QOG

UNIVERSITY OF GOTHENBURG

THE ROOTS OF FEMALE EMANCIPATION

The Initializing Role of Cool Water

MANUEL SANTOS SILVA

AMY C. ALEXANDER

STEPHAN KLASEN

CHRISTIAN WELZEL

WORKING PAPER SERIES 2020:14

QOG THE QUALITY OF GOVERNMENT INSTITUTE
Department of Political Science
University of Gothenburg
Box 711, SE 405 30 GÖTEBORG
December 2020
ISSN 1653-8919

The Roots of Female Emancipation: The Initializing Role of Cool Water[1]
Manuel Santos Silva
Amy C. Alexander
Stephan Klasen
Christian Welzel
QoG Working Paper Series 2020:14
December 2020
ISSN 1653-8919

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.

Manuel Santos Silva
University of Münster
manuel.santos.silva@unimuenster.de

Amy C. Alexander
The Quality of Government Institute
Department of Political Science
University of Gothenburg
Amy.alexander@gu.se

Stephan Klasen
University of Goettingen
sklasen@uni-goettingen.de

Christian Welzel
Leuphana University, Lueneburg
National Research University – Higher School of Economics, Moscow
cwelzel@gmail.com

[1] We would like to thank Laura Barros, Lisa Höckel, Krisztina Kis-Katos, Susan Steiner, Holger Strulik, Mikołaj Szołtysek, 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.
M. Santos Silva: manuel.santos.silva@unimuenster.de; A. Alexander: amy.catherine.alexander@gmail.com; C. Welzel: cwelzel@gmail.com.

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; Voigtländer and Voth, 2013; Prettner 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; BenYishay et al., 2017; Grosjean and Khattar, 2019).

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., 2019).[2]

Building on this scholarship, we theorize and provide evidence for an overlooked trajectory originating in a climatic configuration, called the "Cool Water" (CW-) condition, that leads to late female marriages in preindustrial times, which in turn result in 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

[2] Carranza (2014) shows, for India, that soil textures have a *contemporary* effect on gender equality. Relative to clayey soil textures, loamy soil allows for deep tillage, and deep tillage substitutes for tasks typically performed by women (e.g., weeding, fertilizing). Indian districts with a larger fraction of loamy soil textures display rural gender ratios more skewed towards boys and lower relative female employment.

challenges and opportunities ingrained in its climate. The CW condition embodies opportunity endowments that significantly reduced fertility pressures on women and 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.2[3]

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 robust 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 of 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 labor force participation and life expectancy.

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, there is always the risk that unobserved country characteristics might account for the relationship between CW and its hypothesized outcomes. In order to address

[3] 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); Szołtysek 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); Tur-Prats (2019).

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 (2019), we estimate how large the effect of unobservables needs to be to explain away the 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.

Finally, we show that, at the turn of the twentieth century, the CW-condition of the country of origin correlates negatively with the marriage probability among female immigrants (first and second generation) to the US. These individual-level results suggest that the effect of the CW-condition has persisted over time because it became encoded as a cultural norm. Moreover, since immigrants living in a US state face the same institutional environment, the results further alleviate the concern that omitted institutional factors might be driving the cross-country estimates.

The remainder of our article proceeds in the following steps. Section 2 reviews the literature and derives from this discussion our theoretical propositions. Section 3 introduces the data and variables used to demonstrate the empirical validity of our propositions. The fourth section describes the empirical strategy while section 5 presents the findings. We conclude in section 6.

Theoretical discussion

In section 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 societies, highlighting the distinction between those where women marry at relatively later ages vis-à-vis those where they marry immediately after puberty (section 2.2). In late marriage societies, women are relatively more empowered (section 2.3). Finally, in section 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.

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 pressure 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 positively correlated the sex ratio at birth (boys/girls). 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. (2019) 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 positively correlated with 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.

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 times3[4]: (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)

[4] 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). See Hartman (2004, chapter 3) for a review of the historical evidence. Edwards and Ogilvie (2018, pp. 2–6) survey the evidence for England.

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.

Of course, 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, 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 to 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.[5] 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.

[5] For historical evidence on the contractual household system from mid-fourteen century England see Poos (1991). For older, yet more scattered, evidence from ninth century France (around Paris) see Herlihy (1985).

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) revisits and extends his analysis.

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.

[Figure 1 about here]

Third, the institution of contractual farm service is a driver of gender equality norms and behaviors in its own right. As shown in Figure 1, contractual service was a common phase in the life-cycle of *both* men and women across Northwestern Europe.[6] 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, the divergence of life experiences early on in the Middle East, India or China, 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), "While in service, women were not under the control of any male relative. 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."[7]

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

[6] 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).

[7] A recent study on eighteen-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).

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).

To avoid misunderstandings, we are by no means claiming that late female marriage ages created anything close to perfect gender equality. 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 generated opportunities for female individualization that were not available for their early marriage counterparts.[8] 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.

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

[8] 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.

form in an agrarian economy should be a response to the challenges and opportunities embodied in this economy's climatic conditions.

Welzel (2013, 2014) summarizes the particular thermo-hydrological constellation of Northwestern Europe's climate as the "Cool Water" (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*.[9]

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.[10] 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.

[9] See Welzel (2013, 2014) for broader implications of the CW-condition for long-run development. Here, the discussion focuses exclusively on household formation patterns.

[10] 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).

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 determined by infectious diseases (e.g., Deaton, 2013, pp. 81–87).[11] 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).[12]

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

[11] Of course, Northwest European countries were early adopters of modern public health innovations, such as vaccinations, from the second half of the eighteen-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 1750–1960, 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.

[12] 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 areas with high pathogen prevalence, preferences for extended households based on kinship were relatively stronger.

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 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.[13] 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, making 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

[13] In thirteenth- and fourteenth-century England, for example, the labor inputs, per unit of land, required for pastoral agriculture were one-fifth of those required for arable cultivation (Campbell, 2000, p. 10).

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).[14] Thus, if the Black Death was indeed the event responsible for a further push towards the late marriage pattern[15], 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, 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

[14] In regions where animal husbandry depends on great distance traveling, this activity becomes incompatible with childrearing and women lose their comparative advantage.
[15] See Dennison and Ogilvie (2014, p. 673) and Edwards and Ogilvie (2018) for a rebuttal of this claim.

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.[16]

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, representing 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.

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 A1.

---

[16] 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).

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 EAPEP 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 1) with the period 1960–1980 accounting for 80 percent of the observations.[17]

[Table 1 about here]

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 on historical demography and harmonize the data by regressing the marital ages on several characteristics of the sources from which they were extracted.[18] 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 overall trends.[19] Table 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

[17] 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 4; instead we use data for the next available year, 1982.

[18] Carmichael et al. (2016, p. 200) discuss limitations of these data.

[19] 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).

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.[20] We interpret these data with caution, since they are collected from a variety of sources and, in some cases, supplemented by qualitative adjustments or backward extrapolations by Gapminder.[21]

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 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.

[20] Armenia, Azerbaijan, Bangladesh, China, Egypt, Georgia, India, Japan, Kazakhstan, Pakistan, Sri Lanka, and United States.
[21] For more details, see the original documentation of these data at https://www.gapminder.org/wpcontent/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 Online Appendix.

To capture the CW-condition's thermal aspect, we take latitude as the starting point, using each country-centroid's latitude in degrees.[22] 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 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[23], summer heat is uncorrelated with latitude.[24] 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[25] 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

---

[22] 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.
[23] The country-centroids' latitudes correlate with the countries' annual mean temperatures at $r = -0.89$ (N=183; p-value = 0.001, 2-tailed).
[24] $r = 0.07$ (N=183; p-value = 0.336, 2-tailed).
[25] Average annual rainfall per month and latitude correlate at $r = -0.53$ (N=177; p-value = 0.00, 2-tailed).

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.[26] 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, 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).[27]

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).[28] 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

[26] Continuous rain in this definition is literally uncorrelated with average annual rainfall per month.
[27] 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.
[28] 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).

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.[29]

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.[30]

We also acknowledge that 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. In the Online Appendix, we deal with this issue.[31] 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.[32]

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

---

[29] 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.
[30] 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).
[31] 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.
[32] For a world distribution of the CW-index see Figure 5.

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.[33]

[Figures 2 and 3 about here]

For Europe, we find that ages at first marriage for women have persisted over centuries. Figure 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). 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.

Empirical strategy

We test our hypotheses by estimating regressions of the form:

$$y_i = \alpha + H_i' \beta + X_i^{H\prime} \gamma + X_i^{C\prime} \delta + \theta_c + \epsilon_i \quad (1)$$

where y_i is the outcome variable of interest for country *i*; $\boldsymbol{H_i}$ is a vector of potential deep determinants—*CW-index*, *years of agriculture*, *plow*, and *agricultural suitability*; $\boldsymbol{X_i^H}$ is a vector of historical controls of country *i*'s ancestral population; $\boldsymbol{X_i^C}$ is a vector of contemporary control variables; θ_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

[33] 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 A1.

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 geo-climatic 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, 2019). 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.

Results

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 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.[34] 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.[35] 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.[36]

[Table 2 about here]

As shown in Table 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 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.[37]

[34] 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.

[35] 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.

[36] 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.

[37] If we exclude the MENA countries for the regression in Table 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.

Overall, the estimated coefficients of Table 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. (2019). 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.

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 A2.

We explore the sensitivity of these results to subsample analyses (Table A3). Excluding Sub-Saharan Africa, the Americas and Oceania (i.e., the New World), or Europe 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.

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 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 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-month reduction of the average gap between male and female ages at first marriage.[38] 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 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.[39] 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.[40] The results hold for the usual three subsamples: without Sub-Saharan Africa, without the New World, and without Europe, as shown in Table A5.

