Schedule: 2014-10-30, 11:15-12:30

Plase: Building 72, room 465

Patrilocality and Missing Women*

Avraham Ebenstein
Hebrew University of Jerusalem

April 2014

Abstract: Recent scholarship has documented an alarming increase in the sex ratio at birth in parts of East Asia, South Asia and the South Caucuses. I argue that parents engage in sex selection because of patrilocal norms that dictate elderly coresidence between parents and sons. Sex ratios and coresidence rates are positively correlated when looking across countries, within countries across districts, and within districts across ethnic groups. I examine the origins of patrilocality, and find it is most common among ethnic groups which practiced intensive agriculture. I conclude with an examination of how parents respond to changes in public pension programs.

* Corresponding author: ebenstein@mscc.huji.ac.il. I am deeply indebted to Monica Das Gupta for critical insight throughout the project. I would also like to thank Hoyt Bleakely, David Card, Janet Currie, Larry Katz, Ronald Lee,

Avi Simhon, and Charles Westoff for helpful discussions. Excellent research assistance was provided by Susan Schwartz, Noam Buckman, and Martha Jachimski.

I. Introduction

Following the fall of the Berlin Wall in 1989, political upheaval ensued in many states of the Former Soviet Union, including in the South Caucasus region. The collapse of the guarantees associated with the Communist regime and armed conflict between neighboring Armenia and Azerbaijan led to dramatic fertility decline in both countries, falling below two births per woman. Coincident with this decline, the sex ratio at birth (males per 100 females) began an alarming rise throughout the South Caucuses, cresting above 110 in Armenia, Azerbaijan, and Georgia as a result of sex-selective abortion (Duthe et al. 2011, Guilmoto 2012).[1] While this trend has attracted less attention than the increasing SRB in other parts of Asia, such as China and India, it is no less disturbing and arguably more puzzling. Women in the South Caucuses enjoy relatively high status: women have higher educational attainment than men, work at similar rates, and participate in civic society.[2] However, the sex ratio is similar to what is found in China and India. Why, then, do parents in the South Caucuses abort girls?

When Amartya Sen coined the phrase "missing girls" in the New York Review of Books (1990), he attributed the phenomenon to widespread "female neglect." Others have attributed the missing women issue to cultural "son preference." However, these descriptions are inadequate for understanding the patterns in the South Caucasus, where girls are schooled at an equal or higher rate than boys. These explanations also seem insufficient for explaining the patterns in China and India, which are unremarkable in terms of their levels of gender inequality relative to many other countries, but are responsible for the vast majority of "missing girls." As I will demonstrate, globally, gender inequality is only weakly correlated with the sex ratio at birth, a

[1] The natural sex ratio at birth is roughly 106 boys per 100 girls (Coale and Banister 1994).
[2] All three countries rank in the middle of countries in terms of gender equality as measured by the World Economic Forum: http://www.weforum.org/issues/global-gender-gap. In fact, none of the countries with elevated sex ratios at birth are in the bottom quartile of their measures of gender equality.

result difficult to reconcile with an explanation that missing women are the result of widespread chauvinism. The similarity of the demographic patterns between the South Caucuses and other parts of Asia provides an opportunity to analyze distorted sex ratios across different social and economic contexts and identify a common factor shared by each of these regions that could potentially explain why parents abort girls.

In this paper, I argue that parents abort girls because of patrilocality: a cultural norm in which sons provide care for their elderly parents, and daughters leave the home following marriage to provide care for their in-laws. This custom implies that parents with a son will have two caregivers in their elder years, whereas those with daughters will have zero, resulting in a significant difference in the expected value of sons relative to daughters. While this is widely recognized by the literature as a factor in determining parental son preference, in this paper I demonstrate that it is the single common factor across countries with high sex ratios, and as such, should be viewed as the primary factor in explaining the phenomenon.[3] Patrilocality is the single feature common to the social norms of Christians in Armenia, Muslims in Azerbaijan, Hindus in India and Buddhists in China – all live with their sons when they are old.

I present several empirical results implicating the custom of patrilocality in explaining the "missing women" phenomenon. First, using international IPUMS and DHS data across 141 surveys and 108 unique countries, I find that countries where the elderly live with sons at higher rates have higher sex ratios at birth (SRBs). As shown in Figure I, *every* country with abnormally high sex ratios at birth in the samples has a high proportion of elderly living with sons. As I will

[3] Other studies have generally attributed son preference to a collection of religious and cultural norms that make sons desirable, but do not single out coresidence as singularly important. For example, Zeng et al. (1993) attribute son preference to "social and cultural traditions and daily living conditions [that] make it very important for families to have a son, especially in rural areas." To my knowledge, the only other examination of variation in coresidence patterns and sex ratios is by Guilmoto (2012) who examines Vietnam. He finds that regions with higher rates of coresidence between parents and sons have higher sex ratios at birth.

later show, this correlation is robust to the inclusion of continent fixed effects, and survives several robustness checks.[4] Importantly, I also find that gender equity is only weakly correlated with the sex ratio at birth across countries, and our results are robust to the inclusion of various measures of gender equity, including relative education and employment rates. This further implicates patrilocality in explaining the "missing women" phenomenon, rather than alternative explanations focusing on expected income.[5]

Second, I examine the relationship between the sex ratio at birth and elderly coresidence patterns within several countries. While parents are vastly more likely to live with sons than daughters in every country with a high sex ratio at birth, there is substantial variation within these countries in the strength of this pattern. For example, different ethnic groups in China and different religious groups in India exhibit different norms of coresidence; among groups with higher coresidence with sons, the sex ratio at birth is higher. Interestingly, I also find variation in coresidence rates and sex ratios in countries that have no aggregate bias in the sex ratio at birth. For example, the normal overall sex ratios in many Sub-Saharan African countries masks considerable variation across regions within these countries and across different ethnic groups. The presence of different groups living in the same modern country that exhibit different coresidence norms enables me to demonstrate a link between coresidence norms and sex ratios at

[4] Interestingly, the figure also reveals that several African countries have sex ratios below normal, the only region in the world where this is observed. While sex selection is probably rare in Africa, in countries with high infant mortality, parents make choices every day that have relevance for determining the sex ratio of children. Since boys are more fragile than girls, parents being indifferent to boys or girls will result in sex ratios below those found in developed countries with low childhood mortality (Waldron 1985). This pattern in Africa was also noted by Sen (1990) in his original article bringing attention to the missing women phenomenon.

[5] Other scholars present compelling evidence in favor of other explanations. Qian (2008) exploits variation in the suitability of land for planting crops in which women have a relative advantage. Using the agricultural reforms in China as a "natural experiment," she finds that the sex ratio at birth is sensitive to the relative wages for men and women. In this paper, I argue that relative wage considerations are of secondary importance to those related to old age care.

birth in models with country-fixed effects. The results indicate a strong relationship between sex ratios and coresidence norms, even within countries.

