Notes:* Authors' calculations based on BOOST and Moldovan Ministry of Education data for 2010. School budgets are executed school budgets during 2010 calculated on a per student basis. Successfully matched schools attended by children from the household survey only. All monetary values are expressed in Moldovan Lei.



TABLE 4.A.3: The effect of migration on private education inputs: reduced form estimates

	School enrollment	Payments to teachers		Out-of-school tutoring		Transportation expenditure		Caregiver time
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		log	$D(Y > 0)$	log	$D(Y > 0)$	log	$D(Y > 0)$	
Network-Growth	-0.000** (0.000)	-0.005** (0.002)	-0.001** (0.000)	-0.002*** (0.001)	-0.000** (0.000)	0.000 (0.002)	0.000 (0.000)	0.003*** (0.001)
<i>Child characteristics</i>								
Age	-0.008*** (0.001)	0.049*** (0.017)	0.008** (0.003)	0.026* (0.013)	0.003* (0.002)	0.181*** (0.020)	0.026*** (0.003)	-0.277*** (0.013)
Male	-0.020*** (0.007)	-0.389*** (0.109)	-0.078*** (0.023)	-0.218*** (0.082)	-0.033*** (0.012)	-0.042 (0.091)	-0.008 (0.014)	0.028 (0.076)
Serious illness	-0.007 (0.008)	0.351* (0.179)	0.040 (0.033)	0.273** (0.111)	0.035** (0.016)	0.232 (0.156)	0.033 (0.024)	-0.019 (0.109)
Distance to school (log)	-0.004 (0.006)	0.034 (0.089)	0.001 (0.017)	0.015 (0.055)	0.003 (0.008)	0.626*** (0.105)	0.087*** (0.015)	-0.019 (0.060)
<i>Household characteristics</i>								
Mean years education	0.006*** (0.002)	0.073** (0.034)	0.011 (0.007)	0.100*** (0.028)	0.014*** (0.004)	0.131*** (0.029)	0.018*** (0.004)	0.076*** (0.023)
Older siblings	0.045*** (0.007)	-0.233** (0.099)	-0.034* (0.018)	-0.178** (0.076)	-0.024** (0.011)	-0.154* (0.092)	-0.024* (0.014)	-0.223*** (0.063)
Household size	-0.005** (0.002)	-0.039 (0.046)	-0.008 (0.010)	0.010 (0.029)	0.002 (0.004)	0.015 (0.036)	0.002 (0.005)	-0.028 (0.036)
Parents divorced	-0.011 (0.014)	0.417* (0.212)	0.083* (0.042)	-0.000 (0.137)	0.001 (0.020)	0.094 (0.166)	0.015 (0.025)	-0.240* (0.136)
Urban	0.005 (0.007)	0.936*** (0.288)	0.126** (0.050)	1.127*** (0.225)	0.158*** (0.032)	0.889*** (0.253)	0.142 (0.043)	0.062 (0.147)
<i>Main migration destinations</i>								
Migrant share Italy	-0.000 (0.000)	0.007 (0.004)	0.002* (0.001)	-0.004 (0.003)	-0.001 (0.000)	-0.001 (0.004)	-0.000 (0.001)	0.001 (0.003)
Migrant share Ukraine	0.001** (0.001)	0.010 (0.014)	0.003 (0.003)	-0.004 (0.009)	-0.001 (0.001)	-0.024* (0.012)	-0.004* (0.002)	-0.032** (0.015)
Migrant share Romania	0.003*** (0.001)	0.004 (0.047)	0.001 (0.010)	-0.002 (0.015)	-0.000 (0.002)	-0.005 (0.024)	-0.000 (0.004)	0.015 (0.018)
Migrant share Russia	0.001** (0.000)	0.024** (0.010)	0.005** (0.002)	0.010*** (0.004)	0.001*** (0.001)	-0.004 (0.010)	-0.001 (0.002)	-0.015*** (0.006)
N	2,224	2,148	2,148	2,170	2,170	2,168	2,168	2,162
R <sup>2</sup>	0.064	0.068	0.042	0.095	0.084	0.173	0.168	0.265

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. Columns show reduced form regressions where the outcome of the second stage is regressed on the instrument (*Network Growth*) and the endogenous variable (migration) is excluded. Note that interpreting the size of the instrumental variable is not easy, because it is a sum of Network-Growth Interactions. Differences in missing values for the dependent variables explain the different number of observations across columns. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively.

TABLE 4.A.4: The effect of migration on private education inputs: first stage IV regression

	Migration	
	(1)	(2)
<i>Instrument</i>		
Network-Growth	0.001*** (0.000)	0.001*** (0.000)
<i>School budgets (per student)</i>		
Wages (log)		0.024 (0.072)
Teaching materials (log)		0.015 (0.036)
Utilities (log)		-0.010 (0.024)
Transports (log)		-0.016 (0.011)
Maintenance (log)		-0.011 (0.011)
<i>Child characteristics</i>		
Age	0.005 (0.003)	0.004 (0.005)
Male	-0.002 (0.018)	0.033 (0.027)
Serious illness	-0.008 (0.023)	0.003 (0.030)
Distance to school (log)	-0.014 (0.017)	-0.006 (0.024)
<i>Household characteristics</i>		
Mean years education	0.014** (0.006)	0.014* (0.008)
Older siblings	0.013 (0.017)	0.014 (0.024)
Household size	0.040*** (0.011)	0.044*** (0.015)
Parents divorced	-0.026 (0.044)	-0.054 (0.051)
Urban	-0.125*** (0.033)	-0.118** (0.056)
<i>Main migration destinations</i>		
Migrant share Italy	-0.000 (0.001)	-0.002 (0.001)
Migrant share Ukraine	0.001 (0.003)	-0.012* (0.007)
Migrant share Romania	-0.005 (0.005)	-0.004 (0.005)
Migrant share Russia	-0.004** (0.002)	-0.004** (0.002)
Constant	-0.165 (0.127)	-0.317 (0.576)
<i>N</i>	2,148	1,148
<i>F-stat</i>	10.6	9.1

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively.

TABLE 4.A.5: The effect of migration on private education inputs

Second stage IV regressions	School enrollment	Payments to teachers		Out-of-school tutoring		Transportation expenditure		Caregiver time
		(1)	(2) log	(3) $D(Y > 0)$	(4) log	(5) $D(Y > 0)$	(6) log	
Migration	-0.129 (0.097)	-4.430*** (1.717)	-0.829** (0.332)	-1.987** (0.869)	-0.274** (0.126)	0.247 (1.905)	0.091 (0.304)	2.667* (1.449)
<i>Child characteristics</i>								
Age	-0.007*** (0.002)	0.072*** (0.023)	0.012*** (0.004)	0.037** (0.015)	0.005** (0.002)	0.180*** (0.022)	0.026*** (0.003)	-0.294*** (0.019)
Male	-0.019*** (0.007)	-0.400*** (0.142)	-0.080*** (0.028)	-0.214** (0.089)	-0.033** (0.013)	-0.042 (0.090)	-0.008 (0.013)	0.039 (0.096)
Serious illness	-0.006 (0.009)	0.314* (0.181)	0.033 (0.033)	0.262** (0.113)	0.033** (0.016)	0.234 (0.156)	0.033 (0.024)	0.014 (0.131)
Distance to school (log)	-0.006 (0.008)	-0.028 (0.116)	-0.011 (0.022)	-0.018 (0.061)	-0.001 (0.009)	0.630*** (0.113)	0.088*** (0.017)	0.021 (0.087)
<i>Household characteristics</i>								
Mean years education	0.008*** (0.002)	0.136*** (0.048)	0.023** (0.009)	0.127*** (0.031)	0.017*** (0.004)	0.128*** (0.043)	0.017** (0.007)	0.042 (0.029)
Older siblings	0.047*** (0.008)	-0.177 (0.127)	-0.023 (0.023)	-0.151* (0.086)	-0.021 (0.013)	-0.157* (0.092)	-0.025* (0.014)	-0.264*** (0.091)
Household size	0.000 (0.005)	0.139 (0.093)	0.025 (0.018)	0.089 (0.055)	0.012 (0.008)	0.005 (0.090)	-0.001 (0.014)	-0.136* (0.079)
Parents divorced	-0.014 (0.014)	0.303 (0.282)	0.061 (0.055)	-0.044 (0.174)	-0.005 (0.025)	0.100 (0.180)	0.018 (0.028)	-0.177 (0.168)
Urban	-0.012 (0.013)	0.382 (0.361)	0.022 (0.066)	0.875*** (0.237)	0.123*** (0.033)	0.920*** (0.313)	0.153*** (0.050)	0.409* (0.224)
<i>Main migration destinations</i>								
Migrant share Italy	-0.000 (0.000)	0.005 (0.004)	0.001 (0.001)	-0.005* (0.003)	-0.001* (0.000)	-0.001 (0.004)	-0.000 (0.001)	0.004 (0.003)
Migrant share Ukraine	0.001* (0.001)	0.014 (0.022)	0.004 (0.005)	-0.002 (0.012)	-0.000 (0.002)	-0.025* (0.014)	-0.004* (0.002)	-0.033* (0.019)
Migrant share Romania	0.003** (0.001)	-0.017 (0.035)	-0.003 (0.008)	-0.011 (0.012)	-0.001 (0.002)	-0.004 (0.025)	0.000 (0.004)	0.025 (0.022)
Migrant share Russia	0.000 (0.000)	0.007* (0.004)	0.001* (0.001)	0.002 (0.002)	0.000* (0.000)	-0.003 (0.003)	-0.000 (0.001)	-0.004 (0.003)
N	2,223	2,148	2,148	2,170	2,170	2,168	2,168	2162
K-P weakid	9.6	10.6	10.6	11	11	10.7	10.7	10.6
95% CLR confidence set	[-0.4, 0.06]	[-9.25, -1.88]	[-1.76, -0.32]	[-4.99, -0.01]	[-0.71, 0.02]	[-2.34, 2.82]	[-0.28, 0.49]	[0.99, 5.57]
CLR test p-value	0.17	0.00	0.00	0.05	0.07	0.83	0.60	0.00
Cluster-robust 95% AR conf.	[-0.51, 0.02]	[-9.12, -0.83]	[-1.68, -0.08]	[-5.12, -0.65]	[-0.72, -0.07]	[-3.45, 5.91]	[-0.47, 1.02]	[0.66, 8.35]
Cluster-robust AR p-value	0.11	0.03	0.04	0.01	0.01	0.90	0.76	0.01