[Table 3 about here]

[38] $\widehat{\beta_{cw}} * \sigma_{cw} = -5.902 * 0.145 \approx -0.856\, years\ (\approx -10.27\, months)$.

[39] 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.

[40] In Table A4, 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).

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 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 A6 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 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.[41] 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* weighting decisions. Second, we show that our results are not dependent on

[41] 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.

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 A7).[42] 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 4, 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 coefficient 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 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 A5 and A7 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).[43] 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,

[42] 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.
[43] 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).

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; 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.

[Table 4 about here]

In addition to the regressions of Table 4, we perform further robustness checks and present the results in Table A8 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)[44] to account for the possibility that animal husbandry delays ages at marriage 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

[44] Unless otherwise noted, all additional variables included are from Alesina et al. (2013). For the original source and construction method see their Online Appendix.

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 A8).[45] Overall, after controlling for many 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 (δ) 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, 2019). Table A9 presents the estimates for δ, using the method described in Oster (2019).[46] Altonji et al. (2005) propose $\delta = 1$ as an *ad hoc* cutoff for robustness. For reasonable upper bounds on the R^2 and models that exclude religious variables—as recommended by most of the literature—, the estimates of δ are close to the cutoff, ranging from 0.808 to 1.009.

[45] Reestimating the specifications of Tables 4 and A5 using ancestry-adjusted CW index and years of agriculture produces even stronger CW effects, which are always highly statistically significant. Results available upon request.
[46] See section A.2 of the Online Appendix for more details.

Cool Water and historic late marriages

Preindustrial Europe We now turn to Europe to show that the relationship between the CW-condition and ages at first marriage is deep rooted in history. Between 1500–1900, for a sample of 27 European countries, the CW-index is a positive determinant of female marital ages (Table 5). Despite the much smaller sample, the CW coefficient continues to be statistically significant in all specifications, except when years of agriculture are included in columns 3 and 4. In those specifications, while the size of the CW effect is not much affected, its standard errors increase due to multicollinearity between the CW-index and the timing of the Neolithic revolution ($r \approx -0.70$) for this sample of European countries.

[Table 5 about here]

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 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 A10). In this setting, the CW-effect retains significance, even taking the timing of the Neolithic Revolution into account.[47]

First and second-generation immigrants in the US, 1880–1930 We now provide evidence, at the individual level, that the CW-condition of the country of origin correlates with an immigrant's age at marriage. We use data from five US censuses: 1880, 1900, 1910, 1920, and 1930.[48] We chose these censuses because they are the only to include mother's and father's country of birth, years since

---

[47] Once again, the results are robust to ancestry-adjusting the CW index and years of agriculture.
[48] All census samples are from IPUMS-US. We use the following samples: 1880, 10 percent sample; 1900, 5 percent sample; 1910, 1 percent sample; 1920, 1 percent sample; 1930, 5 percent sample.

immigration, and marital status. We construct two samples of women aged 15–39: (1) immigrants to the US (first generation), and (2) US-born children of immigrants, defined by either paternal or maternal foreign birthplace (second generation). We then estimate linear probability models on whether a woman has ever married, using as right-hand-side variables the same set of controls of the cross-country regressions, this time defined at the immigrant's country of origin. In addition, we control for individual characteristics (age, age squared, literacy fixed effects), US state fixed effects, and census-year fixed effects.

At arrival, female immigrants originating from areas with higher CW-condition were less likely to be married at all ages (Table 6). The negative effect of the ancestral CW weakens slightly with years lived in the US, as shown by the positive interaction effect with years since immigration.[49] This could either suggest assimilation effects or capture different marriage selectivity of migrant cohorts. The implied effect is relevant in magnitude. At age 25–29, 77 percent of female immigrants in the (unweighted) sample were ever married (column 4). At arrival in the US, one standard deviation increase in the CW-index of the country of origin is associated with a 33 percentage point reduction in the likelihood of being ever married. For those living in the US for 15 years, the effect weakens to (minus) 26 percentage points.

The negative correlation between CW and marriage probability persists for the next generation. Among female children of immigrants, the CW of their fathers' birthplace correlates negatively with marriage at all ages (Table 7). Among women born to a foreign father, 68 percent were ever married at ages 25–29 (unweighted sample average). A one standard deviation increase in the CW-index of the father's country of birth is associated with a 28 percentage point reduction in marriage probability.[50]

[Tables 6 and 7 about here]

---

[49] See Table A11 for estimates without the years since immigration interaction term.

[50] In these regressions, we also include mother birthplace fixed effects as controls. Among women born to a foreign mother, the CW-index of the mother's country of birth is also negatively correlated with the daughter's marriage probability, although the effects are generally smaller (in absolute terms) in comparison to father's ancestral CW (Table A12).

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 8, 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.

[Table 8 about here]

The CW-index is also a significant positive correlate of contemporary female-male ratio in life expectancy (Table 8, 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 8, Panel C). The only robust negative deep determinant of gender equality in education is years of agriculture.

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.

In support of these hypotheses, we provide cross-country evidence that, in pre-industrial Europe, women married later in countries with a higher CW-index. We are able to trace this effect, at the turn of the twentieth century, 1880–1930, for a sample of first and second-generation immigrants in the US. Female immigrants whose country of origin had a higher CW-condition were less likely to marry over the 15–39 age distribution. Finally, we show that the CW effect has persisted into the contemporary period. Using a cross-section of world countries, from 1960 onwards, we estimate a sizable effect of a country's CW score on female ages at first marriage. Altogether, the empirical evidence spans a long period of time, involving several independent datasets and levels of analysis (country, ancestry, individual).

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.

REFERENCES

Acemoglu, D., Autor, D.H., Lyle, D., 2004. Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury. Journal of Political Economy 112, 497–551.

Ager, P., Worm Hansen, C., Sandholt Jensen, P., 2018. Fertility and Early-Life Mortality: Evidence from Smallpox Vaccination in Sweden. Journal of the European Economic Association 16, 487–521.

Alesina, A., Giuliano, P., Nunn, N., 2013. On the Origins of Gender Roles: Women and the Plough. Quarterly Journal of Economics 128, 469–530.

Alexander, A.C., Welzel, C., 2015. Eroding Patriarchy: the Co-evolution of Women's Rights and Emancipative Values. International Review of Sociology 25, 144–165.

Altonji, J.G., Elder, T.E., Taber, C.R., 2005. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. Journal of Political Economy 113, 151–184.

Banerji, M., Martin, S., 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.

Barro, R.J., Lee, J.W., 2013. A New Data Set of Educational Attainment in the World, 1950–2010. Journal of Development Economics 104, 184–198.

Bentzen, J.S., Kaarsen, N., Wingender, A.M., 2017. Irrigation and Autocracy. Journal of the European Economic Association 15, 1–53.

BenYishay, A., Grosjean, P., Vecci, J., 2017. The Fish is the Friend of Matriliny: Reef Density and Matrilineal Inheritance. Journal of Development Economics 127, 234–249.

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, 68–92.

Branisa, B., Klasen, S., Ziegler, M., 2013. Gender Inequality in Social Institutions and Gendered Development Outcomes. World Development 45, 252–268.

Campbell, B.M.S., 2000. English Seigniorial Agriculture, 1250–1450. Cambridge University Press, Cambridge.

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, 416–436.

Carmichael, S.G., de Pleijt, A., van Zanden, J.L., De Moor, T., 2016. The European Marriage Pattern and Its Measurement. The Journal of Economic History 76, 196–204.

Carranza, E., 2014. Soil Endowments, Female Labor Force Participation, and the Demographic Deficit of Women in India. American Economic Journal: Applied Economics 6, 197–225.

Cashdan, E., 2014. Biogeography of Human Infectious Diseases: A Global Historical Analysis. PLoS ONE 9, 1–11.

Cavalcanti, T.V.d.V., Tavares, J., 2008. Assessing the "Engines of Liberation": Home Appliances and Female Labor Force Participation. The Review of Economics and Statistics 90, 81–88.

Cinnirella, F., Klemp, M., Weisdorf, J., 2017. Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England. Demography 54, 413–436.

Coen-Pirani, D., León, A., Lugauer, S., 2010. The Effect of Household Appliances on Female Labor Force Participation: Evidence from Microdata. Labour Economics 17, 503–513.

Coleman, I., 2004. The Payoff from Women's Rights. Foreign Affairs 83, 80–95.

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, 143–162.

Davis, N.Z., 1977. Ghosts, Kin, and Progeny: Some Features of Family Life in Early Modern France. Daedalus 106, 87–114.

De Moor, T., 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, 1–33.

Deaton, A., 2013. The Great Escape: Health, Wealth, and the Origins of Inequality. Princeton University Press, Princeton, NJ.

Dennison, T., Ogilvie, S., 2014. Does the European Marriage Pattern Explain Economic Growth? The Journal of Economic History 74, 651–693.

Diamond, J., 1997. Guns, Germs, and Steel: the Fates of Human Societies. WW Norton & Company, New York.

Diebolt, C., Perrin, F., 2013. From Stagnation to Sustained Growth: The Role of Female Empowerment. American Economic Review 103, 545–549.

Dilli, S., 2016. Family Systems and the Historical Roots of Global Gaps in Democracy. Economic History of Developing Regions 31, 82–135.

Dilli, S., Rijpma, A., Carmichael, S.G., 2015. Achieving Gender Equality: Development Versus Historical Legacies. CESifo Economic Studies 61, 301–334.

Doepke, M., Tertilt, M., 2009. Women's Liberation: What's in It for Men? Quarterly Journal of Economics 124, 1541–1591.

Duflo, E., 2012. Women, Empowerment, and Economic Development. Journal of Economic Literature 50, 1051–1079.

Duranton, G., Rodríguez-Pose, A., Sandall, R., 2009. Family Types and the Persistence of Regional Disparities in Europe. Economic Geography 85, 23–47.