Third, I examine the roots of patrilocality, in an attempt to deal with potential endogeneity between son preference and coresidence norms. A reasonable concern is that ethnic groups which prefer sons also choose to live with sons, resulting in a spurious correlation between coresidence patterns and sex ratios. In order to discern whether coresidence patterns are a byproduct of son preference or a root cause, I examine the root causes of patrilocal norms directly. I present evidence supporting a hypothesis put forward by Holden, Sear, and Mace (2003), who argue that patrilocal norms are adopted as a result of cattle-based agriculture. They argue that adoption of cattle increases the value of land, which serves to increase the benefits of engaging in patrilocality. Their framework is examined in the next section, and generates a prediction that modern coresidence with sons will be more common in areas with long traditions of intensive agriculture, which serves to strengthen customs favoring sons in land inheritance.[6]

I explore the relationship between intensive agriculture and patrilocal norms empirically in three ways. First, using the Ethnographic Atlas, I demonstrate that intensive agriculture is associated with patrilocal norms across the 1,265 ethnic groups. This result is found even in models with country fixed effects, where I examine patrilocal norms across groups living in the same modern country. Second, using Murdock's (1959) Ethnographic Map of Africa in combination with modern census data, I demonstrate that areas populated by patrilocal ethnic groups have higher modern coresidence rates and sex ratios at birth, demonstrating a link

[6] This is discussed at length in the next section. First, when land is valuable, it is sensible for sons to remain near their parent's land to help defend in times of conflict. Second, when land is valuable, parents can bequeath this land to sons and enable them to engage in polygyny. Daughters, however, are more limited in their reproductive potential. So, for wealthy parents, having sons becomes advantageous insofar as they wish to maximize their number of offspring. This generates a prediction that intensive agriculture will be associated with partrilocal norms, and these norms will be most rigid when land is most valuable

between tribal customs and modern behavior. Third, I examine data on the timing of the Neolithic transition (Putterman 2008). If the paper's hypothesis is correct, and the roots of the custom of living with sons is found in intensive agriculture, then it stands to reason that areas which adopted agriculture earlier will have stronger patrilocal norms.[7] Indeed, I find a striking relationship between the timing of each country's Neolithic transition and modern coresidence norms. Interestingly, China, India, and the countries in the South Caucuses, which all exhibit the highest rates of coresidence with sons and the highest sex ratios at birth, all had Neolithic transitions roughly 9,000 years ago, making them some of the oldest civilizations in the world to engage in agriculture.[8] This is consistent with an interpretation that land endowments resulted in norms favoring sons in land inheritance, and is an important explanation for the modern phenomenon of "missing girls."

In the last portion of the analysis, I examine two different contexts in which the availability of old age support from the government changed sharply. I focus on South Korea and Armenia, two countries where it is customary for the elderly to live with sons. If, as this paper claims, high sex ratios are due to patrilocal norms where the son provides parental old-age support, changes in the availability of government old age support should affect the sex ratio at birth. South Korea's rapid expansions in national pension programs during the 1990s provide a

[7] Considerable variation exists in the timing of the adoption of agriculture due to variation in weather patterns, exposure to disease, and topography. In a recent paper, Alsan (2013) argues that the exposure of Sub-Saharan Africa to the Tse Tse fly prevented adoption of agriculture for thousands of years. Ashraf and Galor (2011) use the timing of the Neolithic transition to test the Malthusian prediction that population growth reduces the standard of living. They demonstrate that the timing of the revolution is related to biogeographical factors, such as prehistoric domesticability of species of wild plants and animals.

[8] The exact timing and diffusion of the Neolithic transition is the subject of academic debate. While agriculture is thought to have developed first in the Fertile Crescent roughly 12,000 years ago, this region would not necessarily have developed son preference like other regions, because climatic changes resulted in a drier context that led inhabitants to take their domesticated animals with them in search of better conditions. The groups that remained, like the Bedouin, were less likely to practice intensive agriculture given the reduced land quality and therefore would not have been as concerned with bequeathing their land and more concerned with bequeathing their cattle. These groups would presumably be patrilocal but have weaker customs of coresidence, since cattle can be taken to new land.

unique opportunity to test the hypothesis in a differences in differences setup, where I exploit differences in the pension program expansion for self-employed versus salaried workers (who had been required to participate prior to the expansion). Since the cohorts who experienced the pension expansion are still in their childbearing years, it is infeasible to test whether this group will have lower coresidence rates. However, the data available implicate the large scale social insurance program in South Korea led to a normalization of the sex ratio at birth. As predicted, the groups targeted in the pension expansion experienced a large decline in the sex ratio at birth.

I perform a symmetric analysis on Armenia, which has experienced a striking increase in the sex ratio at birth in the wake of the fall of the Soviet Union. One explanation is that the anticipated need to rely on children increased in the wake of a diminished expectation of state support, resulting in more parents wanting to insure the birth of a son. I present two forms of evidence consistent with this interpretation. First, within Armenia, I exploit variation in coresidence across different districts and find that districts with higher coresidence rates have a higher sex ratio at birth. Second, I present a differences in differences setup in which I compare the lower educated and higher educated in Armenia. These groups currently differ sharply in the rates of coresidence, but had similar sex ratios prior to 1991. Indeed, I find a sharp divergence in the sex ratio after the fall of communism, with the sex ratio rising for the low educated and remaining normal among those with more education. This is consistent with an interpretation that the decline in faith in the state encouraged lower income individuals to secure care in their retirement through non-governmental channels, such as having a son.

The rest of the paper is organized as follows. In Section II, I present the paper's conceptual framework. In Section III, I present the empirical results on the relationship between coresidence and sex ratios using IPUMS and DHS data. In Section IV, the paper turns to an analysis of the historical origins of patrilocality using the Ethnographic Atlas and GIS data. In Section V, I present an examination of the recent experience in South Korea and Armenia. I conclude in Section VI.

II. Conceptual Framework

The first human societies are believed to have been based on matrilineal descent. As first described by Bachofen (1861), humanity's forebears are thought to have lived in a context of promiscuity, where sexual partners were shared among community members (Lubbock 1870). The resulting ambiguity over paternity would have led to stronger bonds forming between mothers and children, and weaker ties between children and their fathers. This led to a matrifocal (mother-centered) structure with strong kinship ties among those sharing the same mother. Evidence of these patterns is found among primate social groups like chimpanzees, in which food-sharing takes places with matrifocal groups rather than between sexual mates (Engels 1902). However, at some point in history, sexual norms began to evolve, with pre-historical promiscuity giving way to family structures that took forms more similar to those found in modern history, including monogamy, polygyny and polyandry (Morgan 1877). Kinship norms also began to evolve, taking on very different rules across different groups. What determined the evolution of kinship norms, and which groups would have adopted patrilineal lines of descent?

In this paper, I focus on a line of thought first put forward first by Engels, who argued that the intensification of agriculture led to a shift from matrilineal kinship norms to patrilineal lines of descent.[9] In his view, the emergence of private property and the mastery of animal

[9] Alternative theories for the factors that contribute to the adoption of matrilocal versus patrilocal norms focus on the relative contribution of men and women to subsistence (Murdock 1949). However, empirical support for these theories is limited (Ember and Ember 1971). In a more recent paper, Guha (2010) argues that patrilocality emerged

husbandry and intensive agriculture allowed men to seize control of this wealth, and claim exclusive paternity of their children. It is believed that the widespread adoption of cattle-based farming generated a switch from matriliny to patriliny. As laid out by Holden and Mace (2003), the domestication of animals led to greater accumulation of wealth, and enabled the wealthiest men to engage in polygyny. [10] Thus, the desire for a large clan by the wealthy eventually led to their shift towards male-biased inheritance customs. This is a modified version of the 'Trivers-Willard' hypothesis, which predicts that there is a feedback between the sex ratio at birth and the anticipated health of offspring. The logic is that in a competitive mating environment, males in good condition out-produce females in good condition, but females in poor condition out-produce males in poor condition. In analogous fashion, the adoption of cattle and the resulting heterogeneity in wealth would have served to support patrilocal norms. Wealthier parents would prefer bequeathing their land to sons who could support multiple wives. The strength of patrilocal norms is therefore intimately related to the fertility of land to support a dense population, as is found in many fertile areas of China and India.

Patrilocal norms will also emerge in response to intensive agriculture as a result of the increased importance of local defense when land is valuable. Since intensive agriculture requires repeated cropping of land, it generates a fixity of settlement that necessitates defense of one's land. Attacks by nomads and groups looking to expand their territory would have led to both the need for sons for defense, and patrilocal norms, as they would need to remain nearby for defense. Therefore, rigid patrilocality and having history of clan-based conflict are likely to be correlated.

---

as a way of dealing with paternity ambiguity, since the men working in the fields would have been unable to monitor their wives sexual activity, requiring the husband's parents to maintain this role.

[10] Holden and Mace (2003) also present empirical evidence supporting this hypothesis. They exploit variation across Africa in tribe norms and find that the innovation of cattle directly led Bantu- and Bantoid speaking tribes from Sub-Saharan Africa to switch from matriliny to patriliny.