Notes: Authors' calculations based on CHLB2012. Standard errors are clustered at the community level. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ . K-P weakid is the Kleibergen-Paap weak identification statistic. Migration is instrumented using a network-growth interaction IV. For first stage, refer to Table 4.A.4. The CLR test is the confidence region and test statistic from the "condivreg" package (Mikushova and Poi, 2006). The cluster-robust AR 95% confidence set is from the "rivestf" package (Finlay and Magnusson, 2009).

**TABLE 4.A.6:** The effect of migration on education expenditures: average marginal effects after IV probit

	Payments to teachers		Out-of-school tutoring		Transportation expenditure	
	(1) $D(Y > 0)$	(2) $D(Y > 0)$	(3) $D(Y > 0)$	(4) $D(Y > 0)$	(5) $D(Y > 0)$	(6) $D(Y > 0)$
Migration	-0.536*** (0.069)	-0.433*** (0.157)	-0.369** (0.177)	-0.277 (0.174)	0.120 (0.305)	0.243 (0.280)
<i>School budgets</i>	No	Yes	No	Yes	No	Yes
<i>Child characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Household characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Main migration destinations</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	2,148	1,158	2,170	1,158	2,168	1,158

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. *School budgets*: (log) per student school budgets for (i) staff wages, (ii) teaching materials, (iii) utilities, (iv) transports, (v) maintenance. *Child characteristics*: age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics*: mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations*: 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). Migration is instrumented using a network-growth interaction IV.

TABLE 4.A.7: The effect of migration on private education inputs: OLS regressions

	School enrollment (1)	Payments to teachers		Out-of-school tutoring		Transportation expenditure		Caregiver time (8)
		(2) log	(3) $D(Y > 0)$	(4) log	(5) $D(Y > 0)$	(6) log	(7) $D(Y > 0)$	
Migration	0.024*** (0.008)	-0.142 (0.141) (0.135)	-0.026 (0.028) (0.027)	0.018 (0.099) (0.096)	0.007 (0.015) (0.015)	0.277** (0.134) (0.131)	0.043** (0.020) (0.020)	-0.053 (0.098) (0.098)
<i>Child characteristics</i>								
Age	-0.008*** (0.001)	0.050*** (0.017)	0.008** (0.003)	0.026* (0.013)	0.003 (0.002)	0.180*** (0.020)	0.026*** (0.003)	-0.276*** (0.013)
Male	-0.019*** (0.007)	-0.397*** (0.110)	-0.079*** (0.023)	-0.222*** (0.082)	-0.034*** (0.012)	-0.043 (0.090)	-0.008 (0.013)	0.031 (0.076)
Serious illness	-0.004 (0.008)	0.370** (0.180)	0.044 (0.033)	0.283* (0.111)	0.036** (0.016)	0.235 (0.158)	0.033 (0.024)	-0.032 (0.108)
Distance to school (log)	-0.003 (0.006)	0.035 (0.086)	0.001 (0.017)	0.017 (0.055)	0.003 (0.008)	0.630*** (0.106)	0.087*** (0.015)	-0.024 (0.061)
<i>Household characteristics</i>								
Mean years education	0.006*** (0.002)	0.083** (0.034)	0.013* (0.007)	0.104*** (0.028)	0.014*** (0.004)	0.128*** (0.030)	0.017*** (0.005)	0.070*** (0.023)
Older siblings	0.044*** (0.007)	-0.235** (0.100)	-0.034* (0.018)	-0.180** (0.077)	-0.025** (0.011)	-0.158* (0.092)	-0.025* (0.014)	-0.220*** (0.062)
Household size	-0.006** (0.002)	-0.039 (0.044)	-0.009 (0.009)	0.006 (0.029)	0.001 (0.004)	0.004 (0.036)	0.001 (0.005)	-0.022 (0.036)
Parents divorced	-0.011 (0.014)	0.418** (0.206)	0.083** (0.041)	0.003 (0.138)	0.002 (0.020)	0.101 (0.168)	0.017 (0.025)	-0.246* (0.138)
Urban	0.006 (0.008)	0.878*** (0.298)	0.115** (0.052)	1.110*** (0.229)	0.156*** (0.032)	0.923*** (0.250)	0.148*** (0.042)	0.081 (0.152)
<i>Main migration destinations</i>								
N	2,223	2,148	2,148	2,170	2,170	2,168	2,168	2,162
R <sup>2</sup>	0.068	0.063	0.037	0.094	0.082	0.175	0.171	0.261
F-stat	4.43	6.74	4.81	3.96	4.27	13.07	13.76	39.76

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively.

**TABLE 4.A.8:** Does migration predict perceived barriers to education?

Panel A: Perceived barriers to child's education (%)			
D.048: Is there anything standing in child's way of achieving your desired level of education?	Non-migrant	Migrant	Total
Yes	51.7	34.6	46.8
No	45.8	63.3	50.8
Does not apply	1.3	0.7	1.1
Does not know	1.1	1.5	1.2
Total	100	100	100
N	1,521	610	2,131
D.049: If yes, what?			
Financial reasons	87.6	81.5	86.3
Child's ability	7.3	6.8	7.2
Other reasons	2.6	5.0	3.1
No access to required level due to distance	0.5	2.7	1.0
Does not know	2.0	4.1	2.4
Total	100	100	100
N	808	222	1,030

Panel B: Second-stage IV regressions - LPM and IV Probit				
	D(Barrier to child's education)			
	(1) LPM	(2) LPM	(3) Probit <sup>a</sup>	(4) Probit <sup>a</sup>
Migration	-1.930*** (0.471)	-1.924*** (0.567)	-0.612*** (0.010)	-0.618*** (0.016)
School budgets	No	Yes	No	Yes
Child characteristics	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes
Main migration destinations	Yes	Yes	Yes	Yes
N	2,008	1,057	2,008	1,057
K-P weakid	22.37	17.15	–	–

*Notes:* Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. *School budgets:* (log) per student school budgets for (i) staff wages, (ii) teaching materials, (iii) utilities, (iv) transports, (v) maintenance. *Child characteristics:* age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics:* mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations:* 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. K-P weakid refers to the Kleibergen-Paap weak identification statistic. Migration is instrumented using a network-growth interaction IV. <sup>a</sup> average marginal effect reported for IV probit.

**TABLE 4.A.9:** Main problems with local public schools: community leaders' perceptions

	Main problem for this/these school(s)	
	Frequency	Share (in %)
<i>Primary schools</i>		
Lack of teaching materials (books, etc)	373	41.49
Lack of health and hygiene facilities	173	19.24
Other (specify)	172	19.13
Lack of water	123	13.68
Lack of teachers	58	6.45
Total	899	100
<i>Secondary schools</i>		
Lack of teaching materials (books, etc)	372	41.01
Other (specify)	200	22.05
Lack of health and hygiene facilities	129	14.22
Lack of water	128	14.11
Lack of teachers	78	8.60
Total	907	100

Source: Community questionnaires of CELB 2012 that asked community leaders about schools in the community in general. The difference in number of observations between here and the sample used in Table 4.A.10 is due to some communities not having a school or the community leader not providing an answer to the question.