Dyble, M., Salali, G.D., Chaudhary, N., Page, A., Smith, D., Thompson, J., Vinicius, L., Mace, R., Migliano, A.B., 2015. Sex Equality Can Explain the Unique Social Structure of Hunter-Gatherer Bands. Science 348, 796–798.

Ebenstein, A., 2014. Patrilocality and Missing Women. Mimeo. Hebrew University.

Edlund, L., Lagerlöf, N.P., 2006. Individual Versus Parental Consent in Marriage: Implications for Intra-Household Resource Allocation and Growth. American Economic Review 96, 304–307.

Edwards, J., Ogilvie, S., 2018. Did the Black Death Cause Economic Development by "Inventing" Fertility Restriction? Working Paper 7016. CESifo.

Engels, F., 1902. The Origin of the Family, Private Property, and the State. Charles H. Kerr & Company Cooperative, Chicago.

Fincher, C.L., Thornhill, R., Murray, D.R., Schaller, M., 2008. Pathogen Prevalence Predicts Human Cross-cultural Variability in Individualism/Collectivism. Proceedings of the Royal Society of London B: Biological Sciences 275, 1279–1285.

Fish, M.S., 2002. Islam and Authoritarianism. World Politics 55, 4–37.

Gallup, J.L., Mellinger, A.D., Sachs, J.D., 2010. Geography Datasets. Center for International Development, Harvard University.

Galor, O., Weil, D.N., 1996. The Gender Gap, Fertility, and Growth. American Economic Review 86, 374–387.

Giuliano, P., 2007. Living Arrangements in Western Europe: Does Cultural Origin Matter? Journal of the European Economic Association 5, 927–952.

Giuliano, P., 2015. The Role of Women in Society: from Preindustrial to Modern Times. CESifo Economic Studies 61, 33–52.

Goldin, C., 1991. The Role of World War II in the Rise of Women's Employment. American Economic Review 81, 741–756.

Goldin, C., Katz, L.F., 2002. The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. Journal of Political Economy 110, 730–770.

Greenwood, J., Seshadri, A., Yorukoglu, M., 2005. Engines of Liberation. Review of Economic Studies 72, 109–133.

Grosjean, P., Khattar, R., 2019. It's Raining Men! Hallelujah? The Long-Run Consequences of Male-Biased Sex Ratios. The Review of Economic Studies 86, 723–754.

Guernier, V., Hochberg, M.E., Guégan, J.F., 2004. Ecology Drives the Worldwide Distribution of Human Diseases. PLOS Biology 2, 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., Eversley, D.E.C. (Eds.), Population in History: Essays in Historical Demography. Edward Arnold Ltd, London. chapter 6, pp. 101–143.

Hajnal, J., 1982. Two Kinds of Preindustrial Household Formation System. Population and Development Review 8, 449–494.

Hansen, C.W., Jensen, P.S., Skovsgaard, C.V., 2015. Modern Gender Roles and Agricultural History: the Neolithic Inheritance. Journal of Economic Growth 20, 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., Sarangi, S., 2015. The Role of Historical Resource Scarcity in Modern Gender Inequality. Working Paper 2015-06. Louisiana State University.

Hazarika, G., Jha, C.K., Sarangi, S., 2019. Ancestral Ecological Endowments and Missing Women. Journal of Population Economics 32, 1101–1123.

Herlihy, D., 1985. Medieval Households. Cambridge University Press, Cambridge.

Inglehart, R., Norris, P., 2003. The True Clash of Civilizations. Foreign Policy 135, 62–70.

Jayachandran, S., 2015. The Roots of Gender Inequality in Developing Countries. Annual Review of Economics 7, 63–88.

Jones, E.L., 1981. The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia. Cambridge University Press, Cambridge.

Kaufmann, D., Kraay, A., Mastruzzi, M., 2011. The Worldwide Governance Indicators: Methodology and Analytical Issues. Hague Journal on the Rule of Law 3, 220–246.

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, 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, 345–373.

Klasen, S., 2003. Weibliche Übersterblichkeit in Entwicklungsländern: Eine ökonomische Analyse, in: Neuere Ansätze der theoretischen und empirischen Entwicklungsländerforschung. Duncker und Humblot, Berlin, pp. 257–282.

Klasen, S., Lamanna, F., 2009. The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries. Feminist Economics 15, 91–132.

Klein Goldewijk, K., Beusen, A., 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, 565–573.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., Vishny, R., 1999. The Quality of Government. Journal of Law, Economics, and Organization 15, 222–279.

Lagerlöf, N.P., 2003. Gender Equality and Long-Run Growth. Journal of Economic Growth 8, 403–426.

Marshall, A., 2013. Principles of Economics. 8*th* ed., Palgrave Macmillian.

Mitterauer, M., 2010. Why Europe?: The Medieval Origins of Its Special Path. University of Chicago Press, Chicago.

Murray, D.R., 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, 99–108.

Nunn, N., Puga, D., 2012. Ruggedness: The Blessing of Bad Geography in Africa. Review of Economics and Statistics 94, 20–36.

Olsson, O., Paik, C., 2016. Long-run Cultural Divergence: Evidence from the Neolithic Revolution. Journal of Development Economics 122, 197–213.

Oster, E., 2019. Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics 37, 187–204.

Parker, P.M., 2000. Physioeconomics: The Basis for Long-run Economic Growth. MIT Press, Cambridge, Massachusetts.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated World Map of the Köppen-Geiger Climate Classification. Hydrology and Earth System Sciences Discussions 4, 439–473.

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., Strulik, H., 2017. Gender Equity and the Escape from Poverty. Oxford Economic Papers 69, 55–74.

Putterman, L., Trainor, C.A., 2006. Agricultural Transition Year Country Data Set. Online resource.

Putterman, L., Weil, D.N., 2010. Post-1500 Population Flows and the Long-run Determinants of Economic Growth and Inequality. Quarterly Journal of Economics 125, 1627–1682.

Reher, D.S., 1998. Family Ties in Western Europe: Persistent Contrasts. Population and Development Review 24, 203–234.

Rijpma, A., Carmichael, S.G., 2016. Testing Todd and Matching Murdock: Global Data on Historical Family Characteristics. Economic History of Developing Regions 31, 10–46.

Ross, M.L., 2008. Oil, Islam, and Women. American Political Science Review 102, 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, 249–273.

Smith, R.M., 1981. Fertility, Economy, and Household Formation in England over Three Centuries. Population and Development Review 7, 595–622.

Spolaore, E., Wacziarg, R., 2009. The Diffusion of Development. Quarterly Journal of Economics 124, 469–529.

Spolaore, E., Wacziarg, R., 2013. How Deep Are the Roots of Economic Development? Journal of Economic Literature 51, 325–369.

Szołtysek, M., 2014. Toward a Conceptual Framework for the Variation in Historical Family and Household Systems Across Eurasia. Przeszlosc Demograficzna Polski 36, 55–86.

Szołtysek, M., Klüsener, S., Poniat, R., Gruber, S., 2017. The Patriarchy Index: A New Measure of Gender and Generational Inequalities in the Past. Cross-Cultural Research 51, 228–262.

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., Buhaug, H., 2012. PRIO-GRID: A Unified Spatial Data Structure. Journal of Peace Research 49, 363–374.

Tur-Prats, A., 2019. Family Types and Intimate Partner Violence: a Historical Perspective. Review of Economics and Statistics 101, 878–891.

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., Voth, H.J., 2013. How the West "Invented" Fertility Restriction. American Economic Review 103, 2227–2264.

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, 2016. World Development Indicators. [accessed on 24-09-2016].

Figures and Tables

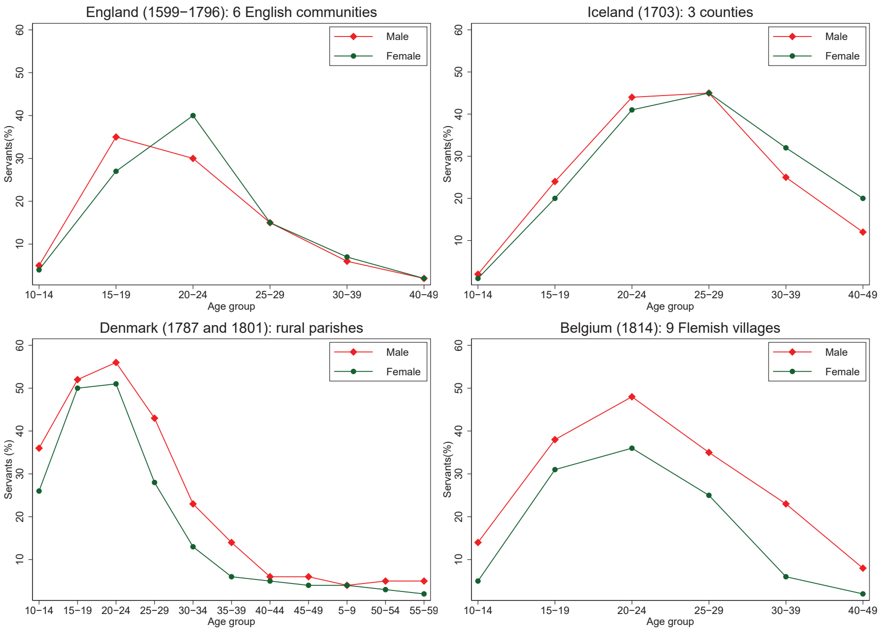


Figure 1: Service prevalence by age and sex for selected Northwest European countries. *Source:* Hajnal (1982): Tables 13, 14, and 17.

Table 1: Descriptive statistics for selected variables

	Mean	(Std. Dev.)	Min.	Max.	N
Average female-male ratio in 1990–2010:					
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
Ages at first marriage:					
Contemporary					
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
Historical (Europe only, ref. = England)					
Female	-2.07	(3.1)	-6.81	2.36	28
Deep determinants:					
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.