Note that concerns over frequent attacks by nomadic groups were a primary motivation for the construction of the Great Wall of China, consistent with this logic.[11]

The adoption of patrilocal norms serves to keep the son close for defense, to provide him wealth to support one or more wives, and to care for his elderly parents in their old age. But why does the son remain in this role of caregiver for his parents on their land, instead of heading for "greener pastures" as was common among Nomadic tribes? The answer lies in the valuable land he stands to inherit from his parents and the intensivity of agriculture. When land is scarce, it becomes valuable and this creates strong ties between fathers and sons, where fathers need their sons in times of war and as caregivers in their old age, and sons need their fathers' land. Often, this implicit barter will be further reinforced by religious norms that develop to enforce the agreement, such as Confucianism's focus on filial piety, or the need for a son to pray for the dead in Hinduism. While some scholars have looked at these ideologies as factors explaining son preference, I would argue that they are more appropriately thought of as symptoms of the economic system which requires faithful participation by each generation. The central role of respect for elders in Confucian society is logical, and particularly important because intensive agriculture generates a retirement window. Unlike African or Native American tribes where people worked until their death, intensive agriculture provides the elderly an asset (the land) that can provide income even after they are unable to work. The elders become the holders of capital, and can exchange this for their children's labor, mediated through a norm of coresidence.

It is also worth noting that this framework generates the slightly unintuitive prediction that societies with intensive agriculture will have relative gender equality in comparison to

[11] Globally, many agricultural groups feared attacks by nomadic groups during harvest. This occurred frequently among Native American tribes in North America, where conflict between groups with a fixed location and nomadic tribes was frequent (Flannery and Marcus 2012).

cultures with extensive agriculture, or pastoralism.[12] Since there is a lower marginal productivity of labor in intensive agriculture relative to shifting agriculture, intensive agriculture requires *everyone* to participate in preparing the land for cropping (Boserup 1965). This may lead to women participating at higher rates in subsistence farming. Insofar as those norms persist, it may be that countries with missing women ironically will not have particularly strong attitudes against women in the workplace.[13]

III. "Missing Women" and Kinship Patterns

A. Data

The data are taken from all IPUMS and DHS countries that are available, with the most recent sample used in each instance. This forms a sample of 141 surveys from 108 unique countries (Figure II).[14] In these data, I calculate for each country the coresidence rates between elderly (age 60+) men and their adult children of each sex (age 25+), with coresidence defined as either living with a son, living with a daughter, living without any children, or living with both sons and daughters. I also calculate for each country the sex ratio of children (ages 0-5), and several measures of gender equality: completed years of schooling, rates of employment, and

---

[12] In fact, the framework's prediction is that patriarchal norms will actually be strongest among those participating in pastoral forms of subsistence, such as shepherding. Since shepherds have significant wealth, and wealth will be heterogenous in the population, it stands to reason that men will dominate society, for the reasons outlined earlier. However, shepherds will have weaker incentives to bequeath their land to their children given the absence of a fixed land area. This may explain why many Middle Eastern countries have normal sex ratios at birth but much stricter norms against female involvement in the labor force or public affairs.

[13] This line of reasoning that today's norms are the historical legacy of norms across ethnic groups builds on a growing literature explaining cultural norms today with the norms of ancestral groups. For example, Alesina et al. (2013) argue that groups which adopted the plough continue to have attitudes more discriminatory against women in the workplace. They propose that the plow made sons more valuable as workers, and led to a decreased status of women. This is compatible with what I present in this paper, because empirically, places with more valuable land will have stronger patrilocal norms, value girls less, and use the plow. In my model, adoption of the plow is more of a proximate factor than an ultimate factor, however. The plow is adopted in areas where cattle has already increased the value of land and generated patrilocal norms. Therefore, land quality is the ultimate factor whereas norms dictating gender roles, such as plow use, should be thought of as proximate factors within the construct of my conceptual framework.

[14] Details on the data set are available in the online data appendix.

literacy rates among those ages 25-45. Summary statistics for all countries are available in the online appendix.

There are several comments on the data worth making prior to reporting the summary statistics. First, the paper's key insight regarding coresidence patterns is facilitated greatly by the IPUMS family relationship pointer variables, which enable a user to identify relationships between family members.[15] A similar method of using household relationship codes was applied to the DHS, which generated pointer variables between parents and children using the household roster. The details of this process and code are made available in the online appendix. Second, it is worth noting that the analysis will primarily focus on the sex ratio of children under 5. This essentially treats any discrimination that leads to higher female mortality up to age 5 as equivalent, whether that is through abortion, infanticide, or deprivation of resources. Strong evidence suggests that the vast majority of "missing girls" are aborted, but the child sex ratio, without assumption regarding how they went "missing," will include discrimination of any form that leads to lower fractions of girls in the population of children under 5. Third, my primary measure of coresidence is the fraction of men over 60 who live with their son who is at least 25 and do *not* live with a daughter who is at least 25. The results are very robust to logical alternatives to this choice, such as defining "elderly coresidence" as only occurring between a married son and his parents, or considering a parent to be coresiding if there are both adult sons and adult daughters in the household. Note also that the choice of a person's father as defining coresidence instead of a mother was less critical, as divorce is rare in the countries in question, implying that fathers and mothers both live with adult sons. Last, the decision that a son who is

[15] https://usa.ipums.org/usa/chapter5/chapter5.shtml

25 is “coresiding” is rather arbitrary. The results are highly robust to using different assumptions regarding the exact definition of coresidence.[16]

B. Summary Statistics

In Table I, I present sample averages for select IPUMS countries. In Panel A, I report summary statistics for the countries in the sample with high sex ratios, including Armenia, China, India, and Vietnam. Notably, all have high rates of elderly coresidence with sons. The countries are not, however, remarkable in terms of gender equality. For example, Armenian women earn more years of schooling than Armenian men (13.34 versus 12.99 years). Chinese women enjoy employment rates of 90%. While gender equality is still a challenge and women are by no means equal in these countries, they are unremarkable in their relative standing in our large sample of countries. In Panel B, I report summary statistics for countries where women appear to enjoy lower equity: Iran, Morocco, and Turkey. In all three, men are more educated and have much higher employment rates. For example, in Turkey, women have 2 fewer years of education and are 44 percentage points less likely to work than men. The contrast is even sharper in Morocco, where women are 62 percentage points less likely to work. Yet the sex ratio in Turkey and Morocco is normal. Interestingly, both countries have elderly coresidence rates that are more balanced than the countries with higher sex ratios. For example, 22% of Moroccan elderly live with sons and 14% live with daughters, suggesting parents are likely to coreside with a child of either gender.

C. Examining Variation Across Countries

In Table II, I examine the relationship between coresidence rates and sex ratios of children across countries. The left-hand side in every regression is the sex ratio of children

[16] Results using different ages, or excluding unmarried sons, are available upon request.

(males per 100 females) and the right-hand side is the share of elderly men living with an adult son. In the first specification, which includes all countries and weights them equally, I estimate that an additional 10 percentage points of men who coreside with a son is associated with 1.3 more boys per 100 girls. To give a sense of what this implies, note that in Table I it was reported that China has roughly 6 more boys per 100 girls than the US (110 versus 104). This suggests that nearly the entire difference in sex ratios between China and the US could be explained by the additional 41 percent of Chinese elderly living with sons (48% versus 7%). The results are highly robust to weighting the regressions (column 2), continent fixed effects (column 3), changing the sample to only IPUMS (columns 4-6), only DHS (7-8), or dropping OECD countries (9)-(11).