**TABLE 4.A.10:** The effect of migration on private education inputs: controlling for public school expenditures

<i>Second stage IV regressions</i>	Payments to teachers		Out-of-school tutoring		Transportation expenditure		Caregiver time
	(1) log	(2) $D(Y > 0)$	(3) log	(4) $D(Y > 0)$	(5) log	(6) $D(Y > 0)$	(7)
Migration	-2.576* (1.494)	-0.470* (0.285)	-0.926 (0.810)	-0.123 (0.116)	0.897 (1.890)	0.168 (0.295)	1.189 (1.263)
<i>School budgets (per pupil)</i>							
Wages (log)	0.049 (0.454)	-0.007 (0.088)	-0.341 (0.391)	-0.048 (0.056)	0.535 (0.456)	0.067 (0.069)	-0.553* (0.291)
Teaching materials (log)	-0.081 (0.248)	-0.016 (0.045)	-0.013 (0.191)	-0.006 (0.026)	-0.222 (0.210)	-0.024 (0.030)	-0.043 (0.117)
Utilities (log)	-0.252 (0.179)	-0.051 (0.033)	-0.168 (0.110)	-0.021 (0.016)	-0.306** (0.138)	-0.048** (0.021)	0.099 (0.084)
Transports (log)	0.043 (0.083)	0.014 (0.015)	-0.046 (0.053)	-0.005 (0.008)	-0.212** (0.092)	-0.030** (0.013)	0.075 (0.048)
Maintenance (log)	-0.062 (0.078)	-0.004 (0.014)	-0.028 (0.071)	-0.005 (0.011)	0.065 (0.068)	0.007 (0.011)	0.070 (0.052)
<i>Child characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Household characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Main migration destinations</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,148	1,148	1,158	1,158	1,158	1,158	1,177
K-P weakid	9.092	9.092	9.112	9.112	8.934	8.934	8.889

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. *Child characteristics*: age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics*: mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations*: 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. K-P weak id is the Kleibergen-Paap weak identification statistic. Migration is instrumented using a network-growth interaction IV.

TABLE 4.A.11: Determinants of caregiver time: OLS results; sample split by migration status

Migrants Estimator	Caregiver time			
	(1) no OLS	(2) yes OLS	(3) no OLS	(4) yes OLS
<i>School budgets (per student)</i>				
Wages (log)	-0.633** (0.270)	-0.693 (0.438)		
D(School wage bills per student above median)			-0.314* (0.163)	-0.171 (0.233)
Teaching materials (log)	-0.127 (0.157)	0.218 (0.199)	-0.244 (0.148)	0.182 (0.194)
Utilities (log)	0.0858 (0.105)	0.0960 (0.147)	0.105 (0.111)	0.0706 (0.148)
Transports (log)	0.0685 (0.0451)	0.0320 (0.0763)	0.0761* (0.0446)	0.0192 (0.0737)
Maintenance (log)	0.0696 (0.0536)	0.139** (0.0672)	0.0856 (0.0541)	0.146** (0.0697)
<i>Child characteristics</i>				
Age	-0.269*** (0.0177)	-0.269*** (0.0309)	-0.272*** (0.0179)	-0.268*** (0.0310)
Male	0.128 (0.106)	0.407* (0.220)	0.127 (0.107)	0.429* (0.224)
Serious illness	-0.0358 (0.155)	-0.300 (0.214)	-0.0369 (0.152)	-0.306 (0.211)
Distance to school (log)	-0.0880 (0.0975)	0.00838 (0.143)	-0.0946 (0.0970)	0.00763 (0.146)
<i>Household characteristics</i>				
Household size	0.0266 (0.0574)	-0.00339 (0.0862)	0.0247 (0.0570)	0.00357 (0.0882)
Mean years education	0.0430 (0.0342)	0.124** (0.0566)	0.0436 (0.0338)	0.126** (0.0556)
Older siblings	-0.219** (0.103)	-0.457** (0.186)	-0.228** (0.102)	-0.470** (0.190)
Parents divorced	-0.470** (0.202)	-0.340 (0.374)	-0.435** (0.205)	-0.335 (0.376)
Urban	0.538** (0.261)	0.579 (0.413)	0.560** (0.264)	0.560 (0.394)
Constant	11.81*** (1.911)	9.052** (3.596)	7.256*** (1.265)	3.687** (1.760)
N	862	315	862	315
R <sup>2</sup>	0.227	0.285	0.226	0.282

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively.



**TABLE 4.A.12:** Household-level IV: second stage

<i>Second stage IV regression</i>	<i>Payments to teachers</i>	
	(1) log	(2) $D(Y > 0)$
$X_h$ = Mean years education		
Migration	-4.161*** (1.593)	-0.776*** (0.299)
<i>Child characteristics</i>		
Age	0.071*** (0.022)	0.012*** (0.004)
Male	-0.382*** (0.137)	-0.076*** (0.027)
Serious illness	0.325* (0.177)	0.036 (0.033)
Distance to school (log)	-0.016 (0.109)	-0.009 (0.021)
<i>Household characteristics</i>		
Mean years education	0.232*** (0.078)	0.038*** (0.014)
Older siblings	-0.186 (0.124)	-0.025 (0.023)
Household size	0.116 (0.084)	0.020 (0.016)
Parents divorced	0.305 (0.275)	0.060 (0.054)
Urban	0.461 (0.351)	0.038 (0.063)
<i>Main migration destinations</i>		
Migrant share Italy	0.030* (0.015)	0.005 (0.003)
Migrant share Italy * $X_h$	-0.002 (0.001)	-0.000 (0.000)
Migrant share Ukraine	-0.094** (0.040)	-0.022*** (0.008)
Migrant share Ukraine * $X_h$	0.010** (0.005)	0.002** (0.001)
Migrant share Romania	0.324*** (0.113)	0.067*** (0.023)
Migrant share Romania * $X_h$	-0.032*** (0.010)	-0.006*** (0.002)
Migrant share Russia	0.008 (0.013)	0.001 (0.003)
Migrant share Russia * $X_h$	-0.000 (0.001)	-0.000 (0.000)
Constant	-1.170 (1.001)	-0.080 (0.186)
$N$	2,148	2,148
K-P weakid	12.3	12.3

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. Mean years education is the mean years of education of the adult household members. K-P weakid refers to the Kleibergen-Paap weak identification statistic. Migration is instrumented using a network-growth \*  $X_h$  interaction IV.

TABLE 4.A.13: Second stage on perceived school quality

Second stage IV regression	Perceived school quality	
	(1)	(2)
Migration	0.433 (0.786)	0.560 (0.798)
<i>School budgets</i>	No	Yes
<i>Child characteristics</i>	Yes	Yes
<i>Household characteristics</i>	Yes	Yes
<i>Main migration destinations</i>	Yes	Yes
<i>N</i>	2,205	1,174
<i>K-P weakid</i>	10.15	8.8

*Notes:* Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. *School budgets:* (log) per student school budgets for (i) staff wages, (ii) teaching materials, (iii) utilities, (iv) transports, (v) maintenance. *Child characteristics:* age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics:* mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations:* 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. K-P weakid refers to the Kleibergen-Paap weak identification statistic. Migration is instrumented using a network-growth interaction IV. *Perceived school quality* is the caregiver's perception of the public school's quality in a 5-item scale from *very poor* ... *very good*.

**TABLE 4.A.14:** The effect of migration on private education inputs: when the caregiver is a biological parent

	Second stage IV regressions		Payments to teachers		Out-of-school tutoring		Transportation expenditure		Caregiver time
	(1) log	(2) $D(Y > 0)$	(3) log	(4) $D(Y > 0)$	(5) log	(6) $D(Y > 0)$			
Migration	-4.640*** (1.354)	-0.858*** (0.256)	-2.128** (0.960)	-0.276** (0.135)	-0.214 (1.533)	0.006 (0.228)		2.160* (1.154)	
Child characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Main migration destinations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
N	1,849	1,849	1,865	1,865	1,859	1,859		1,860	
K-P weakid	13,857	13,857	14,114	14,114	13,820	13,820		12,736	

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors are clustered at the community level in parentheses. *Child characteristics*: age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics*: mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations*: 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. K-P weakid is the Kleibergen-Paap weak identification statistic. Migration is instrumented using a network-growth interaction IV.