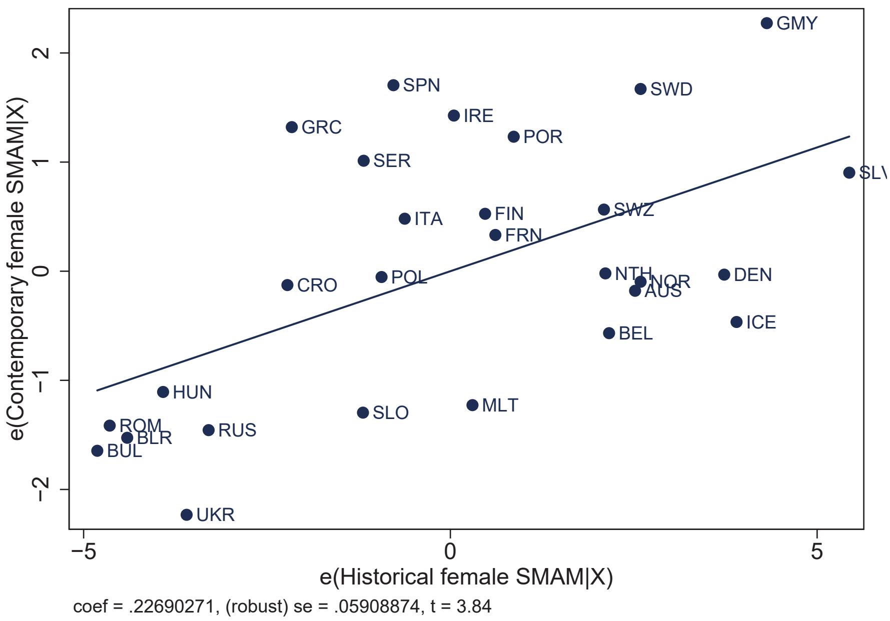
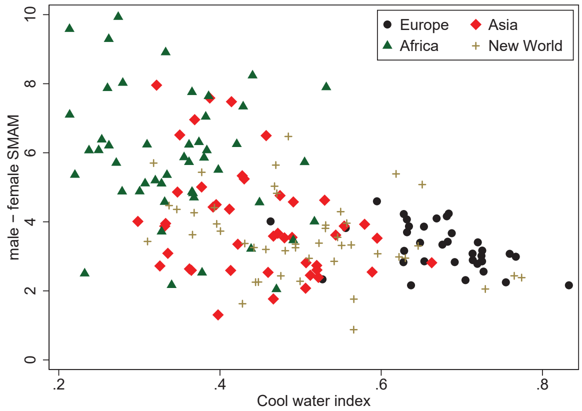
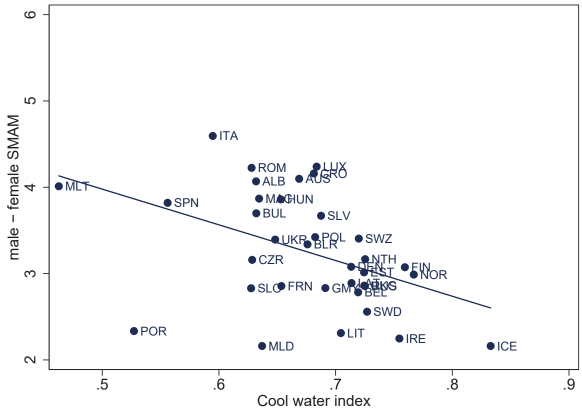


Figure 2: Historical and contemporary female SMAM for 28 European countries

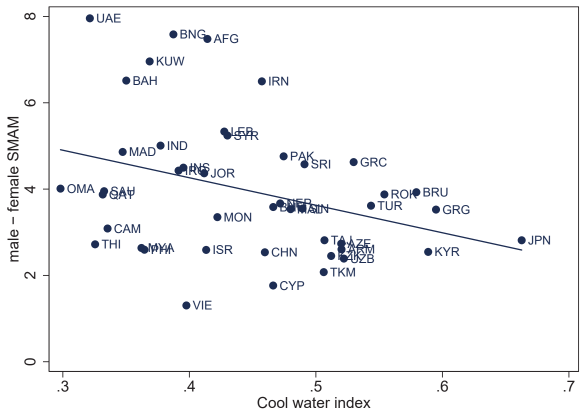
Notes: Linear correlation controlling for the year of observation of contemporary female SMAM, 1966–2002. Historical data from Dennison and Ogilvie (2014, Table 2); contemporary data from UN (2009). SMAM is singulate mean age at first marriage.



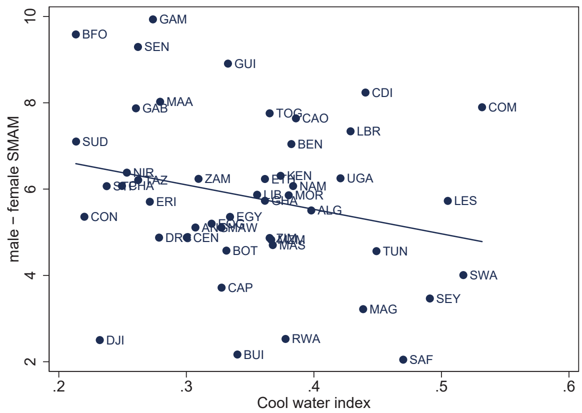
(a) World



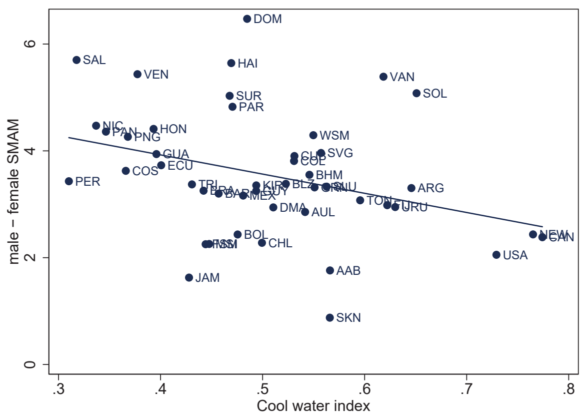
(b) Europe



(c) Asia



(d) Africa



(e) New World

Figure 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.