In Table III, I further examine the hypothesis that coresidence and kinship patterns better explain missing women than gender inequality. In column 1, I reproduce a preferred specification from Table II, in which I use the full sample of IPUMS/DHS countries and weight each sample by the number of observations. I then add an increasing rich set of measures of gender equality, adding relative education, employment rates, and literacy rates in columns 2-4. The coefficients are actually larger once controls are included. This may be surprising but since education gaps are actually *lower* in countries with high sex ratios, it leads to a negative sign on the education ratio, increasing the point estimate of coresidence (see Figure III). In specifications with the aforementioned controls for gender equality, I estimate that an additional 10 percentage points of coresidence would increase sex ratios by 1.0 – 2.4 male children per 100 female children. Finally, I perform an additional check of running the regressions separately for all countries in Asia versus those in non-Asia, including all the gender equality controls. The coefficient estimates are surprisingly similar (24.5 vs 20.7) between the two groups, suggesting that coresidence is correlated with sex ratios even outside of Asia.

D. Examining Variation Within Countries

In order to further explore the relationship between coresidence patterns and sex ratios, I examine several sources of sub-national variation. In China and India, the data provide indicators of an important marker of culture: ethnicity and religion. Chinese ethnic groups vary substantially in their norms regarding coresidence, though most continue to favor living with sons (Figure IV). Similarly, in India, members of different religions exhibit different coresidence patterns; for example, Jains exhibit high coresidence rates with sons and high sex ratios, with over 60% of elderly men living with a son. At the other end of the spectrum, Zoroastrians have normal sex ratios and low coresidence rates. I probe this relationship more rigorously in Table IV, where I estimate models similar to those presented earlier, but each observation is a Chinese ethnic group or Indian religion. The results indicate that even after controlling for measures of gender equity, groups with higher coresidence rates have higher sex ratios at birth.[17]

In Table V, I explore the relationship between coresidence patterns and sex ratios in Sub-Saharan Africa. In Africa, many countries are home to a diverse group of cultures living within the same modern borders. This presents an opportunity to exploit sharper differences in coresidence norms, even within the same country, than what is available in China or India. In Table V, I estimate models using a pooled sample of each African country in the IPUMS sample, with each observation being a region of a country. The regions are specified to be the most granular geographic region available by country. In the table, I present results with country fixed

[17] In other specifications, I imputed coresidence rates to individual observations from the average coresidence rate for the religious group. This was found to be correlated with the sex ratio of children, even in models with district fixed effects. Results are available from the author upon request.

effects for three age groups: ages 0-4, 5-9, 10-14. Since most regions of sub-Saharan Africa do not have access to ultrasound or sex-selective abortion, it is likely that gender preferences manifest themselves over the course of a child's life. In particular, the table reveals that sex ratios of children are *rising* at older ages: the sex ratio of children rises from 102.2 to 103.6 to 106.4 in the three age groupings. This is not biologically normal, since boys are more fragile than girls and should suffer higher mortality at every age. This suggests that son preference is manifesting itself in a less obvious manner, relative to the context in China and India. However, it is observable that the sex ratio of children is higher in regions with higher coresidence rates with sons. As shown in columns 1 and 2, the sex ratios among young children (ages 0-4) are not correlated with coresidence patterns. However, in columns 3-6, the results indicate a significant positive association between areas with higher sex ratios and higher coresidence rates with sons. While causality is not demonstrated by these results, they demonstrate that coresidence norms are intimately linked to the cultural roots of son preference. In the next section, I examine whether coresidence norms are the result of son preference, or are actually the primary factor behind son preference, as this paper claims.

IV. Why Sons? Examining the Origins of Patrilocality

In light of the evidence that the "missing girls" phenomenon is related to the custom of patrilocality, it begs the question: why do elderly parents live with sons and not daughters? In this section, I explore the historical origins of this custom and its relationship to modern day sex ratios.[18] I present evidence that the phenomenon of son preference is rooted in a tradition of old age support provided by sons in areas where land is valuable. I first present general evidence from the international context using Murdock's cultural atlas. Second, I examine variation across

[18] See Das Gupta et al. (2003) for a thorough discussion of kinship systems in countries with high sex ratios.

regions of Africa which are known to have varying degrees of patrilocality (Murdock 1959). Further evidence supporting my hypothesis is available in the Online Appendix.

As shown in Figure VI, patrilocal norms (instead of norms towards balance or matrilocality) are increasingly more common with respect to each ethnic group's reliance on agriculture and animal husbandry, and decline with respect to their reliance on hunting, gathering, and fishing. This is consistent with the notion that patrilocality is rooted in the value of land, as societies with agriculture and animal husbandry will be less likely to be nomadic and will invest more heavily in their land. I present additional evidence of the deep connection between patrilocality and intensive agriculture in the Online Appendix.

But what is the relevance of patrilocal norms to demographic patterns today? While modern China and India do not have sufficient variation in their origin cultures to explore this further, Sub-Saharan Africa is home to over 800 original tribes which can be grouped into 105 original culture groups, according to Murdock's analysis of Africa (1959). In Figure VII, we use a digitized version of Murdock's map provided by Nunn (2010) to link different parts of Africa to the groups that historically inhabited these areas (Nunn 2010).[19] For each culture group, I observe each tribes residence norm following marriage (e.g. patrilocal or not) and calculate the fraction of tribes that were patrilocal in each culture group; this is the background of Figure VII, which generates the colors of the areas corresponding to tribe norms.[20] Overlaid on this figure are the modern boundaries of IPUMS regions, allowing me to consider whether historical patrilocal norms are in fact responsible for the modern phenomenon of living with sons. The conceptual framework predicts lower rates of elderly coresidence in Africa, in light of the fact

[19] Special thanks to Nathan Nunn, who digitized the original Murdock map of Africa by tribe, and the Harvard Data Library for hosting the file.
[20] Not all of the 800 tribes contain information on their residence norms, so this aggregation allows me to assign a patrilocal tribe share to all of Africa.

that the land is less conducive for intensive agriculture. While coresidence rates *are* lower in Africa (relative to Asia), there is considerable variation within the continent in this pattern. Using sub-national boundaries available from IPUMS, I calculate the coresidence rates for the modern inhabitants of these regions. These regions are then assigned a patrilocal tribe share according to the overlap of tribe location within the IPUMS region. Note that this since the culture group boundaries do not overlap perfectly with the regional boundaries, this requires in certain cases that the calculated patrilocal tribe share in a region be weighted by the share of the territory, calculated using a spatial averaging with ArcMap GIS software.[21] The results of this test are shown in Figure VIII, where we observe a striking relationship between the modern coresidence rates and the areas that were home to patrilocal tribes (Panel A). This can be compared to a "first stage," where patrilocal tribe norms are treated as an instrument for modern coresidence. In Panel B, I reproduce the OLS relationship between coresidence and the sex ratio of children observed in Table V. In Panel C, I present the "reduced form," where I present evidence that patrilocal tribe norms continue to have relevance today; the Murdock coding of tribe norms has predictive power of modern sex ratios. While this does not resolve whether son preference is causing coresidence or the reverse, this is suggesting that norms that have existed for thousands of years have persistence in modern times.

V. What Happens When Coresidence Expectations Change? Two Case Studies

A natural concern regarding this paper's central argument is that son preference is not the result of parents living with sons, but rather parents choose to live with sons because they prefer

[21] This was executed using an ArcGIS zonal command where the IPUMS region was assigned an average patrilocal tribe share using the proportion of each language group found in the region. So, for example, if a region was 2/3 covered by the Berbers and 1/3 covered by the Akan, the zone would be assigned a patrilocal tribe share equal to 2/3 the Berber rate and 1/3 the Akan rate.

sons. In this line of thinking, coresidence patterns are the byproduct of preferences rather than the main driver. A natural way to test this hypothesis is to examine a context where parental coresidence patterns changed abruptly, and to test whether there was a corresponding change in the sex ratio at birth. If, as this paper argues, son preference is primarily related to concerns over old age support, then expansions to social insurance will lead to normalizations in the sex ratio at birth, and declines in the availability of public support will lead to increases in the sex ratio at birth. The recent experiences of South Korea and Armenia provide context in which policy change dramatically altered perceptions of the state's role in caring for the elderly, and provide an opportunity to examine this paper's hypothesis in a "natural experiment" framework.