TABLE 4.A.15: Determinants of grade point average (GPA)

	GPA (0-10)					
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
Migration	0.041 (0.077)	0.044 (0.077)	0.004 (0.069)	-0.828 (0.533)	-0.832 (0.518)	-1.527** (0.600)
Household asset index in 1999 (log)	-0.037 (0.070)	-0.072 (0.071)		-0.075 (0.073)	-0.109 (0.072)	
<i>Endogenous outcomes</i>						
Household asset index in 2011 (log)			0.310*** (0.079)			0.539*** (0.141)
Payments to teachers (log)		0.007 (0.014)	0.012 (0.014)		-0.002 (0.017)	0.000 (0.021)
Tutoring expenditures (log)		0.041** (0.016)	0.027* (0.016)		0.037** (0.017)	0.022 (0.021)
Transportation expenditures (log)		0.038** (0.016)	0.019 (0.014)		0.047** (0.018)	0.032* (0.019)
Caregiver time		0.032 (0.023)	0.029 (0.021)		0.034 (0.026)	0.025 (0.025)
<i>School budgets</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Child characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Household characteristics</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Main migration destinations</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	804	799	1,019	804	799	1019
adj. $R^2$	0.200	0.215	0.230	0.071	0.083	-0.160
K-P weakid	—	—	—	12.767	12.414	7.201

Notes: Authors' calculations based on CELB 2012. Heteroskedasticity robust standard errors that cluster at the community level in parentheses. *School budgets*: (log) per student school budgets for (i) staff wages, (ii) teaching materials, (iii) utilities, (iv) transports, (v) maintenance. *Child characteristics*: age, gender, serious illness in the past 12 months (dummy variable), and the (log) distance to school in minutes. *Household characteristics*: mean years of education of adult members, older siblings (dummy variable), household size, parents divorced, and urban residence status. *Main migration destinations*: 2004 share of the community's population that is a migrant to Italy, Ukraine, Romania and Russia (4 variables). \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively. K-P weakid refers to the Kleibergen-Paap weak identification statistic. Migration is instrumented using a network-growth interaction IV. Please interpret columns 2, 3, 5, and 6 with caution since they include endogenous covariates for illustrative purposes.

**TABLE 4.A.16:** Summary statistics: sample selection from matching

	<i>N</i>	Matched Mean	(SD)	<i>N</i>	Matching failed Mean	(SD)	Mean equality (t-test)
<i>Child characteristics</i>							
Age	1,265	11.96	(3.21)	1,413	13.61	(4.41)	***
Male	1,265	0.50		1,413	0.51		
Serious illness (past year)	1,265	0.31		1,413	0.27		**
Distance to school (min)	1,190	19.18	(12.87)	1,296	22.90	(24.9)	***
<i>Household characteristics</i>							
Migration	1,265	0.26		1,413	0.32		***
Household size	1,265	4.80	(1.57)	1,413	4.84	(1.48)	
Mean years education	1,265	10.81	(2.28)	1,412	10.68	(2.22)	
Urban	1,265	0.19		1,413	0.23		**
Older siblings	1,260	0.61		1,241	0.57		*
Parents divorced	1,265	0.12		1,413	0.11		
<i>Private inputs to child's education</i>							
Caregiver time	1,189	4.00	(1.9)	1,153	3.24	(1.97)	***
Payments to teachers	1,151	80.23	(236.3)	1,186	90.68	(266.59)	
Tutoring expenditures	1,161	198.74	(1,240.23)	1,202	133.65	(693.28)	
Transportation expenditures	1,162	145.78	(604.34)	1,197	315.93	(1,002.93)	***

Notes: Authors' calculations based on CELB 2012. All monetary values are expressed in Moldovan Lei. \*, \*\*, and \*\*\* indicate  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively.



# Bibliography

- Aaronson, D., Dehejia, R., Jordan, A., Pop-Eleches, C., Samii, C., and Schulze, K. (2017). The Effect of Fertility on Mothers' Labor Supply over the Last Two Centuries. IZA Discussion Paper No. 10559, IZA Institute of Labor Economics.
- Acemoglu, D., Autor, D. H., and Lyle, D. (2004). Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury. *Journal of Political Economy*, 112(3):497–551.
- Adams, R. H. and Page, J. (2005). Do International Migration and Remittances Reduce Poverty in Developing Countries? *World Development*, 33(10):1645–1669.
- Ager, P., Worm Hansen, C., and Sandholt Jensen, P. (2018). Fertility and Early-Life Mortality: Evidence from Smallpox Vaccination in Sweden. *Journal of the European Economic Association*, 16(2):487–521.
- Agüero, J. M. and Marks, M. S. (2011). Motherhood and Female Labor Supply in the Developing World: Evidence from Infertility Shocks. *Journal of Human Resources*, 46(4):800–826.
- Akee, R. and Kapur, D. (2012). Remittances and Rashomon. Working Paper 285, Center for Global Development.
- Alesina, A., Giuliano, P., and Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *Quarterly Journal of Economics*, 128(2):469–530.
- Alexander, A. C. and Welzel, C. (2015). Eroding Patriarchy: the Co-evolution of Women's Rights and Emancipative Values. *International Review of Sociology*, 25(1):144–165.
- Altonji, J. G. and Blank, R. M. (1999). Chapter 48 - Race and Gender in the Labor Market. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3, pages 3143–3259. Elsevier, Amsterdam.
- Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1):151–184.
- Alves, J. E. D. (2005). A Definição de Família Convivente do IBGE: Cuidados Metodológicos Necessários. *Aparte-Inclusão Social em Debate*, Rio de Janeiro, UFRJ. Available at: [http://www.ie.ufrj.br/aparte/pdfs/notametodologica\\_familiasconviventes.pdf](http://www.ie.ufrj.br/aparte/pdfs/notametodologica_familiasconviventes.pdf). [accessed on March 2017].

- Angrist, J. D. and Evans, W. N. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *American Economic Review*, 88(3):450–477.
- Antman, F. M. (2011). The Intergenerational Effects of Paternal Migration on Schooling and Work: What Can We Learn from Children's Time Allocations? *Journal of Development Economics*, 96(2):200–208.
- Antman, F. M. (2012). Gender, Educational Attainment, and the Impact of Parental Migration on Children Left Behind. *Journal of Population Economics*, 25(4):1187–1214.
- Ashraf, N., Field, E., and Lee, J. (2014). Household Bargaining and Excess Fertility: An Experimental Study in Zambia. *American Economic Review*, 104(7):2210–2237.
- Assaad, R., Hendy, R., and Yassine, S. (2014). Gender and the Jordanian Labor Market. In Assaad, R., editor, *The Jordanian Labor Market in the New Millennium*, pages 105–143. Oxford University Press, Oxford.
- Bandiera, O. and Does, A. N. (2013). Does Gender Inequality Hinder Development and Economic Growth? Evidence and Policy Implications. *World Bank Research Observer*, 28(1):2–21.
- Banerji, M., Martin, S., and Desai, S. (2008). Is Education Associated with a Transition Towards Autonomy in Partner Choice? A Case Study of India. India Human Development Survey Working Paper No. 8.
- Bansak, C. and Chezum, B. (2009). How Do Remittances Affect Human Capital Formation of School-Age Boys and Girls? *American Economic Review*, 99(2):145–148.
- Bardhan, P. (1997). Corruption and Development: a Review of Issues. *Journal of Economic Literature*, 35(3):1320–1346.
- Barro, R. J. and Lee, J. W. (2013). A New Data Set of Educational Attainment in the World, 1950–2010. *Journal of Development Economics*, 104:184–198.
- Barsbai, T., Rapoport, H., Steinmayr, A., and Trebesch, C. (2017). The Effect of Labor Migration on the Diffusion of Democracy: Evidence from a Former Soviet Republic. *American Economic Journal: Applied Economics*, 9(3):36–69.
- Batista, C. and Vicente, P. C. (2011). Do Migrants Improve Governance at Home? Evidence from a Voting Experiment. *World Bank Economic Review*, 25(1):77–104.
- Baumann, H. (1928). The Division of Work According to Sex in African Hoe Culture. *Africa: Journal of the International African Institute*, 1(3):289–319.
- Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful Women: Does Exposure Reduce Bias? *Quarterly Journal of Economics*, 124(4):1497–1540.



- Becker, G. S. (1960). An Economic Analysis of Fertility. In *Demographic and Economic Change in Developed Countries*, pages 209–240. National Bureau of Economic Research, New York.
- Becker, G. S. (1981). *A Treatise on the Family*. Harvard University Press, Cambridge, Massachusetts.
- Becker, G. S. and Barro, R. J. (1988). A Reformulation of the Economic Theory of Fertility. *Quarterly Journal of Economics*, 103(1):1–26.
- Beine, M., Docquier, F., and Schiff, M. (2013). International Migration, Transfer of Norms and Home Country Fertility. *Canadian Journal of Economics*, 46(4):1406–1430.
- Bentzen, J. S., Kaarsen, N., and Wingender, A. M. (2017). Irrigation and Autocracy. *Journal of the European Economic Association*, 15(1):1–53.
- Berry, J. W. (1997). Immigration, Acculturation, and Adaptation. *Applied Psychology*, 46(1):5–34.
- Bertrand, M. (2011). Chapter 17 - New Perspectives on Gender. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4, pages 1543–1590. Elsevier, Amsterdam.
- Bertrand, M. and Duflo, E. (2017). Chapter 8 - Field Experiments on Discrimination. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Field Experiments*, volume 1 of *Handbook of Economic Field Experiments*, pages 309–393. North-Holland, Amsterdam.
- Besley, T., Folke, O., Persson, T., and Rickne, J. (2017). Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden. *American Economic Review*, 107(8):2204–2242.
- Blau, F. D. and Kahn, L. M. (2007). Changes in the Labor Supply Behaviour of Married Women: 1980-2000. *Journal of Labor Economics*, 25(3):393–438.
- Blecker, R. A. and Seguino, S. (2002). Macroeconomic Effects of Reducing Gender Wage Inequality in an Export-Oriented, Semi-Industrialized Economy. *Review of Development Economics*, 6(1):103–119.
- Bloom, D. E., Kuhn, M., and Prettnner, K. (2015). The Contribution of Female Health to Economic Development. NBER Working Paper 21411, National Bureau of Economic Research.
- Blundell, R. and Macurdy, T. (1999). Chapter 27 - Labor Supply: A Review of Alternative Approaches. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3, pages 1559–1695. Elsevier, Amsterdam.