Table 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)
Ages at first marriage:									
Female	0.037***			0.006***			0.017*		
	(0.010)			(0.002)			(0.009)		
Male	-0.032***			-0.006***			-0.005		
	(0.012)			(0.002)			(0.010)		
Female/male		1.014***			0.164***			0.425*	
		(0.269)			(0.043)			(0.245)	
Male−female			-0.035***			-0.006***			-0.014
			(0.010)			(0.002)			(0.009)
Year of obs.	-0.002	-0.002	-0.002	0.001	0.000	0.001	-0.003	-0.003	-0.002
	(0.003)	(0.003)	(0.003)	(0.000)	(0.000)	(0.000)	(0.003)	(0.003)	(0.003)
Deep determinants:									
Years of agriculture	-0.041***	-0.040***	-0.041***	-0.004**	-0.004**	-0.004**	-0.024**	-0.025**	-0.026**
	(0.010)	(0.010)	(0.010)	(0.002)	(0.002)	(0.002)	(0.010)	(0.010)	(0.011)
Plow	-0.096*	-0.095*	-0.099*	-0.004	-0.002	-0.003	0.023	0.017	0.015
	(0.052)	(0.052)	(0.052)	(0.009)	(0.009)	(0.009)	(0.098)	(0.096)	(0.096)
Agricultural suitability	0.115**	0.118**	0.118**	-0.001	-0.001	-0.001	-0.046	-0.044	-0.045
	(0.052)	(0.053)	(0.052)	(0.010)	(0.010)	(0.010)	(0.071)	(0.069)	(0.069)
Historical controls:									
Large animals	0.018	0.015	0.023	-0.020	-0.022	-0.020	-0.190*	-0.187*	-0.186*
	(0.081)	(0.083)	(0.081)	(0.020)	(0.020)	(0.020)	(0.104)	(0.106)	(0.106)
Political hierarchies	-0.005	-0.006	-0.005	0.001	0.001	0.001	0.006	0.007	0.008
	(0.020)	(0.020)	(0.020)	(0.003)	(0.003)	(0.003)	(0.016)	(0.016)	(0.016)
Economic complexity	0.012	0.012	0.013	-0.002	-0.002	-0.002	-0.011	-0.010	-0.010
	(0.013)	(0.013)	(0.013)	(0.002)	(0.002)	(0.002)	(0.014)	(0.014)	(0.014)
Tropical climate	-0.145***	-0.140***	-0.139***	-0.024***	-0.025***	-0.025***	-0.048	-0.036	-0.036
	(0.050)	(0.047)	(0.048)	(0.009)	(0.008)	(0.008)	(0.048)	(0.048)	(0.048)
Contemporary controls:									
Income per capita (log)	-0.547***	-0.557***	-0.527***	0.137***	0.130***	0.135***	0.588***	0.618***	0.629***
	(0.159)	(0.158)	(0.160)	(0.033)	(0.030)	(0.031)	(0.192)	(0.183)	(0.183)
(Income per capita (log))^2	0.028***	0.029***	0.027***	-0.008***	-0.008***	-0.008***	-0.030***	-0.031***	-0.031***
	(0.009)	(0.009)	(0.009)	(0.002)	(0.002)	(0.002)	(0.011)	(0.011)	(0.011)
Continent dummies:	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	131	131	131	132	132	132	119	119	119
R^2	0.581	0.585	0.580	0.571	0.574	0.571	0.630	0.627	0.625
adj. R^2	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 singulate 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 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 the same time period as the dependent variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: 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)
Deep determinants:								
Cool water	6.727***	4.316*	1.949	-1.555	0.195***	0.221***	-4.672***	-5.902***
	(2.546)	(2.494)	(2.423)	(2.402)	(0.047)	(0.075)	(1.302)	(2.078)
Years of agriculture		-0.312**		-0.159		-0.007*		0.154*
		(0.147)		(0.112)		(0.004)		(0.092)
Plow		-1.105		-0.253		-0.034		0.890
		(1.025)		(0.860)		(0.026)		(0.686)
Agricultural		0.310		0.728		-0.019		0.523
suitability		(0.906)		(0.824)		(0.024)		(0.646)
Historical controls:								
Large animals	1.074	0.731	0.694	0.336	0.020	0.019	-0.361	-0.395
	(1.121)	(1.363)	(1.035)	(1.096)	(0.023)	(0.035)	(0.575)	(0.877)
Political hierarchies	0.658***	0.753**	0.329	0.254	0.013**	0.020**	-0.303*	-0.482**
	(0.245)	(0.295)	(0.203)	(0.259)	(0.006)	(0.008)	(0.163)	(0.218)
Economic complexity	-0.242	-0.061	-0.030	0.064	-0.007	-0.003	0.190	0.099
	(0.156)	(0.121)	(0.106)	(0.110)	(0.005)	(0.004)	(0.128)	(0.112)
Contemporary controls:								
Income p.c. (log)	5.726**	6.092***	6.471***	6.632***	-0.002	0.010	0.918	0.759
	(2.369)	(2.113)	(1.951)	(1.807)	(0.062)	(0.061)	(1.625)	(1.631)
(Income p.c. (log))^2	-0.307**	-0.321***	-0.350***	-0.349***	0.000	-0.001	-0.052	-0.039
	(0.137)	(0.119)	(0.115)	(0.102)	(0.003)	(0.003)	(0.092)	(0.092)
Year of SMAM obs.	0.080***	0.059***	0.061***	0.050**	0.001**	0.001	-0.025	-0.015
	(0.022)	(0.022)	(0.020)	(0.021)	(0.001)	(0.001)	(0.016)	(0.016)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	138	125	134	121	134	121	134	121
R^2	0.459	0.433	0.312	0.253	0.495	0.517	0.468	0.498
adj. R^2	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. "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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: 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***	-6.054**	-6.041***	-5.921***	-5.291**	-5.335**	-6.450***	-5.781***	-4.016**	-5.464**	-4.870**
	(2.078)	(2.315)	(2.087)	(2.051)	(2.395)	(2.404)	(2.372)	(2.056)	(1.787)	(2.173)	(2.353)
Years of agriculture	0.154*	0.155*	0.068	0.133	0.095	0.090	0.141	0.116	-0.071	0.149*	-0.130
	(0.092)	(0.093)	(0.109)	(0.094)	(0.099)	(0.105)	(0.100)	(0.097)	(0.098)	(0.090)	(0.100)
Plow	0.890	0.892	0.687	1.032	0.867	0.881	0.999	1.064	0.431	0.877	1.297*
	(0.686)	(0.689)	(0.734)	(0.715)	(0.766)	(0.786)	(0.743)	(0.671)	(0.669)	(0.679)	(0.754)
Agricultural suitability	0.523	0.523	0.556	0.362	0.606	0.621	0.551	0.699	0.685	0.574	0.754
	(0.646)	(0.649)	(0.643)	(0.667)	(0.716)	(0.723)	(0.656)	(0.628)	(0.572)	(0.650)	(0.620)
Tropical climate		-0.091									-0.882
		(0.443)									(0.603)
F_{ST} from U.K. (weighted)			-5.727								1.599
			(3.534)								(4.385)
Population density in 1500				0.018							0.023*
				(0.014)							(0.012)
Total years of schooling in 1950 (log)					-0.327						-0.036
					(0.250)						(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				0.045*
							(0.024)				(0.026)
Rule of law in 2000								-0.360			-0.598*
								(0.239)			(0.303)
Catholic shares in 1980									-0.009*		-0.006
									(0.005)		(0.007)
Protestant shares in 1980									-0.017**		-0.013
									(0.007)		(0.008)
Muslim shares in 1980									0.022***		0.024***
									(0.006)		(0.008)
Oil production (per capita)										1.057	1.687
										(1.038)	(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^2	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_{ST} 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Europe: historical female ages at first marriage

	Historical female age at first marriage, 1500–1900							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Deep determinants:								
Cool water	14.424**	20.872***	13.733	18.809	12.814**	19.033**	12.873**	18.894***
	(5.175)	(6.894)	(10.357)	(11.132)	(5.708)	(7.291)	(5.443)	(6.492)
Years of agriculture			0.105	-0.005				
			(0.703)	(0.643)				
Plow					-1.668	-2.028		
					(1.168)	(1.206)		
Agricultural suitability							-2.988*	-4.361**
							(1.728)	(1.613)
preindustrial development:								
Population density in 1500		0.091***		0.093***		0.093***		0.098***
		(0.032)		(0.032)		(0.033)		(0.031)
N	27	26	26	25	27	26	27	26
R^2	0.152	0.299	0.105	0.267	0.161	0.313	0.187	0.374
adj. R^2	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. 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. "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). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Marriage probability of immigrants in the US

	Female immigrants: 1880–1930					
	(1)	(2)	(3)	(4)	(5)	(6)
Age group:	15–39	15–19	20–24	25–29	30–34	35–39
Cool water	-1.928**	-1.096***	-3.135***	-2.298**	-1.232**	-0.825
	(0.771)	(0.211)	(1.073)	(0.948)	(0.597)	(0.548)
Cool water × Years since immigration	0.034***	0.026***	0.040***	0.032**	0.016***	0.004
	(0.010)	(0.006)	(0.015)	(0.013)	(0.004)	(0.004)
Years since immigration	-0.022***	-0.019***	-0.028***	-0.021**	-0.010***	-0.002
	(0.006)	(0.004)	(0.010)	(0.008)	(0.003)	(0.003)
Other deep determinants	Yes	Yes	Yes	Yes	Yes	Yes
Ancestry-country controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Census-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	295,864	29,598	53,128	65,705	70,616	76,817
R^2	0.370	0.129	0.151	0.098	0.059	0.047
adj. R^2	0.370	0.127	0.150	0.097	0.058	0.046
Sample share of ever-married	0.723	0.117	0.502	0.769	0.885	0.921

Notes: OLS estimates are reported with standard errors clustered at the ancestry-country level in parentheses. The unit of observation is a female immigrant in the US. The data are from US censuses samples from IPUMS-US for 1880, 1900, 1910, 1920, and 1930. The dependent variable is a dummy taking value one if the woman has ever married and zero otherwise. "Years since immigration" is the number of years since the woman migrated to the US. All deep determinants are measured at the country of ancestry. "Cool water" is the cool water index described in section 3. *Other deep determinants* are: the number of years (in thousands) since the Neolithic revolution (from 1500) from Putterman and Trainor (2006), the proportion of population with ancestors that used the plow in preindustrial agriculture from Alesina et al. (2013), and agricultural suitability from Alesina et al. (2013). *Ancestry-country controls* are: ancestral domestication of large animals, ancestral settlement patterns, ancestral political complexity, fraction of ancestral land that was tropical or subtropical, the natural log of per capita income and its square, and continent dummies. *Individual controls* are: age, age squared, and literacy fixed effects. US state and census-year fixed effects included. Constants are not reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Marriage probability of children of immigrants in the US, defined by the father's country of origin

	Female children of immigrants: 1880–1930					
	(1)	(2)	(3)	(4)	(5)	(6)
Age group:	15–39	15–19	20–24	25–29	30–34	35–39
Deep determinants:						
Cool water	-1.304***	-0.282**	-1.695***	-1.933***	-1.346***	-1.696***
	(0.408)	(0.116)	(0.619)	(0.554)	(0.382)	(0.288)
Years of agriculture	-0.009	-0.005*	-0.010	-0.012	-0.004	-0.010*
	(0.007)	(0.002)	(0.011)	(0.009)	(0.006)	(0.005)
Plow	0.616*	0.179	0.251	1.245*	1.231**	1.627***
	(0.334)	(0.111)	(0.533)	(0.713)	(0.563)	(0.186)
Agricultural suitability	-0.481**	-0.106	-0.471	-0.782***	-0.651***	-0.871***
	(0.199)	(0.067)	(0.305)	(0.277)	(0.195)	(0.111)
Ancestry-country controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Census-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mother BPL fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	662,914	185,709	162,830	125,311	100,343	88,721
R^2	0.402	0.073	0.092	0.060	0.046	0.041
adj. R^2	0.402	0.073	0.091	0.059	0.044	0.040
Sample share of ever-married	0.472	0.057	0.385	0.680	0.798	0.839