A. Pension Reform in South Korea

South Korea is a country with historically high rates of coresidence between sons and elderly parents that engaged in a massive expansion in pension coverage in 1995. The expansion applied to self-employed workers and required every citizen to contribute 9.5% of income to the pension program, representing a landmark shift in how the country provided for its elderly population. This also obviated the need for elderly to live with sons, since it provided a guaranteed stream of wealth during retirement for all citizens. As shown in Figure IX, coresidence rates have begun to decline in South Korea. Insofar as this paper's main hypothesis regarding the motivation for having sons is related to coresidence, sub-populations which will have larger declines in anticipated coresidence rates should have larger decreases in the sex ratio at birth.

In Table VI, I examine the impact of South Korea's pension expansion in a differences-in-differences setup. While wage and salary workers were covered by existing pension schemes, the addition of self-employed workers to the mandatory pension program was a significant change in policy.[22] Roughly a quarter of workers in South Korea are self-employed, making this an ideal opportunity to consider whether changing the probability of needing to rely on sons for old age support changes the sex ratio of children. As shown in Table VI, while 54.8% of births to self-employed persons were male prior to the expansion, only 52.9% were male after 1995. This can be compared to a relatively stable ratio of male births among salaried workers who had been covered, which increased slightly from 53.5% to 53.8%. I estimate that the introduction of mandatory social security was associated with a 2.21 percentage point decline in the male fraction of births [(54.8-52.9)-(53.5-53.8)], and this point estimate is statistically significant at the 5% level.

While the data available in the South Korean census are not sufficiently rich to conduct a more complete analysis of trends in births prior to the change, the evidence presented here is strongly suggestive that parents in Korea were responsive to the policy change.[23] In fact, the rapid normalization of the sex ratio suggests that the "missing girls" phenomenon can be properly framed as a problem of missing social insurance, consistent with evidence presented from China that parents without sons are deeply concerned about the prospect of retirement with no son to provide care (Ebenstein and Leung 2010). While other scholars have focused on the decline in the intensity of son preference in Korean fertility surveys in terms of the proportion of participants who report that they "must have a son," it may be that the underlying *reason* parents feel less obliged to have a son is related to mitigated concerns over financial support when they are elderly (Chung and Das Gupta 2007).

---

[22] The pension scheme in Korea was initially introduced for public officials and military personnel in the early 1960s. The National Pension Program, as a compulsory scheme for private firms with ten or more employees, was introduced in 1988. In 1992, the compulsory coverage was expanded to cover workplaces with five or more. In July 1995, the compulsory coverage was extended further to cover the self-employed living in rural areas, including farmers and fishermen. I exploit the last phase of the rollout.

[23] I also have little information regarding people's expectations of the policy change. Expectations of an expansion would presumably attenuate the estimated effect of the policy, suggesting that the policy's estimated impact on the sex ratio would be even larger if parents had children in the early 90s expecting the policy change.

B. Collapse of the Soviet Union and Sex Ratio Imbalance in Armenia

The sudden collapse of the Soviet Union in 1991 led to independence for Armenia, and many other countries in the Former Soviet Union. The guarantees of social protections vanished as well, leaving citizens concerned for their welfare. Roughly coincident with the collapse, Armenia's sex ratio at birth began to rise, as did the sex ratio in neighboring Georgia and Azerbaijan. Scholarship has been puzzled by the roots of this phenomenon, but I argue in this paper that this is the result of coresidence norms favoring sons. I examine this first in Figure X, where I plot coresidence rates and the sex ratio of children by districts of Armenia's capital Yerevan, using the 2001 Armenia census. The survey was carried out well after the collapse of the Soviet Union, where parents realized that old age support would be provided by the family rather than the state. As shown in Figure X, districts with higher coresidence rates have higher sex ratios at birth, consistent with an interpretation that concerns over old age support are driving sex selection. In Table VII, I extend this argument a step further and examine whether the sudden exposure of Armenians to having no social protections was partly responsible for the increase in the sex ratio at birth. As a proxy for parental coresidence expectations, I stratify the sample by the father's education. I posit that Armenians with less education would be expected to have their coresidence expectations increase more than more educated Armenians for two reasons. First, lower-educated parents are more traditional and presumably are more likely to participate in traditional residence norms where children live with and care for their elderly parents. Second, with the transition to a market economy, those less educated presumably fared worse, and so the switch would make them feel more vulnerable than their more-educated counterparts. This makes education a natural way to stratify the data.

As shown in Table VII, prior to the Soviet collapse, the sex ratio of children was normal. The male fraction of births among low educated and high educated parents was 50.6% and 51.0% respectively. However, among children born between 1992 and 2001, the sex ratio jumps precipitously in the lower educated group to 53.1%. Among the high educated, the sex ratio remains near the natural rate, at 51.2%. The differences in differences estimate of the impact of the collapse is 2.2%, and is statistically significant at the 1% level. The result strongly suggests that parents desire for a son increased in the wake of the Soviet collapse and since abortion was widely available during the Soviet period, it is unlikely that this phenomenon is explained by changing access to ultrasound technology. It is difficult to prove that the demand for sons increased as a result of the Soviet collapse, but the timing of the increase, and the affected group of parents are both consistent with an interpretation that parents in Armenia went through an experience that was the mirror image of what occurred in South Korea. In this interpretation, the removal of the safety net leaves families scrambling to provide for their retirement through family, and results in sex-selective abortion to ensure the birth of a son.

VI. Conclusion

This paper examines the relationship between "missing girls" and cultural norms favoring coresidence with sons during old age. I present evidence that sex ratios and coresidence rates are positively correlated when looking across countries, within countries across districts, and within districts across ethnic groups. Standard measures of gender inequality, such as relative education or employment rates, are only weakly correlated with sex ratios, suggesting that attributing the phenomenon to "son preference" is insufficient. In light of the relationship between patrilocal norms and the phenomenon of "missing girls," I examine the roots of this custom. I find that ethnic groups practicing intensive agriculture were more likely develop norms dictating that sons inherit land, and this is linked to the modern expectation that sons will care for elderly parents. I then turn to a two case studies of changing coresidence expectations. I present evidence that South Korea's pension expansion led to a decline in the sex ratio at birth, whereas the Soviet Union's collapse is implicated in the rise in the sex ratio at birth in Armenia.

The findings of this paper lend further support to the belief that the "missing girls" phenomenon is related less to religion or culture and more to the role played by sons in caring for their elderly parents. While initiatives aimed at improving female status are noble, expanding social insurance is likely to be more effective at normalizing the sex ratio. The recent experience in South Korea suggests that social change can occur quite rapidly when effective public policy reduces the economic need for sons. Likewise, the Armenian experience documents the dangers facing policymakers when social protections are removed among parents who have access to sex selection technology.

References

Alesina, Alberto, Paola Giuliano and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *The Quarerly Journal of Economics* 128(2): 469-530.

Alsan, Marcella. 2013. "The Effect of the Tse Tse Fly on African Development". Unpublished manuscript.

Ashraf, Quamrul and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review, American Economic Association* 105(10):2003-41.

Bachofen, Johann Jakob. 1861. "An English Translation of Bachofen's Mutterecht (Mother Right): A Study of the Religious and Juridical Aspects of Gynecocracy in the Ancient World." Edwin Mellow Publishers, Translated by David Partenheimer.

Boserup, Ester. 1965. *The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure*. London: George Allen & Unwin Ltd.

Chung, Woojin and Monica Das Gupta. 2007 "The decline of son preference in South Korea: The roles of development and public policy." *Population and Development Review* 33 (2007): 757–783.

Coale, Ansley and Judith Banister. 1994. "Five decades of missing females in China." *Demography* 31(3): 459–479.

Das Gupta, Monica, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung and Bae Hwa -Ok, 2003. "Why is Son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea." *Journal of Development Studies* 40(2):153-187.