- Blundell, R., MaCurdy, T., and Meghir, C. (2007). Chapter 69 - Labor Supply Models: Unobserved Heterogeneity, Nonparticipation and Dynamics. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6, pages 4667–4775. Elsevier, Amsterdam.
- Böhme, M., Persian, R., and Stöhr, T. (2015). Alone But Better Off? Adult Child Migration and Health of Elderly Parents in Moldova. *Journal of Health Economics*, 39:211–227.
- Böhme, M. and Stöhr, T. (2014). Household Interview Duration Analysis in CAPI Survey Management. *Field Methods*, 26(4):390–405.
- Borcan, O., Lindahl, M., and Mitrut, A. (2017). Fighting Corruption in Education: What Works and Who Benefits? *American Economic Journal: Economic Policy*, 9(1):180–209.
- Borrowman, M. and Klasen, S. (2017). Drivers of Gendered Sectoral and Occupational Segregation in Developing Countries. Courant Research Centre: Poverty, Equity and Growth—Discussion Papers No. 222, University of Goettingen.
- Boserup, E. (1970). *Woman's Role in Economic Development*. George Allen and Unwin Ltd, London.
- Boter, C. (2017). Marriages are Made in Kitchens: The European Marriage Pattern and Life-Cycle Servanthood in Eighteenth-Century Amsterdam. *Feminist Economics*, 23(2):68–92.
- Branisa, B., Klasen, S., and Ziegler, M. (2013). Gender Inequality in Social Institutions and Gendered Development Outcomes. *World Development*, 45:252–268.
- Branson, N. and Wittenberg, M. (2014). Reweighting South African National Household Survey Data to Create a Consistent Series Over Time: A Cross-Entropy Estimation Approach. *South African Journal of Economics*, 82(1):19–38.
- Brollo, F. and Troiano, U. (2016). What Happens When a Woman Wins an Election? Evidence from Close Races in Brazil. *Journal of Development Economics*, 122:28–45.
- Burger, R. and Yu, D. (2006). Wage Trends in Post-Apartheid South Africa: Constructing an Earnings Series from Household Survey Data. Stellenbosch Economic Working Papers: 10/06, Stellenbosch University.
- Calero, C., Bedi, A. S., and Sparrow, R. (2009). Remittances, Liquidity Constraints and Human Capital Investments in Ecuador. *World Development*, 37(6):1143–1154.
- Cameron, L., Erkal, N., Gangadharan, L., and Zhang, M. (2015). Cultural Integration: Experimental Evidence of Convergence in Immigrants' Preferences. *Journal of Economic Behavior & Organization*, 111:38–58.
- Carasciuc, L. (2001). Corruption and Quality of Governance: the Case of Moldova. Monograph, Transparency International.

- Careja, R. and Emmenegger, P. (2012). Making Democratic Citizens The Effects of Migration Experience on Political Attitudes in Central and Eastern Europe. *Comparative Political Studies*, 45(7):875–902.
- Carmichael, S. G. (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. G., de Pleijt, A., van Zanden, J. L., and De Moor, T. (2016). The European Marriage Pattern and Its Measurement. *The Journal of Economic History*, 76(1):196–204.
- Cashdan, E. (2014). Biogeography of Human Infectious Diseases: A Global Historical Analysis. *PLoS ONE*, 9(10):1–11.
- Cavalcanti, T. and Tavares, J. (2016). The Output Cost of Gender Discrimination: A Model-based Macroeconomics Estimate. *Economic Journal*, 126(590):109–134.
- Cavalcanti, T. V. d. V. and Tavares, J. (2008). Assessing the “Engines of Liberation”: Home Appliances and Female Labor Force Participation. *The Review of Economics and Statistics*, 90(1):81–88.
- Chauvet, L. and Mercier, M. (2014). Do Return Migrants Transfer Political Norms to Their Origin Country? Evidence from Mali. *Journal of Comparative Economics*, 42(3):630–651.
- Chiappori, P.-A. (1988). Rational Household Labor Supply. *Econometrica*, 56(1):63–90.
- Chiappori, P.-A. (1992). Collective Labor Supply and Welfare. *Journal of Political Economy*, 100(3):437–467.
- Cinnirella, F., Klemp, M., and Weisdorf, J. (2017). Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England. *Demography*, 54(2):413–436.
- Coen-Pirani, D., Len, A., and Lugauer, S. (2010). The Effect of Household Appliances on Female Labor Force Participation: Evidence from Microdata. *Labour Economics*, 17(3):503–513.
- Coleman, I. (2004). The Payoff from Women’s Rights. *Foreign Affairs*, 83(3):80–95.
- Corbacho, A., Gingerich, D. W., Oliveros, V., and Ruiz-Vega, M. (2016). Corruption as a Self-Fulfilling Prophecy: Evidence from a Survey Experiment in Costa Rica. *American Journal of Political Science*, 60(4):1077–1092.
- Cornell, L. L. (1987). Hajnal and the Household in Asia: a Comparativist History of the Family in Preindustrial Japan, 1600–1870. *Journal of Family History*, 12(1-3):143–162.
- Cortes, P. (2015). The Feminization of International Migration and its Effects on the Children Left Behind: Evidence from the Philippines. *World Development*, 65:62–78.
- Cuberes, D. and Teignier, M. (2014). Gender Inequality and Economic Growth: a Critical Review. *Journal of International Development*, 26(2):260–276.

- Cuberes, D. and Teignier, M. (2016). Aggregate Effects of Gender Gaps in the Labor Market: A Quantitative Estimate. *Journal of Human Capital*, 10(1):1–32.
- Cuberes, D. and Teignier, M. (2018). Macroeconomic Costs of Gender Gaps in a Model with Entrepreneurship and Household Production. *B.E. Journal of Macroeconomics*. forthcoming.
- Davis, N. Z. (1977). Ghosts, Kin, and Progeny: Some Features of Family Life in Early Modern France. *Daedalus*, 106(2):87–114.
- De la Croix, D. and Vander Donckt, M. (2010). Would Empowering Women Initiate the Demographic Transition in Least Developed Countries? *Journal of Human Capital*, 4(2):85–129.
- De Moor, T. and Van Zanden, J. L. (2010). Girl Power: The European Marriage Pattern and Labour Markets in the North Sea Region in the Late Medieval and Early Modern Period. *Economic History Review*, 63(1):1–33.
- Deaton, A. (2013). *The Great Escape: Health, Wealth, and the Origins of Inequality*. Princeton University Press, Princeton, NJ.
- Dennison, T. and Ogilvie, S. (2014). Does the European Marriage Pattern Explain Economic Growth? *The Journal of Economic History*, 74(3):651–693.
- Diamond, J. (1997). *Guns, Germs, and Steel: the Fates of Human Societies*. WW Norton & Company, New York.
- Diebolt, C. and Perrin, F. (2013). From Stagnation to Sustained Growth: The Role of Female Empowerment. *American Economic Review*, 103(3):545–549.
- Dilli, S. (2016). Family Systems and the Historical Roots of Global Gaps in Democracy. *Economic History of Developing Regions*, 31(1):82–135.
- Dilli, S., Rijpma, A., and Carmichael, S. G. (2015). Achieving Gender Equality: Development Versus Historical Legacies. *CESifo Economic Studies*, 61(1):301–334.
- Docquier, F. and Rapoport, H. (2012). Globalization, Brain Drain, and Development. *Journal of Economic Literature*, 50(3):681–730.
- Doepke, M. and Kindermann, F. (2016). Bargaining over Babies: Theory, Evidence, and Policy Implications. NBER Working Paper 22072, National Bureau of Economic Research.
- Doepke, M. and Tertilt, M. (2009). Women’s Liberation: What’s in It for Men? *Quarterly Journal of Economics*, 124(4):1541–1591.
- Doepke, M. and Tertilt, M. (2014). Does Female Empowerment Promote Economic Development? NBER Working Paper 19888, National Bureau of Economic Research.