Notes: OLS estimates are reported with standard errors clustered at the ancestry-country level in parentheses. The unit of observation is a US-born female child of a immigrant, as defined by the father's country of origin. The data are from US censuses samples from IPUMS-US for 1880, 1900, 1910, 1920, and 1930. The dependent variable is a dummy taking value one if the woman has ever married and zero otherwise. *Deep determinants* are all measured at the father's country of origin. "Cool water" is the cool water index described in section 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). *Ancestry-country controls* are: ancestral domestication of large animals, ancestral settlement patterns, ancestral political complexity, fraction of ancestral land that was tropical or subtropical, the natural log of per capita income and its square, and continent dummies. *Individual controls* are: age, age squared, and literacy fixed effects. Mother's birthplace (*Mother BPL*), US state, and census-year fixed effects included. Constants are not reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: 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							
Deep determinants:								
Cool water	0.626***	0.447***	0.715***	0.480**	0.612***	0.310*	0.578***	0.492***
	(0.178)	(0.165)	(0.195)	(0.185)	(0.175)	(0.171)	(0.199)	(0.182)
Years of agriculture		-0.055***			-0.048***	-0.054***		-0.048***
		(0.011)			(0.011)	(0.011)		(0.011)
Plow			-0.190***		-0.188***		-0.181***	-0.170***
			(0.062)		(0.056)		(0.055)	(0.054)
Agricultural suitability				0.149**		0.136**	0.136**	0.104*
				(0.067)		(0.062)	(0.059)	(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
N	156	146	156	156	146	146	156	146
R^2	0.282	0.452	0.332	0.309	0.493	0.471	0.354	0.504
	Panel B: Life expectancy							
Deep determinants:								
Cool water	0.073***	0.060**	0.078***	0.061**	0.073***	0.052*	0.066**	0.068**
	(0.025)	(0.025)	(0.026)	(0.027)	(0.025)	(0.028)	(0.029)	(0.028)
Years of agriculture		-0.005***			-0.005**	-0.005***		-0.005**
		(0.002)			(0.002)	(0.002)		(0.002)
Plow			-0.011		-0.015		-0.011	-0.015
			(0.011)		(0.011)		(0.010)	(0.011)
Agricultural suitability				0.012		0.007	0.011	0.004
				(0.010)		(0.012)	(0.010)	(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
N	158	147	158	158	147	147	158	147
R^2	0.408	0.466	0.413	0.413	0.474	0.468	0.418	0.475
	Panel C: Years of education							
Deep determinants:								
Cool water	0.245	0.114	0.255	0.254	0.087	0.119	0.265	0.085
	(0.205)	(0.200)	(0.190)	(0.198)	(0.189)	(0.195)	(0.179)	(0.175)
Years of agriculture		-0.029***			-0.030***	-0.029***		-0.030***
		(0.008)			(0.009)	(0.008)		(0.009)
Plow			-0.017		0.029		-0.017	0.029
			(0.073)		(0.085)		(0.074)	(0.087)
Agricultural suitability				-0.009		-0.004	-0.010	0.001
				(0.057)		(0.057)	(0.056)	(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
N	130	127	130	130	127	127	130	127
R^2	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. “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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix

The Roots of Female Emancipation

A.1 Data

For a list of the variables used in this article, some descriptive statistics, a short description, and original sources see Table A1.

Here, we extend the discussion on the CW-index from section 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 6 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.

A.2 Selection on unobservables

Here, we estimate how large the ratio (δ) 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, 2019). Table A9 presents the estimates for δ, using the method described in Oster

(2019). In our restricted regression, the male-female difference in SMAM is regressed on the CW-index and continent dummies. The controlled regressions are column 11 of Table 4, with and without the religion share variables—since most literature argues that religion is a transmission channel rather than a determinant—, and column 10 of Table A8.

We estimate δ for different values of R_{max}. R_{max} 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 38–45 percent that of selection on the included observables for the CW effect to be zero. However, Oster (2019) 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 that the R^2 of the regression with the full set of observables.[50] When implementing this 1.3 factor, the minimum δ is 0.636, when religious variables are controlled for; without religion, δ is above 0.80 (column 4). Column 5 shows that, if the factor is reduced to 1.25, δ is 0.75 for the regression with religion and around 1.00 for the regressions without religion (column 5).

Altonji et al. (2005) propose $\delta = 1$ as an *ad hoc* cutoff for robustness. As evident from Table A9, whether this cutoff is met depends on the choice of R_{max} and on whether religion is included in the full model. Nevertheless, given that, in our most conservative regressions, the full set of controls includes 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.

A.3 Additional tables and figures

[50] Oster (2019) calculates this multiplicative factor of 1.3 as the cutoff for which δ is less than unity in 90 percent of the results in a random sample of top-published randomized control trials in economics.

Table A1: Description of variables used and their source

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
Average female-male ratio in 1990–2010:							
Labor force participation	0.68	(0.21)	0.15	1.01	191	Female / male: % of ages 25-59 in labor force.	ILO Laborsta EAPEP 6^{th} 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)
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

Continued on next page

Table A1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
Cool water	0.48	(0.15)	0.21	0.83	183	Cool water index; see section 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)
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)

Continued on next page

Table A1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
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)
Years of schooling in 1950 (log):							
Total	0.5	(1.2)	-4.44	2.19	146	Log of mean years of schooling, ages 25+.	Barro and Lee (2013)
Male	0.73	(1.09)	-4.4	2.21	146	Log of male mean years of schooling, ages 25+.	Barro and Lee (2013)

Continued on next page

Table A1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
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)
Religious shares in 1980:							
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)

Continued on next page

Table A1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
GDI in 2014	0.93	(0.07)	0.6	1.03	161	Gender Development Index in 2014	UNDP (2015)
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: *Correlates of War Database* 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: *Correlates of War Database* version 4.	Alesina et al. (2013)

Continued on next page

Table A1 – *Continued from previous page*

Variable	Mean	(Std. Dev.)	Min.	Max.	N	Short description	Source
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)
Share of GDP in 2000:							
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)

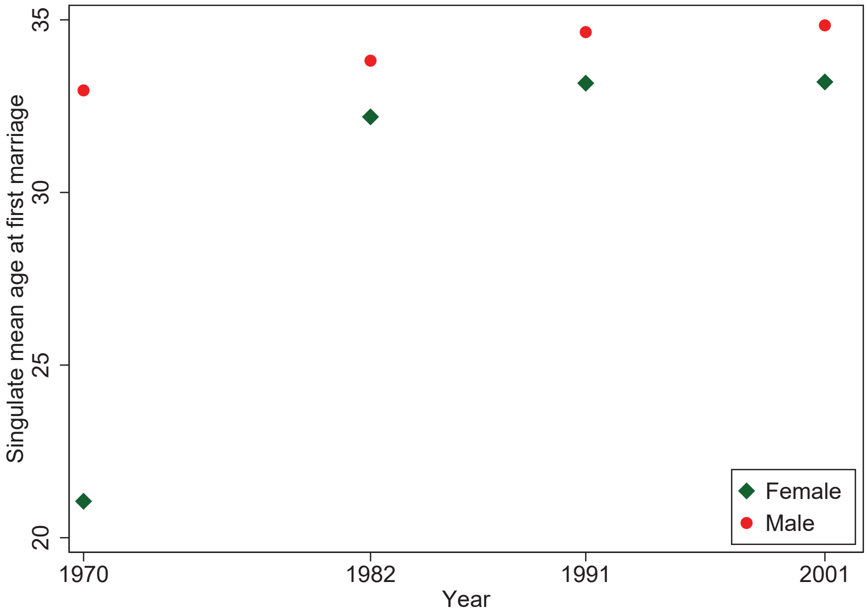


Figure 4: Jamaica: ages at first marriage. 1970 is an outlier

Sources: UN (2009).

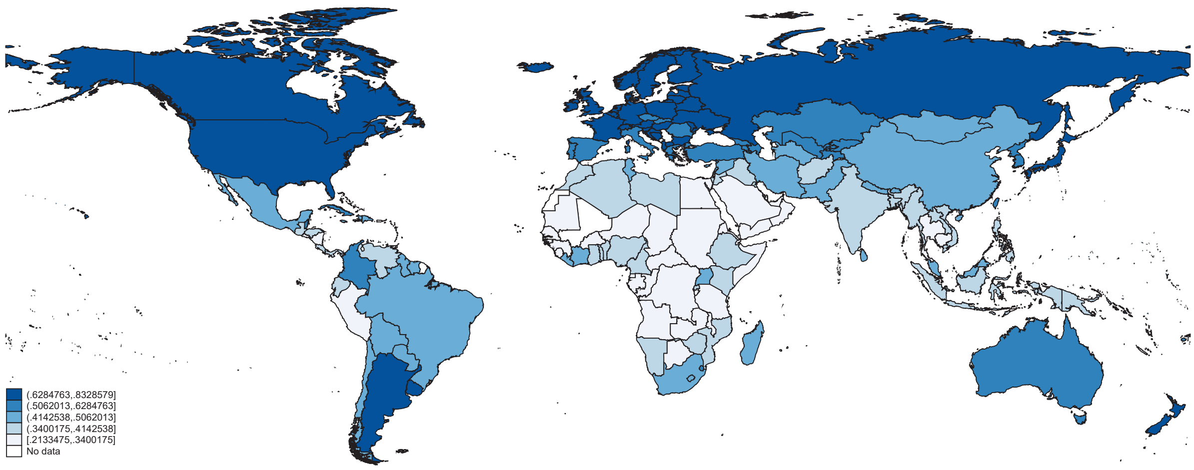
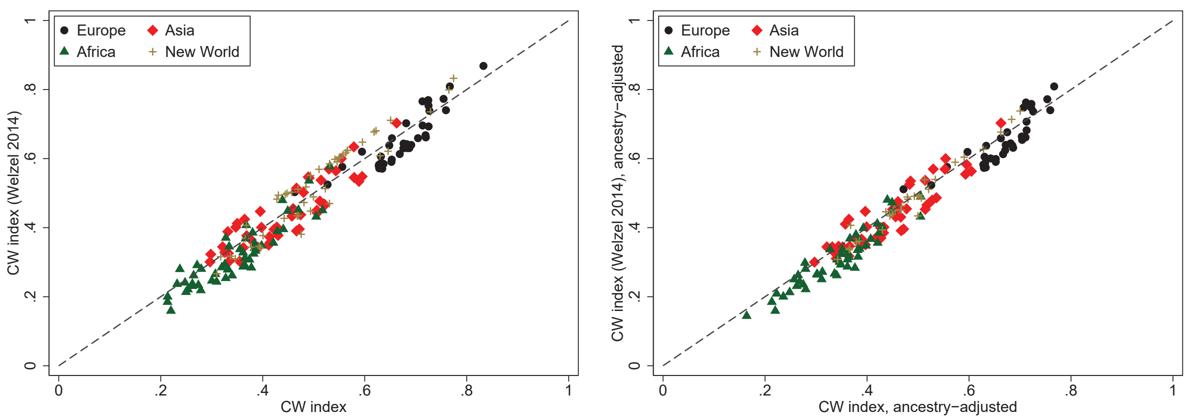