Diamond, Jared M. *Guns, Germs, and Steel: The Fates of Human Societies*. W.W. Norton, New York, 1st edition, 1997.

Duthe, Geraldine, France Mesle, Jacques Vallin, Irina Badurashvili, and Karine Kuyumjyan. 2011. "High Level of Sex Ratio at Birth in the Caucasus. A Persistent Phenomenon?" *Working Paper presented at PAA Conference – Washington D.C, March 2011.*

Ebenstein, Avraham and Steven Leung. 2010. "Son Preference and the Missing Market for Social Insurance: Evidence from China's Rural Pension Program," *Population and Development Review* 36(1):47-70.

Ember, Melvin and Carol R. Ember. 1971. "The Conditions Favoring Matrilocal versus Patrilocal Residence," *American Anthropologist* , New Series, Vol. 73, No. 3, pp. 571-594.

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

Flannery, Kent and Joyce Marcus. 2012. *The Creation of Inequality: How Our Prehistoric Ancestors Set the Stage for Monarchy, Slavery and Empire.* Cambridge, MA: Harvard University Press.

Guha, Brishti. 2010. "*Patrilocal Exogamy as a Monitoring Mechanism: How Inheritance and Residence Patterns Co-evolve."* Research Collection School of Economics, Singapore Management Univeristy.

Guilmoto, Christophe Z. 2012. "Son Preference, Sex Selection, and Kinship in Vietnam." *Population and Development Review* 38(1): 31-54.

Holden, Clare Janaki and Ruth Mace. 2003. "Spread of cattle led to the loss of matrilineal descent in Africa: a coevolutionary analysis," *Proc. R. Soc. Lond. B*, 270:2425-2433.

Holden, Clare Janaki, Rebecca Sear, and Ruth Mace. 2003. "Matriliny as daughter-biased investment." *Evolution and Human Behavior*. Vol 24. 99-112.

Lubbock, Sir John. 1870. *The Origin of Civilisation and the Primitive Condition of Man*, London: Longmans, Green, and Co.

Morgan, Lewis H. 1877. *Ancient Society,* London: MacMillan & Company.

Murdock, George. 1949. *Social Structure*, New York: Macmillan.

Murdock, George. 1959a. *Africa, its Peoples and their Culture History*. McGraw-Hill, New York.

Murdock, George. 1959b. *Tribal Map of Africa*. Digitized copy made available by Harvard University.

Murdock, George and Douglas White. 1969. "Standard Cross-Cultural Sample." *Ethnology* 9:329-369.

Nunn, Nathan. 2010. Online representation of Murdock (1959). Hosted by Harvard University. Downloaded from : http://worldmap.harvard.edu/data/geonode:murdock_ea_2010_3

Putterman, Louis. 2008. "Agriculture, Diffusion and Development: Ripple Effects of the Neolithic Revolution." *Economica* 75(300): 729–748.

Qian, Nancy. 2008. "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Income on Sex Imbalance," *The Quarterly Journal of Economics* 123(3).1251:1285.

Sen, Amartya. 1990. "More than 100 million women are missing." *New York Review of Books* 37: 61–66.

Trivers, R. L. and D.E. Willard. 1973. "Natural selection of parental ability to vary the sex ratio of offspring". *Science* 179(4068): 90–92.

Waldron, I. 1985. “What do we know about causes of sex differences in mortality? A review of the literature.” *Population Bulletin of the United Nations* 18: 59–76.

World Economic Forum: *Global Gender Gap Report 2013.* Accessed 3/12/14: *http://reports.weforum.org/global-gender-gap-report-2013/#*

Zeng, Yi, Tu Ping, Gu Baochang, Xu Yi, Li Bohua and Li Yongpjing. 1993. “Causes and Implications of the Recent Increase in the Reported Sex Ratio in China.” *Population and Development Review* 19(2): 283-302.

Table I

Sex Ratios of Children, Coresidence Patterns, and the Status of Women in IPUMS Countries

Country	Sex Ratio Under 5	Coresidence Differences: Son	Coresidence Differences: Daughter	Coresidence Differences: Diff	Schooling Differences: Men	Schooling Differences: Women	Schooling Differences: Diff	Employment Differences: Men	Employment Differences: Women	Employment Differences: Diff	Gender Inequality Index
Panel A: Select Countries with High Rates of Coresidence with Sons											
Armenia	113	0.49	0.10	0.39	12.99	13.34	-0.3	0.63	0.46	0.16	0.36
China	110	0.48	0.09	0.39	7.79	5.97	1.81	0.98	0.90	0.09	0.21
India	109	0.56	0.06	0.50	5.81	3.60	2.20	0.95	0.44	0.51	0.62
Vietnam	108	0.41	0.11	0.30	7.58	7.36	0.22	0.95	0.86	0.09	0.32
Panel B: Select Countries with Gender Inequality											
Iran	105	0.23	0.14	0.09	8.21	6.51	1.69	0.87	0.14	0.73	0.53
Morocco	105	0.22	0.14	0.09	3.97	2.55	1.42	0.84	0.22	0.62	0.50
Turkey	106	0.28	0.10	0.18	8.21	6.01	2.20	0.85	0.40	0.44	0.42
Panel C: OECD Countries											
Israel	106	0.12	0.09	0.03	11.08	11.42	-0.3	0.80	0.61	0.19	0.14
Italy	105	0.19	0.12	0.07	10.58	10.74	-0.1	0.83	0.56	0.26	0.12
United States	104	0.07	0.06	0.02	9.23	9.02	0.21	0.86	0.69	0.16	0.29

Notes : The statistics are reported for the most recent IPUMS sample for each country. The share of elderly living with a son is defined as the ratio of men age 60+ who coreside with adult sons (age>=25). Schooling and employment are reported for men and women ages 25-45. All differences for schooling and employment are statistically significant at the 1% level. The Gender Equality Index is calculated by the United Nations and reported for 2013. Higher values imply greater levels of gender inequality.

Table II

Cross-Country Relationship Between Share of Elderly Living with Sons and the Sex Ratio of Children

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Dependent Variable: Sex Ratio of Children (ages 0-5)										
Mean of Dep. Var.	103.80	105.20	105.20	104.30	105.20	105.20	103.40	104.20	103.60	105.30	105.30
Share of Elderly Living With Son	12.88*** (3.19)	12.47** (5.13)	10.38** (5.23)	11.30*** (3.83)	12.39** (5.29)	10.48* (5.62)	15.13*** (2.99)	15.13*** (2.99)	15.22*** (3.36)	17.49*** (5.72)	11.29* (6.17)
Weights	No	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes
Continent Fixed Effects	No	No	Yes	No	No	Yes	No	No	No	No	Yes
Sample Selection	Full	Full	Full	IPUMS	IPUMS	IPUMS	DHS	DHS	w/o OECD	w/o OECD	w/o OECD
Observations	140	140	140	69	69	69	71	71	122	122	122
R Squared	0.15	0.26	0.41	0.16	0.26	0.41	0.23	0.23	0.21	0.34	0.39

Notes: Sample is composed all IPUMS and DHS samples for the most recent year available. The share of elderly living with a son is defined as the ratio of men age 60+ who coreside with an adult son (age>=25). Standard errors are reported in parentheses below the coefficients and are heteroskedastic-consistent. Weighted regressions are weighted by the number of observations in the survey. * significant at 10% ** significant at 5%. *** significant at 1%.