- Doepke, M. and Tertilt, M. (2016). Chapter 23 - Families in Macroeconomics. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2, pages 1789–1891. Elsevier, Amsterdam.
- Doepke, M., Tertilt, M., and Voena, A. (2012). The Economics and Politics of Women's Rights. *Annual Review of Economics*, 4(1):339–372.
- Dollar, D., Fisman, R., and 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. and Gatti, R. (1999). Gender Inequality, Income, and Growth: Are Good Times Good for Women? Policy Research Report on Gender and Development Working Paper Series No. 1, The World Bank.
- Dong, B., Dulleck, U., and Torgler, B. (2012). Conditional Corruption. *Journal of Economic Psychology*, 33(3):609–627.
- Duflo, E. (2012). Women, Empowerment, and Economic Development. *Journal of Economic Literature*, 50(4):1051–1079.
- Duranton, G., Rodríguez-Pose, A., and Sandall, R. (2009). Family Types and the Persistence of Regional Disparities in Europe. *Economic Geography*, 85(1):23–47.
- Dyble, M., Salali, G. D., Chaudhary, N., Page, A., Smith, D., Thompson, J., Vinicius, L., Mace, R., and Migliano, A. B. (2015). Sex Equality Can Explain the Unique Social Structure of Hunter-Gatherer Bands. *Science*, 348(6236):796–798.
- Ebenstein, A. (2014). Patrilocality and Missing Women. Mimeo, Hebrew University.
- Echevarria, C. and Moe, K. S. (2000). On the Need for Gender in Dynamic Models. *Feminist Economics*, 6(2):77–96.
- Edlund, L. and Lagerlöf, N.-P. (2006). Individual Versus Parental Consent in Marriage: Implications for Intra-Household Resource Allocation and Growth. *American Economic Review*, 96(2):304–307.
- Emran, M. S., Shilpi, F., and Islam, A. (2013). Admission is Free Only If Your Dad is Rich!: Distributional Effects of Corruption in Schools in Developing Countries. World Bank Policy Research Working Paper 6671, The World Bank.
- Engels, F. (1902). *The Origin of the Family, Private Property, and the State*. Charles H. Kerr & Company Cooperative, Chicago.
- ESP/NEPC (2010). *Drawing the Line: Parental Informal Payments for Education across Eurasia*. Education Support Program (ESP) of the Open Society Institute, Budapest.

- Esteve-Volart, B. (2004). Gender Discrimination and Growth: Theory and Evidence from India. STICERD Research Paper No. DEDPS42, London School of Economics and Political Science.
- Eswaran, M. and Malhotra, N. (2011). Domestic Violence and Women's Autonomy in Developing Countries: Theory and Evidence. *Canadian Journal of Economics*, 44(4):1222–1263.
- Fairlie, R. W. (2006). An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models. IZA Discussion Paper Series No. 1917, IZA Institute of Labor Economics.
- Fehr, E. and Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *The American Economic Review*, 90(4):980–994.
- Fincher, C. L., Thornhill, R., Murray, D. R., and Schaller, M. (2008). Pathogen Prevalence Predicts Human Cross-cultural Variability in Individualism/Collectivism. *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1640):1279–1285.
- Finlay, K. and Magnusson, L. M. (2009). Implementing Weak-instrument Robust Tests for a General Class of Instrumental-variables Models. *Stata Journal*, 9(3):1–26.
- Fish, M. S. (2002). Islam and Authoritarianism. *World Politics*, 55(1):4–37.
- Fortin, N., Lemieux, T., and Firpo, S. (2011). Chapter 1 - Decomposition Methods in Economics. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4, Part A, pages 1–102. Elsevier, Amsterdam.
- Gaddis, I. and Klasen, S. (2014). Economic Development, Structural Change, and Women's Labor Force Participation. *Journal of Population Economics*, 27(3):639–681.
- Gallup, J. L., Mellinger, A. D., and Sachs, J. D. (2010). Geography Datasets. Center for International Development, Harvard University.
- Galor, O. and Weil, D. N. (1996). The Gender Gap, Fertility, and Growth. *American Economic Review*, 86(3):374–387.
- Ganguli, I., Hausmann, R., and Viarengo, M. (2014). Closing the Gender Gap in Education: What is the State of Gaps in Labor Force Participation for Women, Wives and Mothers? *International Labour Review*, 153(2):173–207.
- Gasparini, L. and Marchionni, M. (2015). *Bridging Gender Gaps? The Rise and Deceleration of Female Labor Force Participation in Latin America*. Centre for Distributive Labor and Social Studies (CEDLAS), Universidad Nacional de La Plata, La Plata, Argentina.
- Giuliano, P. (2007). Living Arrangements in Western Europe: Does Cultural Origin Matter? *Journal of the European Economic Association*, 5(5):927–952.

- Giuliano, P. (2015). The Role of Women in Society: from Preindustrial to Modern Times. *CESifo Economic Studies*, 61(1):33–52.
- Goetz, A. M. (2007). Political Cleaners: Women as the New Anti-corruption Force? *Development and Change*, 38(1):87–105.
- Goldin, C. (1990). *Understanding the Gender Gap: An Economic History of American Women*. Oxford University Press.
- Goldin, C. (1991). The Role of World War II in the Rise of Women's Employment. *American Economic Review*, 81(4):741–756.
- Goldin, C. (1995). The U-Shaped Female Labor Force Function in Economic Development and Economic History. In Schultz, T. P., editor, *Investment in Womens Human Capital and Economic Development*, pages 61–90. University of Chicago Press, Chicago.
- Goldin, C. (2006). The Quiet Revolution That Transformed Women's Employment, Education, and Family. *The American Economic Review*, 96(2):1–21.
- Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119.
- Goldin, C. and 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.
- Gould, E. D., Moav, O., and Simhon, A. (2008). The Mystery of Monogamy. *American Economic Review*, 98(1):333–57.
- Greenwood, J., Guner, N., and Vandenbroucke, G. (2017). Family Economics Writ Large. *Journal of Economic Literature*, 55(4):1346–1434.
- Greenwood, J., Seshadri, A., and Yorukoglu, M. (2005). Engines of Liberation. *Review of Economic Studies*, 72(1):109–133.
- Grimm, M. (2003). Family and Economic Growth: a Review. *Mathematical Population Studies*, 10(3):145–173.
- Guernier, V., Hochberg, M. E., and Guban, J.-F. (2004). Ecology Drives the Worldwide Distribution of Human Diseases. *PLOS Biology*, 2(6):e141.
- Haber, S. (2012). Climate, Technology, and the Evolution of Economic and Political Institutions. PERC Research Paper No. 12/2, The Property and Environment Research Center.
- Hajnal, J. (1965). European Marriage Patterns in Perspective. In Glass, D. V. and Eversley, D. E. C., editors, *Population in History: Essays in Historical Demography*, chapter 6, pages 101–143. Edward Arnold Ltd, London.

- Hajnal, J. (1982). Two Kinds of Preindustrial Household Formation System. *Population and Development Review*, 8(3):449–494.
- Hallward-Driemeier, M., Hasan, T., and Rusu, A. B. (2013). Women’s Legal Rights over 50 Years: Progress, Stagnation or Regression? World Bank Policy Research Working Paper 6616, The World Bank.
- Hansen, C. W., Jensen, P. S., and Skovsgaard, C. V. (2015). Modern Gender Roles and Agricultural History: the Neolithic Inheritance. *Journal of Economic Growth*, 20(4):365–404.
- Hartman, M. S. (2004). *The Household and the Making of History: A Subversive View of the Western Past*. Cambridge University Press, Cambridge.
- Hazarika, G., Jha, C. K., and Sarangi, S. (2015). The Role of Historical Resource Scarcity in Modern Gender Inequality. Working Paper 2015-06, Louisiana State University.
- Heim, B. T. (2007). The Incredible Shrinking Elasticities: Married Female Labor Supply, 1978–2002. *Journal of Human Resources*, 42(4):881–918.
- Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, 102(2):994–1028.
- Herlihy, D. (1985). *Medieval Households*. Cambridge University Press, Cambridge.
- Heyneman, S. P., Anderson, K. H., and Nuraliyeva, N. (2008). The Cost of Corruption in Higher Education. *Comparative Education Review*, 52(1):1–25.
- Hiller, V. (2014). Gender Inequality, Endogenous Cultural Norms, and Economic Development. *Scandinavian Journal of Economics*, 116(2):455–481.
- Houtenville, A. J. and Conway, K. S. (2008). Parental Effort, School Resources, and Student Achievement. *Journal of Human Resources*, 43(2):437–453.
- IBGE (2015). Nota Técnica: Principais Diferenças Metodológicas Entre as Pesquisas PME, PNAD e PNAD Contínua. Instituto Brasileiro de Geografia e Estatística (IBGE), Rio de Janeiro.
- ILO (2017). *World Employment and Social Outlook: Trends for women 2017*. International Labour Office, Geneva.
- Inglehart, R. and Norris, P. (2003). The True Clash of Civilizations. *Foreign Policy*, 135(Mar.-Apr.):62–70.
- Ivlevs, A. and King, R. M. (2017). Does Emigration Reduce Corruption? *Public Choice*, 171(3):389–408.
- Jayachandran, S. (2015). The Roots of Gender Inequality in Developing Countries. *Annual Review of Economics*, 7(1):63–88.