Figure 5: World distribution of the CW-condition

Table A2: 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)
SMAM:									
Female	2.765***			1.034*			0.009***		
	(1.024)			(0.524)			(0.003)		
Male	-1.510			-1.284**			-0.006*		
	(1.182)			(0.564)			(0.004)		
Female/male		74.609***			30.177**			0.263***	
		(26.338)			(13.835)			(0.081)	
Male−female			-2.334**			-1.111**			-0.009***
			(0.982)			(0.493)			(0.003)
Year of obs.	-0.283	-0.206	-0.161	-0.045	-0.083	-0.068	-0.001	-0.001	-0.001
	(0.201)	(0.179)	(0.179)	(0.124)	(0.111)	(0.114)	(0.001)	(0.001)	(0.001)
Deep determinants:									
Years of agriculture	-1.011	-0.996	-1.078	-0.801	-0.737	-0.758	-0.012***	-0.012***	-0.012***
	(1.088)	(1.074)	(1.088)	(0.556)	(0.553)	(0.549)	(0.002)	(0.002)	(0.003)
Plow	-11.415**	-11.767**	-12.579**	-4.124*	-3.785	-3.939*	-0.009	-0.010	-0.012
	(5.262)	(5.179)	(5.142)	(2.281)	(2.300)	(2.291)	(0.013)	(0.013)	(0.013)
Agricultural suitability	-3.289	-2.373	-2.350	2.217	2.110	2.115	0.022	0.024*	0.024*
	(6.086)	(6.060)	(6.220)	(2.853)	(2.871)	(2.866)	(0.015)	(0.014)	(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
N	106	106	106	125	125	125	136	136	136
R^2	0.290	0.291	0.275	0.382	0.381	0.381	0.688	0.691	0.684
adj. R^2	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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



(a) Unadjusted

(b) Ancestry-adjusted

Figure 6: CW index comparison: with and without coastal borders

Notes: Comparison of CW index, as described in section 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 A3: 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***			0.006*			0.002		
	(0.016)			(0.003)			(0.013)		
Male	-0.064***			-0.008**			0.011		
	(0.018)			(0.003)			(0.015)		
Female/male		1.751***			0.159**			0.138	
		(0.413)			(0.076)			(0.334)	
Male−female			-0.068***			-0.006**			-0.000
			(0.016)			(0.003)			(0.013)
Deep determinants:									
Years of agriculture	-0.037***	-0.038***	-0.038***	-0.006***	-0.005***	-0.005***	-0.014*	-0.015*	-0.015*
	(0.010)	(0.011)	(0.010)	(0.002)	(0.002)	(0.002)	(0.008)	(0.008)	(0.008)
Plow	0.068	0.065	0.061	0.000	0.005	0.004	0.016	0.003	-0.001
	(0.050)	(0.048)	(0.049)	(0.013)	(0.013)	(0.012)	(0.064)	(0.062)	(0.062)
Agricultural suitability	0.042	0.041	0.050	0.013	0.008	0.009	-0.055	-0.047	-0.045
	(0.059)	(0.056)	(0.056)	(0.014)	(0.013)	(0.013)	(0.057)	(0.057)	(0.057)
N	97	97	97	98	98	98	90	90	90
R^2	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***			0.007***			0.024**		
	(0.012)			(0.002)			(0.011)		
Male	-0.035**			-0.006**			-0.002		
	(0.014)			(0.003)			(0.013)		
Female/male		1.131***			0.195***			0.554**	
		(0.300)			(0.046)			(0.268)	
Male−female			-0.039***			-0.007***			-0.017*
			(0.011)			(0.002)			(0.010)
Deep determinants:									
Years of agriculture	-0.035***	-0.035***	-0.036***	-0.004**	-0.004**	-0.004**	-0.019	-0.020	-0.021
	(0.011)	(0.011)	(0.011)	(0.002)	(0.002)	(0.002)	(0.012)	(0.013)	(0.013)
Plow	-0.122*	-0.117*	-0.121*	-0.013	-0.012	-0.013	0.013	0.019	0.018
	(0.070)	(0.069)	(0.069)	(0.011)	(0.011)	(0.011)	(0.155)	(0.157)	(0.158)
Agricultural suitability	0.097	0.102	0.096	-0.009	-0.008	-0.009	-0.044	-0.055	-0.059
	(0.066)	(0.066)	(0.065)	(0.012)	(0.011)	(0.012)	(0.106)	(0.103)	(0.104)
N	105	105	105	105	105	105	94	94	94
R^2	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***			0.007***			0.014		
	(0.010)			(0.002)			(0.010)		
Male	-0.028**			-0.005**			-0.002		
	(0.012)			(0.002)			(0.011)		
Female/male		0.904***			0.188***			0.349	
		(0.277)			(0.043)			(0.264)	
Male−female			-0.031***			-0.006***			-0.011
			(0.010)			(0.002)			(0.009)
Deep determinants:									
Years of agriculture	-0.045***	-0.045***	-0.046***	-0.003	-0.003	-0.003	-0.026**	-0.028**	-0.029**
	(0.012)	(0.012)	(0.012)	(0.002)	(0.002)	(0.002)	(0.012)	(0.012)	(0.013)
Plow	-0.107*	-0.106*	-0.110**	-0.003	-0.003	-0.004	0.011	0.003	0.001
	(0.055)	(0.054)	(0.054)	(0.009)	(0.009)	(0.009)	(0.102)	(0.099)	(0.099)
Agricultural suitability	0.133**	0.136**	0.136**	-0.004	-0.003	-0.003	-0.043	-0.041	-0.042
	(0.053)	(0.053)	(0.052)	(0.010)	(0.009)	(0.010)	(0.074)	(0.072)	(0.071)
N	101	101	101	102	102	102	90	90	90
R^2	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 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: 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***			-3.156***					
	(0.773)			(0.889)					
Mild summers		-2.542***		-2.868***					
		(0.825)		(0.815)					
Continuous rain			-0.256	-0.861					
			(0.734)	(0.790)					
Coolness-component					-0.362		-2.267**	-2.551	
					(0.930)		(1.067)	(2.024)	
Water-component						-2.376**	-3.272***	-3.472*	
						(0.971)	(1.154)	(1.821)	
Coolness× *Water*								0.535	
								(3.150)	
Cool water									-5.902***
									(2.078)
Years of agriculture	0.269***	0.216**	0.228**	0.212**	0.241***	0.167*	0.143	0.144	0.154*
	(0.082)	(0.085)	(0.092)	(0.092)	(0.084)	(0.091)	(0.094)	(0.094)	(0.092)
Plow	0.612	-0.081	0.064	0.834	0.127	0.093	0.717	0.716	0.890
	(0.658)	(0.594)	(0.645)	(0.643)	(0.691)	(0.601)	(0.689)	(0.692)	(0.686)
Agricultural suitability	0.512	-0.166	-0.082	0.664	0.002	-0.147	0.355	0.349	0.523
	(0.604)	(0.536)	(0.573)	(0.597)	(0.610)	(0.541)	(0.634)	(0.639)	(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^2	0.472	0.497	0.450	0.537	0.450	0.485	0.502	0.502	0.498
adj. R^2	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 for more details. "Cool water" is the cool water index described in section 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 A5: 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							
Deep determinants:								
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)
N	96	85	95	84	95	84	95	84
R^2	0.418	0.406	0.370	0.338	0.344	0.374	0.308	0.355
adj. R^2	0.333	0.277	0.278	0.192	0.248	0.236	0.206	0.213
	Panel B: Old World							
Deep determinants:								
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)
N	104	98	100	94	100	94	100	94
R^2	0.376	0.458	0.256	0.344	0.438	0.491	0.427	0.478
adj. R^2	0.316	0.382	0.181	0.247	0.382	0.416	0.370	0.401
	Panel C: without Europe							
Deep determinants:								
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)
N	115	103	111	99	111	99	111	99
R^2	0.461	0.432	0.315	0.264	0.466	0.488	0.436	0.467
adj. R^2	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. *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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: 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)
Deep determinants:								
Cool water,	5.706**	6.133**	-2.081	-2.570	0.289***	0.331***	-7.837***	-8.938***
ancestry-adjusted	(2.729)	(2.964)	(2.519)	(2.744)	(0.058)	(0.074)	(1.634)	(2.049)
Years of agriculture,		-0.142		-0.071		-0.003		0.066
ancestry-adjusted		(0.152)		(0.117)		(0.003)		(0.088)
Plow		-0.635		0.214		-0.034		0.917
		(0.911)		(0.704)		(0.025)		(0.653)
Agricultural		-0.221		0.166		-0.022		0.522
suitability		(0.805)		(0.711)		(0.021)		(0.560)
Historical controls:								
Large animals	0.278	0.260	-0.404	-0.367	0.024	0.020	-0.666	-0.575
	(0.811)	(0.862)	(0.769)	(0.812)	(0.023)	(0.024)	(0.628)	(0.650)
Political	0.442*	0.575*	0.155	0.110	0.010	0.018**	-0.248	-0.450*
hierarchies	(0.248)	(0.308)	(0.190)	(0.256)	(0.007)	(0.009)	(0.178)	(0.228)
Economic	0.026	0.019	0.169	0.163	-0.003	-0.003	0.115	0.113
complexity	(0.111)	(0.112)	(0.105)	(0.105)	(0.004)	(0.004)	(0.111)	(0.103)
Contemporary controls:								
Income p.c. (log)	4.628**	5.023**	5.399***	5.567***	-0.010	-0.001	0.968	0.766
	(2.306)	(2.412)	(1.990)	(2.108)	(0.059)	(0.059)	(1.602)	(1.600)
(Income p.c. (log))^2	-0.244*	-0.264*	-0.284**	-0.292**	0.001	0.000	-0.051	-0.040
	(0.132)	(0.137)	(0.114)	(0.120)	(0.003)	(0.003)	(0.092)	(0.092)
Year of SMAM obs.	0.066***	0.067***	0.048**	0.048**	0.001**	0.001**	-0.027*	-0.029*
	(0.022)	(0.022)	(0.021)	(0.021)	(0.001)	(0.001)	(0.016)	(0.016)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	125	125	121	121	121	121	121	121
R^2	0.375	0.389	0.210	0.214	0.524	0.545	0.514	0.534
adj. R^2	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 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 A7: 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)
Deep determinants:					
Cool water	-5.902***	-5.640***	-5.672**	-5.372**	-5.534**
	(2.078)	(2.148)	(2.167)	(2.225)	(2.296)
Years of agriculture	0.154*	0.153*	0.157*	0.155*	0.161*
	(0.092)	(0.092)	(0.094)	(0.092)	(0.096)
Plow	0.890	0.852	0.847	0.902	0.916
	(0.686)	(0.698)	(0.702)	(0.721)	(0.735)
Agricultural suitability	0.523	0.506	0.516	0.513	0.532
	(0.646)	(0.652)	(0.656)	(0.656)	(0.656)
Northwest Europe		-0.421		-0.453*	
		(0.267)		(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
N	121	121	112	121	108
R^2	0.498	0.500	0.466	0.501	0.438
adj. R^2	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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Determinants of ages at first marriage: robustness to inclusion of additional controls