Table III

Cross-Country Relationship Between Share of Elderly Living with Sons and the Sex Ratio of Children Controlling for Measures of Gender Equity

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Sex Ratio of Children (ages 0-5)					
Mean of Dep. Var.	105.20	105.20	105.20	105.30	107.20	103.60
Share of Elderly	10.38**	17.73***	19.74***	24.36***	24.52***	20.65***
Living With Son	(5.23)	(5.18)	(4.64)	(5.23)	(7.83)	(6.94)
Education Ratio		-2.90**	-2.39*	-6.90**	-6.73	-3.88
(male yrs / female yrs)		(1.18)	(1.31)	(2.68)	(7.27)	(3.07)
Working Ratio			-0.44***	-0.47***	-0.56***	0.27
(male emp / female emp)			(0.14)	(0.11)	(0.09)	(0.17)
Literacy Ratio				5.05***	2.40	3.06
(male rate / female rate)				(1.78)	(10.44)	(2.40)
Sample Selection	Full	Full	Full	Full	Asia	Non-Asia
Weights	Yes	Yes	Yes	Yes	Yes	Yes
Continent Fixed Effects	Yes	Yes	Yes	Yes	No	No
Observations	140	138	126	102	24	78
R Squared	0.41	0.51	0.62	0.68	0.76	0.33

Notes : Sample is composed all IPUMS and DHS samples for the most recent year available. In row 1, the fraction living with son is defined as the ratio of men age 60+ who coreside with adult sons (age>=25). In row 2, the education ratio is the ratio of average years of education of men to women. In row 3, the employment ratio is the ratio of the employment rate of any type among men to the rate among women. In row 4, the literacy ratio is the ratio of the literacy rate among men to the literacy rate among women. In column 5, the sample is restricted to countries in Asia. In column 6, the sample is restricted to countries not in Asia. Standard errors are reported in parentheses below the coefficients and are heteroskedastic-consistent. All regressions are weighted by the number of observations in the survey used. * significant at 10% ** significant at 5%. *** significant at 1%.

Table IV

Relationship Between Share of Elderly Living with Sons and the
Sex Ratio of Children Across China's Ethnic Groups and India's Religions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable: Sex Ratio of Children (ages 0-5)							
	China's Ethnic Groups				India's Religions			
Share of Elderly	15.77	26.11***	26.17***	34.98***	49.30***	56.76**	80.20**	87.70*
Living With Son	(9.13)	(6.62)	(7.03)	(5.24)	(13.66)	(17.95)	(23.54)	(32.34)
Education Ratio		-3.558***	-3.546**	-13.98**		-6.273*	-8.443**	2.1220
(male yrs / female yrs)		(1.09)	(1.28)	(4.60)		(2.75)	(2.59)	(25.10)
Working Ratio			0.18	-5.58			-2.7430	-3.6640
(male emp / female emp)			(4.83)	(5.38)			(2.53)	(2.96)
Literacy Ratio				10.39**				-15.3100
(male rate / female rate)				(4.16)				(38.41)
Observations	17	17	17	17	9	9	9	9
R Squared	0.15	0.54	0.54	0.66	0.44	0.62	0.68	0.69

Notes: Sample is composed of IPUMS India (1983, 1987, 1993, 1999, 2004) and IPUMS China (1990). In row 1, the fraction living with son is defined as the ratio of men age 60+ who coreside with adult sons (age>=25). In row 2, the education ratio is the ratio of average years of education of men to women. In row 3, the employment ratio is the ratio of the employment rate of any type among men to the rate among women. In row 4, the literacy ratio is the ratio of the literacy rate among men to the literacy rate among women. Standard errors are reported in parentheses below the coefficients and are heteroskedastic-consistent. All regressions are weighted by the number of observations in the survey used. * significant at 10% ** significant at 5%. *** significant at 1%.

Table V

Relationship Between Share of Elderly Living with Sons and the Sex Ratio of Children Across Sub-Saharan Africa

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Sex Ratio of Children ages:					
	Ages 0-4		*Ages 5-9*		*Ages 10-14*	
Mean of Dep. Var.	102.20	102.20	103.60	103.60	106.40	106.40
Share of Elderly Living With Son	7.68	4.32	29.89**	19.75***	44.01**	31.07**
	(4.9)	(4.8)	(11.2)	(5.5)	(20.0)	(13.4)
Gender Controls	No	Yes	No	Yes	No	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1731	1729	1731	1729	1731	1729
R Squared	0.24	0.26	0.29	0.36	0.29	0.37

Notes: Sample is composed of all IPUMS samples in Africa. Each observation represents a sub-national district. Each country is broken down using the most granular geographic unit available. Standard errors are reported in parentheses below the coefficients, are heteroskedastic-consistent, and clustered by country. All regressions are weighted by the number of observations in the district. * significant at 10% ** significant at 5%. *** significant at 1%.

Table VI

Differences in Differences Estimates of the Effect of Pension Expansion on the Male Fraction of Births in Korea

	Self-Employed (added to pension)	Salaried Workers (existing pension)	Difference
	(1)	(2)	(3)
1. Children Born Before Pension Expansion (1990-1994)	0.548	0.535	0.0133**
	(0.005)	(0.003)	(0.005)
2. Children Born After Pension Expansion (1995-2000)	0.529	0.538	-0.009
	(0.006)	(0.004)	(0.007)
3. Change in Male Fraction of Births After Reform	-0.0191**	0.003	-0.0221**
	(0.008)	(0.005)	(0.009)

Notes : Data are taken 68,282 births observed between 1990 and 2000 observed in the Korean 2000 census. Sample is restricted to parents where the father and mother are at least 35 and 30 years old respectively. * significant at 10% ** significant at 5%. *** significant at 1%.

Table VII

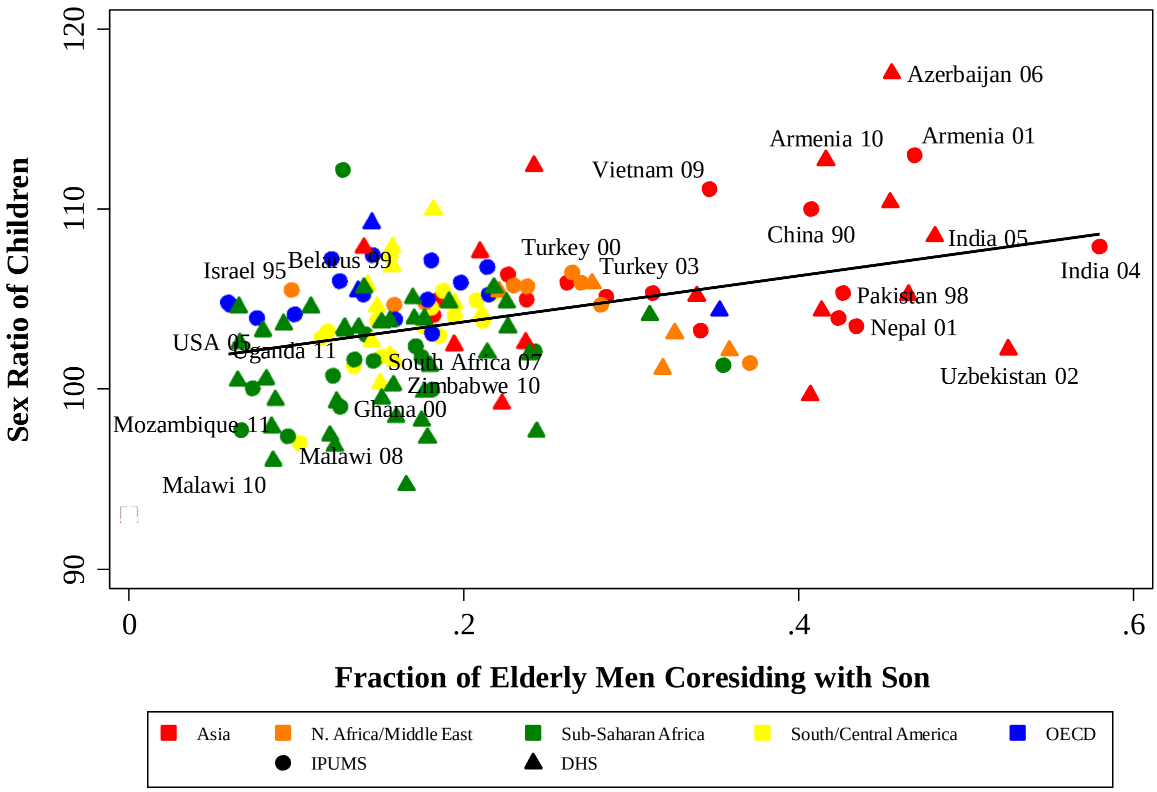
Differences in Differences Estimates of the Effect of Soviet Collapse on the Male Fraction of Births in Armenia

	Lower Educated (higher coresidence)	Higher Educated (lower coresidence)	Difference
	(1)	(2)	(3)
1. Children Born Before Soviet Collapse (1983-1991)	0.506	0.510	-0.004
	(0.003)	(0.002)	(0.004)
2. Children Born After Soviet Collapse (1992-2001)	0.531	0.512	0.0181***
	(0.003)	(0.003)	(0.004)
3. Change in Male Fraction of Births After Collapse	0.0246***	0.003	0.0220***
	(0.004)	(0.003)	(0.006)

Notes : Data are taken 134,257 births observed for children ages 0-18 observed in the Armenian 2001 census. Lower educated is defined as having less than a secondary degree. * significant at 10% ** significant at 5%. *** significant at 1%.