- Jones, E. L. (1981). *The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia*. Cambridge University Press, Cambridge.
- Kabeer, N. (2016). Gender Equality, Economic Growth, and Women's Agency: the Endless Variety and Monotonous Similarity of Patriarchal Constraints. *Feminist Economics*, 22(1):295–321.
- Kabeer, N. and Natali, L. (2013). Gender Equality and Economic Growth: Is there a Win-Win? IDS Working Papers 417, Institute of Development Studies.
- Kaufmann, D., Kraay, A., and Mastruzzi, M. (2011). The Worldwide Governance Indicators: Methodology and Analytical issues. *Hague Journal on the Rule of Law*, 3(2):220–246.
- Kerr, A., Lam, D., and Wittenberg, M. (2016). Post-Apartheid Labour Market Series [dataset]. Version 3.1. Cape Town: DataFirst [producer and distributor].
- Kerr, A. and Wittenberg, M. (2016). A Guide to Version 3.1 of the Post-Apartheid Labour Market Series (PALMS).
- Kimura, M. and Yasui, D. (2010). The Galor-Weil Gender-gap Model Revisited: from Home to Market. *Journal of Economic Growth*, 15:323–351.
- Klasen, S. (1998). Marriage, Bargaining, and Intrahousehold Resource Allocation: Excess Female Mortality Among Adults During Early German Development, 1740–1860. *The Journal of Economic History*, 58(2):432–467.
- 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. (2003). Weibliche Übersterblichkeit in Entwicklungsländern: Eine ökonomische Analyse. In *Neuere Ansätze der theoretischen und empirischen Entwicklungsländerforschung*, pages 257–282. Duncker und Humblot, Berlin.
- Klasen, S. (2018a). The Impact of Gender Inequality on Economic Performance in Developing Countries. *Annual Review of Resource Economics*, 10. forthcoming.
- Klasen, S. (2018b). What Explains Uneven Female Labor Force Participation Levels and Trends in Developing Countries? Courant Research Centre–Poverty, Equity and Growth Discussion Paper No. 246, University of Goettingen.
- Klasen, S. and 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. and Pieters, J. (2015). What Explains the Stagnation of Female Labor Force Participation in Urban India? *World Bank Economic Review*, 29(3):449–478.

- Klein Goldewijk, K., Beusen, A., and Janssen, P. (2010). Long-term Dynamic Modeling of Global Population and Built-up Area in a Spatially Explicit Way: HYDE 3.1. *The Holocene*, 20(4):565–573.
- Kleven, H. and Landais, C. (2017). Gender Inequality and Economic Development: Fertility, Education and Norms. *Economica*, 84(334):180–209.
- Knowles, B. S., Lorgelly, P. K., and Owen, P. D. (2002). Are Educational Gender Gaps a Brake on Economic Development? Some Cross-country Empirical Evidence. *Oxford Economic Papers*, 54(1):118–149.
- Kreibaum, M. and Klasen, S. (2015). Missing Men: Differential Effects of War and Socialism on Female Labour Force Participation in Vietnam. Courant Research Centre-Poverty, Equity and Growth Discussion Paper No. 181, University of Goettingen.
- La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The Quality of Government. *Journal of Law, Economics, and Organization*, 15(1):222–279.
- Lagerlöf, N.-P. (2003). Gender Equality and Long-Run Growth. *Journal of Economic Growth*, 8(4):403–426.
- Lagerlöf, N.-P. (2005). Sex, Equality, and Growth. *Canadian Journal of Economics*, 38(3):807–831.
- Levitt, P. (1998). Social Remittances: Migration Driven Local-Level Forms of Cultural Diffusion. *International Migration Review*, 32(4):926–948.
- Levitt, P. (2001). *The Transnational Villagers*. University of California Press, Berkeley, California.
- Lucas, R. E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(1):3–42.
- Lücke, M., Omar Mahmoud, T., and Peuker, C. (2012). Identifying the Motives of Migrant Philanthropy. Working Paper No. 1790, Kiel Institute for the World Economy.
- Mahmud, S. and Bidisha, S. H. (2016). Female Labor Market Participation in Bangladesh: Structural Changes and Determinants of Labor Supply. In Raihan, S., editor, *Structural Change and Dynamics of Labor Markets in Bangladesh: Studies on Labor and Employment*, pages 33–40. SANEM Publications, Dhaka.
- Marshall, A. (2013). *Principles of Economics*. Palgrave Macmillian, 8<sup>th</sup> edition.
- Martineau, H. (1837). *Society in America*, volume 3. Saunders & Otley, London.
- McKenzie, D. and Rapoport, H. (2010). Can Migration Reduce Educational Attainment? Evidence from Mexico. *Journal of Population Economics*, 24(4):1331–1358.

- Méon, P.-G. and Sekkat, K. (2005). Does Corruption Grease or Sand the Wheels of Growth? *Public Choice*, 122(1):69–97.
- Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., and Aiden, E. L. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, 331(6014):176–182.
- Mikusheva, A. and Poi, B. P. (2006). Tests and Confidence Sets with Correct Size When Instruments are Potentially Weak. *Stata Journal*, 6(3):335–347.
- Mitterauer, M. (2010). *Why Europe?: The Medieval Origins of Its Special Path*. University of Chicago Press, Chicago.
- Moreira, M. J. (2009). Tests With Correct Size When Instruments Can Be Arbitrarily Weak. *Journal of Econometrics*, 152(2):131–140.
- Mountford, A. (1997). Can a Brain Drain Be Good for Growth in the Source Economy? *Journal of Development Economics*, 53(2):287–303.
- MPC (2013). Migration Profile Moldova. Migration Policy Centre: Country Report, European University Institute.
- Mulder, M. B. and Rauch, K. L. (2009). Sexual Conflict in Humans: Variations and Solutions. *Evolutionary Anthropology: Issues, News, and Reviews*, 18(5):201–214.
- Murray, D. R. and Schaller, M. (2010). Historical Prevalence of Infectious Diseases Within 230 Geopolitical Regions: A Tool for Investigating Origins of Culture. *Journal of Cross-Cultural Psychology*, 41(1):99–108.
- Ntuli, M. and Wittenberg, M. (2013). Determinants of Black Women’s Labour Force Participation in Post-Apartheid South Africa. *Journal of African Economies*, 22(3):347–374.
- Nunn, N. and Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics*, 94(1):20–36.
- OAMDI (2015a). Harmonized Labor Force Surveys (HLFS), <http://erf.org.eg/data-portal/>. Version 1.0 of Licensed Data Files; EUS 2006 - Department of Statistics (DOS), The Hashemite Kingdom of Jordan. Egypt: Economic Research Forum (ERF).
- OAMDI (2015b). Harmonized Labor Force Surveys (HLFS), <http://erf.org.eg/data-portal/>. Version 1.0 of Licensed Data Files; EUS 2008 - Department of Statistics (DOS), The Hashemite Kingdom of Jordan. Egypt: Economic Research Forum (ERF).
- OAMDI (2015c). Harmonized Labor Force Surveys (HLFS), <http://erf.org.eg/data-portal/>. Version 1.0 of Licensed Data Files; EUS 2010 - Department of Statistics (DOS), The Hashemite Kingdom of Jordan. Egypt: Economic Research Forum (ERF).