	Singulate mean ages at first marriage: male−female									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cool water	-5.902***	-5.830***	-5.925***	-5.974***	-5.206**	-5.301***	-6.074***	-5.912***	-6.296**	-5.669**
	(2.078)	(2.067)	(2.107)	(2.205)	(2.057)	(1.959)	(2.188)	(2.040)	(2.462)	(2.498)
Years of agriculture	0.154*	0.172*	0.146	0.162*	0.212**	0.160*	0.160*	0.148	0.152	0.231**
	(0.092)	(0.101)	(0.092)	(0.091)	(0.099)	(0.091)	(0.092)	(0.091)	(0.093)	(0.113)
Plow	0.890	1.298*	0.836	0.967	0.760	0.931	1.024	0.966	1.247*	1.650**
	(0.686)	(0.686)	(0.715)	(0.686)	(0.691)	(0.681)	(0.710)	(0.692)	(0.733)	(0.777)
Agricultural suitability	0.523	0.355	0.548	0.636	0.586	0.536	0.616	0.608	0.797	1.024
	(0.646)	(0.617)	(0.644)	(0.681)	(0.653)	(0.632)	(0.670)	(0.632)	(0.637)	(0.648)
Historical controls:										
Intensive agriculture		-1.223**								-1.129*
		(0.602)								(0.630)
Subsistence share from husbandry		1.694								1.266
		(1.689)								(2.072)
Subsistence share from hunting			-5.290							-4.287
			(5.064)							(6.632)
Absence of land inheritance rules				0.813						1.423*
				(0.550)						(0.744)
Contemporary controls:										
Years of civil conflicts (1816-2007)					-0.013					-0.009
					(0.008)					(0.009)
Years of interstate conflicts (1816-2007)					-0.016					-0.011
					(0.014)					(0.017)
Terrain ruggedness index						-0.181				-0.175*
						(0.113)				(0.100)
Fraction of European descent							-0.004			0.005
							(0.008)			(0.007)
Communist dummy								-0.759**		-0.758**
								(0.343)		(0.362)
Agriculture share of GDP in 2000									0.021	0.024*
									(0.014)	(0.014)
Manufacturing share of GDP in 2000									-0.010	0.002
									(0.020)	(0.024)
Services share of GDP in 2000									-0.002	-0.006
									(0.013)	(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^2	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 SMAM" is the gender gap in SMAM data from UN (2009). For a description of the additional control variables see Table A1. *Baseline historical & contemporary controls* are as in Table 3. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Assessing unobservable selection: estimates of δ

Controls in		β $[R^2]$		δ for $\beta = 0$ given R_{max}		
Restricted set	Full set	Restricted (1)	Full (2)	$R_{max} = 1$ (3)	$R_{max} = 1.3\tilde{R}$ (4)	$R_{max} = 1.25\tilde{R}$ (5)
Continent dummies	as in Table 4, 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 above, but without religion	-5.147 [0.423]	-6.701 [0.583]	0.380	0.808 $[R_{max} = 0.757]$	0.935 $[R_{max} = 0.728]$
Continent dummies	as in Table A8, 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. $\tilde{R}$ is the R^2 from the regression with the full set of observable controls. See Oster (2019) for more details on how to estimate δ. Calculations were done using her Stata package *psacalc*.

Table A10: Gapminder: historical female ages at first marriage

	Historical female age at first marriage, 1801–1900							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Deep determinants:								
Cool water	29.490***	28.621***	31.887***	30.966***	29.496***	28.378***	29.252***	28.391***
	(3.670)	(4.202)	(5.543)	(6.484)	(3.703)	(4.149)	(3.283)	(3.760)
Years of agriculture			0.280	0.260				
			(0.321)	(0.346)				
Plow					1.402	1.634		
					(1.061)	(1.203)		
Agricultural suitability							0.280	0.272
							(1.611)	(1.547)
preindustrial development:								
Population density in 1500		-0.019		-0.016		-0.024		-0.019
		(0.028)		(0.028)		(0.028)		(0.029)
N	27	27	26	26	27	27	27	27
R^2	0.761	0.765	0.747	0.750	0.768	0.774	0.761	0.765
adj. R^2	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. "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). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Marriage probability of immigrants in the US, without integration effects

	Female immigrants: 1880–1930					
	(1)	(2)	(3)	(4)	(5)	(6)
Age group:	15–39	15–19	20–24	25–29	30–34	35–39
Deep determinants:						
Cool water	-1.301*	-0.826***	-2.597**	-1.634*	-0.791	-0.653
	(0.678)	(0.230)	(1.014)	(0.905)	(0.548)	(0.439)
Years of agriculture	0.006	-0.007*	-0.002	0.013	0.011	0.005
	(0.009)	(0.004)	(0.015)	(0.013)	(0.007)	(0.006)
Plow	0.275	-0.053	0.658	0.563	0.088	0.123
	(0.334)	(0.170)	(0.507)	(0.431)	(0.293)	(0.260)
Agricultural suitability	-0.426	-0.110	-0.744*	-0.696*	-0.317	-0.251
	(0.265)	(0.102)	(0.398)	(0.371)	(0.216)	(0.178)
Ancestry-country controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Census-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	446,376	49,604	79,688	97,785	105,941	113,358
R^2	0.381	0.121	0.147	0.095	0.058	0.041
adj. R^2	0.381	0.120	0.146	0.094	0.057	0.041
Sample share of ever-married	0.708	0.107	0.486	0.762	0.880	0.918

Notes: OLS estimates are reported with standard errors clustered at the ancestry-country level in parentheses. The unit of observation is a female immigrant in the US. The data are from US censuses samples from IPUMS-US for 1880, 1900, 1910, 1920, and 1930. The dependent variable is a dummy taking value one if the woman has ever married and zero otherwise. *Deep determinants* are all measured at the country of ancestry. "Cool water" is the cool water index described in section 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). *Ancestry-country controls* are: ancestral domestication of large animals, ancestral settlement patterns, ancestral political complexity, fraction of ancestral land that was tropical or subtropical, the natural log of per capita income and its square, and continent dummies. *Individual controls* are: age, age squared, and literacy fixed effects. US state and census-year fixed effects included. Constants are not reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Marriage probability of children of immigrants in the US, defined by the mother's country of origin

	Female children of immigrants: 1880–1930					
	(1)	(2)	(3)	(4)	(5)	(6)
Age group:	15–39	15–19	20–24	25–29	30–34	35–39
Deep determinants:						
Cool water	-0.940***	-0.293**	-0.890*	-1.498***	-1.434***	-0.924***
	(0.301)	(0.133)	(0.476)	(0.392)	(0.269)	(0.333)
Years of agriculture	-0.003	-0.003	-0.004	-0.002	-0.004	-0.004
	(0.005)	(0.003)	(0.009)	(0.007)	(0.005)	(0.006)
Plow	0.327	-0.191	0.949*	-0.462	-0.092	0.812***
	(0.286)	(0.227)	(0.531)	(0.696)	(0.320)	(0.287)
Agricultural suitability	-0.260**	0.046	-0.341*	-0.359***	-0.476***	-0.360*
	(0.120)	(0.083)	(0.175)	(0.132)	(0.134)	(0.185)
Ancestry-country controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Census-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Father BPL fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	593,352	166,492	146,872	112,177	89,296	78,515
R^2	0.402	0.074	0.093	0.062	0.046	0.042
adj. R^2	0.402	0.073	0.091	0.061	0.044	0.040
Sample share of ever-married	0.470	0.057	0.384	0.679	0.798	0.839

Notes: OLS estimates are reported with standard errors clustered at the ancestry-country level in parentheses. The unit of observation is a US-born female child of a immigrant, as defined by the mother's country of origin. The data are from US censuses samples from IPUMS-US for 1880, 1900, 1910, 1920, and 1930. The dependent variable is a dummy taking value one if the woman has ever married and zero otherwise. *Deep determinants* are all measured at the mother's country of origin. "Cool water" is the cool water index described in section 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). *Ancestry-country controls* are: ancestral domestication of large animals, ancestral settlement patterns, ancestral political complexity, fraction of ancestral land that was tropical or subtropical, the natural log of per capita income and its square, and continent dummies. *Individual controls* are: age, age squared, and literacy fixed effects. Father's birthplace (*Father BPL*), state, and US census-year fixed effects included. Constants are not reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.