- OAMDI (2015d). Harmonized Labor Force Surveys (HLFS), <http://erf.org.eg/data-portal/>. Version 1.0 of Licensed Data Files; EUS 2014 - Department of Statistics (DOS), The Hashemite Kingdom of Jordan. Egypt: Economic Research Forum (ERF).
- Olsson, O. and Paik, C. (2016). Long-run Cultural Divergence: Evidence from the Neolithic Revolution. *Journal of Development Economics*, 122:197–213.
- Oostendorp, R. H. (2009). Globalization and the Gender Wage Gap. *World Bank Economic Review*, 23(1):141–161.
- Osipian, A. L. (2009). Corruption Hierarchies in Higher Education in the Former Soviet Bloc. *International Journal of Educational Development*, 29(3):321–330.
- Oster, E. (2017). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*. forthcoming.
- Parker, P. M. (2000). *Physioeconomics: The Basis for Long-run Economic Growth*. MIT Press, Cambridge, Massachusetts.
- Peel, M. C., Finlayson, B. L., and McMahon, T. A. (2007). Updated World Map of the Köppen-Geiger Climate Classification. *Hydrology and Earth System Sciences Discussions*, 4(2):439–473.
- Penn, D. J. and Smith, K. R. (2007). Differential Fitness Costs of Reproduction Between the Sexes. *Proceedings of the National Academy of Sciences*, 104(2):553–558.
- Poos, L. R. (1991). *A Rural Society After the Black Death: Essex: 1350–1525*. Cambridge University Press, Cambridge.
- Powelson, J. P. (1994). *Centuries of Economic Endeavor: Parallel Paths in Japan and Europe and Their Contrast with the Third World*. University of Michigan Press, Ann Arbor, Michigan.
- Prettner, K. and Strulik, H. (2017). Gender Equity and the Escape from Poverty. *Oxford Economic Papers*, 69(1):55–74.
- Priebe, J. (2010). Child Costs and the Causal Effect of Fertility on Female Labor Supply: An Investigation for Indonesia 1993–2008. Courant Research Centre-Poverty, Equity and Growth Discussion Paper No. 45, University of Goettingen.
- Putterman, L. and Trainor, C. A. (2006). Agricultural Transition Year Country Data Set. Online resource.
- Putterman, L. and Weil, D. N. (2010). Post-1500 Population Flows and the Long-run Determinants of Economic Growth and Inequality. *Quarterly Journal of Economics*, 125(4):1627–1682.

- Rahman, R. I. and Islam, R. (2013). Female Labour Force Participation in Bangladesh: Trends, Drivers and Barriers. ILO Asia-Pacific Working Paper Series, International Labour Organization.
- Rees, R. and Riezman, R. (2012). Globalization, Gender, and Growth. *Review of Income and Wealth*, 58(1):107–117.
- Reher, D. S. (1998). Family Ties in Western Europe: Persistent Contrasts. *Population and Development Review*, 24(2):203–234.
- Rijpma, A. and Carmichael, S. G. (2016). Testing Todd and Matching Murdock: Global Data on Historical Family Characteristics. *Economic History of Developing Regions*, 31(1):10–46.
- Rosenzweig, M. and Wolpin, K. (1980). Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica*, 48(1):227–240.
- Ross, M. L. (2008). Oil, Islam, and Women. *American Political Science Review*, 102(1):107–123.
- Ruggles, S. (2009). Reconsidering the Northwest European Family System: Living Arrangements of the Aged in Comparative Historical Perspective. *Population and Development Review*, 35(2):249–273.
- Schaner, S. and Das, S. (2016). Female Labor Force Participation in Asia: Indonesia Country Study. ADB Economics Working Paper Series No. 474, Asian Development Bank.
- Seguino, S. (2000). Gender Inequality and Economic Growth: A Cross-Country Analysis. *World Development*, 28(7):1211–1230.
- Seneviratne, P. (2017). Economic Development or Economic Distress? Explaining the Rise in Married Women's Labor Force Participation in Sri Lanka. Mimeographed, Carleton College.
- Shleifer, A. and Vishny, R. W. (1993). Corruption. *Quarterly Journal of Economics*, 108(3):599–617.
- Sinha, N., Raju, D., and Morrison, A. (2007). Gender Equality, Poverty and Economic Growth. World Bank Policy Research Paper 4349, The World Bank.
- Smith, R. M. (1981). Fertility, Economy, and Household Formation in England over Three Centuries. *Population and Development Review*, 7(4):595–622.
- Spilimbergo, A. (2009). Democracy and Foreign Education. *American Economic Review*, 99(1):528–543.
- Spolaore, E. and Wacziarg, R. (2009). The Diffusion of Development. *Quarterly Journal of Economics*, 124(2):469–529.

- Spolaore, E. and Wacziarg, R. (2013). How Deep Are the Roots of Economic Development? *Journal of Economic Literature*, 51(2):325–369.
- Stotsky, J. G. (2006). Gender and Its Relevance to Macroeconomic Policy: A Survey. IMF Working Paper 06/233, International Monetary Fund.
- Strulik, H. (2018). Desire and Development. *Macroeconomic Dynamics*. forthcoming.
- Swamy, A., Knack, S., Lee, Y., and Azfar, O. (2001). Gender and Corruption. *Journal of Development Economics*, 64(1):25–55.
- Szołtysek, M. (2014). Toward a Conceptual Framework for the Variation in Historical Family and Household Systems Across Eurasia. *Przeszłość Demograficzna Polski*, 36:55–86.
- Szołtysek, M., Klüsener, S., Poniak, R., and Gruber, S. (2017). The Patriarchy Index: A New Measure of Gender and Generational Inequalities in the Past. *Cross-Cultural Research*, 51(3):228–262.
- Tertilt, M. (2005). Polygyny, Fertility, and Savings. *Journal of Political Economy*, 113(6):1341–1371.
- Tertilt, M. (2006). Polygyny, Women’s Rights, and Development. *Journal of the European Economic Association*, 4(2-3):523–530.
- Thompson, G. H. (1951). *The Factorial Analysis of Human Ability*. University of London Press, London.
- Todd, E. (1985). *The Explanation of Ideology: Family Structures and Social Systems*. Blackwell, Oxford.
- Todd, E. (1987). *The Causes of Progress: Culture, Authority and Change*. Blackwell, Oxford.
- Tollefsen, A. F., Strand, H., and Buhaug, H. (2012). PRIO-GRID: A Unified Spatial Data Structure. *Journal of Peace Research*, 49(2):363–374.
- Transparency International (2013). Global Corruption Barometer: Moldova. <https://www.transparency.org/gcb2013/country/?country=moldova>. [accessed on 10.06.2015].
- Tuccio, M. and Wahba, J. (2015). Can I Have Permission to Leave the House? Return Migration and the Transfer of Gender Norms. IZA Discussion Paper No. 9216, IZA Institute of Labor Economics.
- United Nations, Department of Economic and Social Affairs, Population Division (2009). World Marriage Data 2008. (POP/DB/Marr/Rev2008).
- United Nations Development Programme (2015). *Human Development Report 2015: Work for Human Development*. New York, NY.

- Voigtländer, N. and Voth, H.-J. (2013). How the West “Invented” Fertility Restriction. *American Economic Review*, 103(6):2227–2264.
- Von Fintel, D. (2007). Dealing with Earnings Bracket Responses in Household Surveys – How Sharp are Midpoint Imputations? *South African Journal of Economics*, 75(2):293–312.
- von Rueden, C. R. and Jaeggi, A. V. (2016). Men’s Status and Reproductive Success in 33 Nonindustrial Societies: Effects of Subsistence, Marriage System, and Reproductive Strategy. *Proceedings of the National Academy of Sciences*, 113(39):10824–10829.
- Walker, M. (2011). *PISA 2009 Plus Results: Performance of 15-year-olds in Reading, Mathematics and Science for 10 Additional Participants*. Australian Council for Educational Research (ACER) Press.
- Weichselbaumer, D. and Winter-Ebmer, R. (2005). A Meta-Analysis of the International Gender Wage Gap. *Journal of Economic Surveys*, 19(3):479–511.
- Welzel, C. (2013). *Freedom Rising*. Cambridge University Press, Cambridge.
- Welzel, C. (2014). Evolution, Empowerment, and Emancipation: How Societies Climb the Freedom Ladder. *World Development*, 64:33–51.
- Whyte, M. K. (1978). *The Status of Women in Pre-Industrial Societies*. Princeton University Press, Princeton, NJ.
- Wittfogel, K. (1957). *Oriental Despotism: A Comparative Study of Total Power*. Yale University Press, New Haven, Connecticut.
- World Bank (2001). *Engendering Development Through Gender Equality in Rights, Resources, and Voice*. The World Bank and Oxford University Press.
- World Bank (2011). *World Development Report 2012: Gender Equality and Development*. The World Bank, Washington, DC.
- World Bank (2014). World Development Indicators. <http://data.worldbank.org/country/moldova>. [accessed on 10.02.2015].
- World Bank (2016). World Development Indicators. [accessed on 24.09.2016].
- Yang, D. (2008). International Migration, Remittances and Household Investment: Evidence from Philippine Migrants’ Exchange Rate Shocks. *The Economic Journal*, 118(528):591–630.
- Yu, D. (2007). The Comparability of the Statistics South Africa October Household Surveys and Labour Force Surveys. Stellenbosch Economic Working Papers: 17/07, Stellenbosch University.

- Yu, D. (2009). The Comparability of Labour Force Surveys (LFS) and Quarterly Labour Force Surveys (QLFS). Stellenbosch Economic Working Papers: 08/09, Stellenbosch University.
- Zhang, H., Behrman, J. R., Fan, C. S., Wei, X., and Zhang, J. (2014). Does Parental Absence Reduce Cognitive Achievements? Evidence from Rural China. *Journal of Development Economics*, 111:181–195.
- Zhang, J., Zhang, J., and Li, T. (1999). Gender Bias and Economic Development in an Endogenous Growth Model. *Journal of Development Economics*, 59(2):497–525.