Figure I

Coresidence and the Sex Ratio of Children, IPUMS and DHS Samples

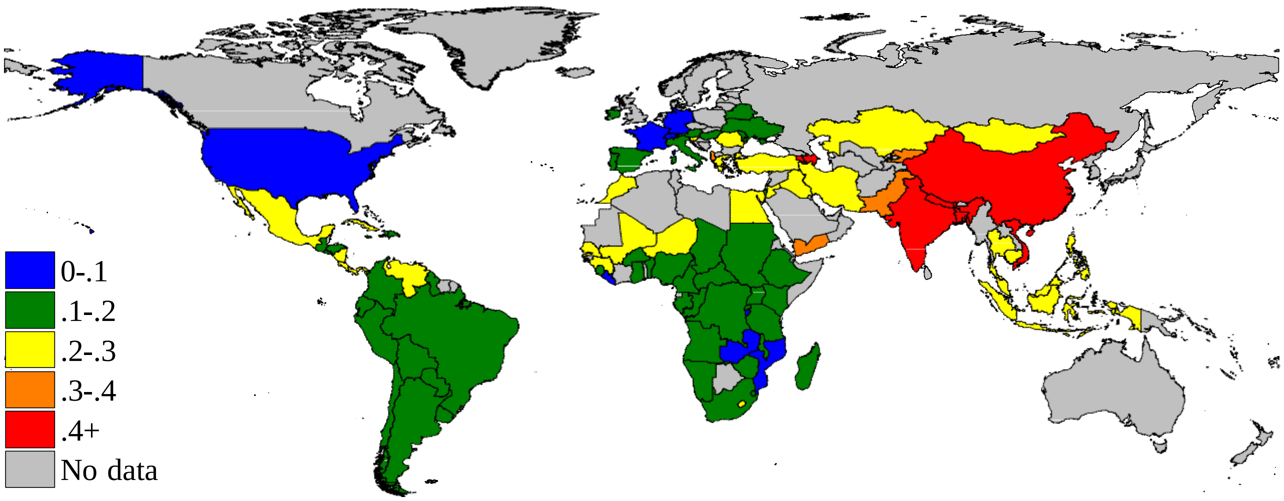


β=12.87, t=4.03, p-value<.001, N=141

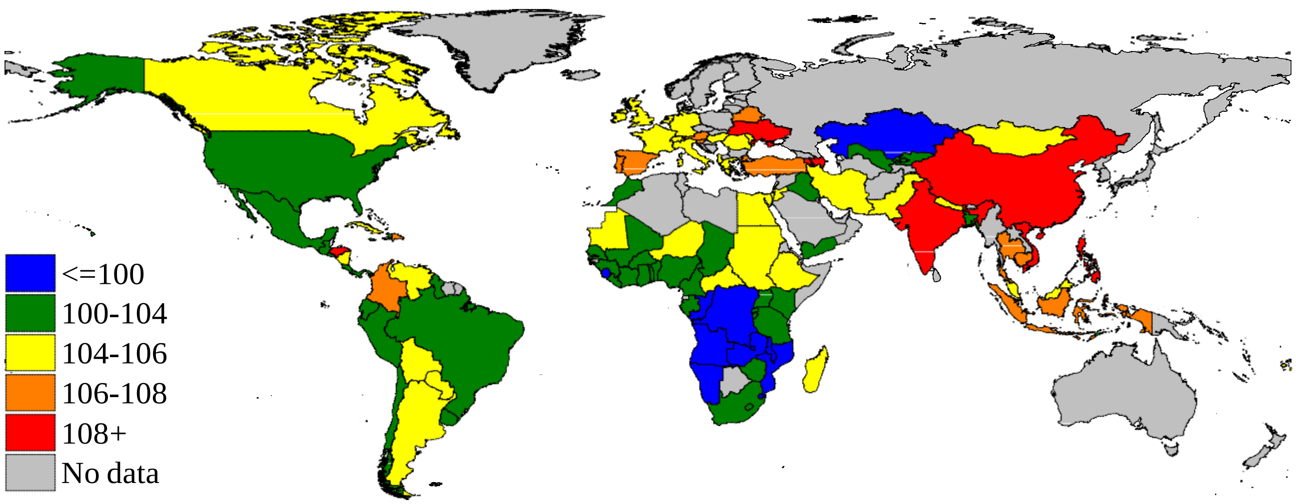
Notes: The sample is composed of 141 of the most recent IPUMS and DHS surveys for 108 unique countries. The x-axis is the fraction of men ages 60+ living with their adult son (age >=25). The y-axis is the sex ratio aamong children ages 0 to 5 in each survey.

Figure II

Panel A: Coresidence Rates by Country



Panel B: Sex Ratios by Country



Notes: Data for 103 countries from IPUMS and DHS using the most recent available survey.

Figure III

Male Fraction of Births and Relationship with Other Measures of Gender Equity

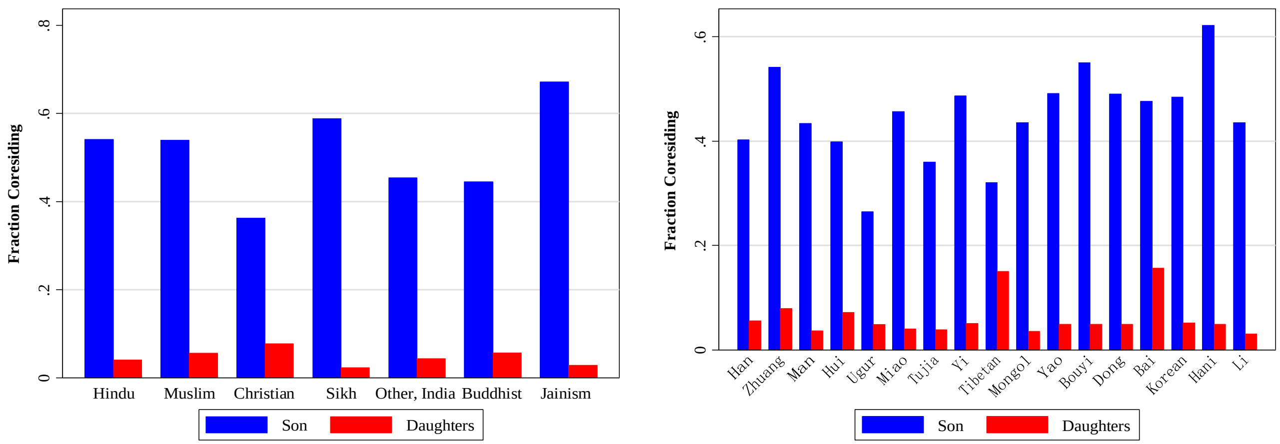
Panel A: Education Ratio *Panel B: Employment Ratio* *Panel C: Literacy Ratio*

Notes : Data for countries from IPUMS and DHS using the most recent available survey. The ratios are defined as (a) male to female education in years (b) male to female employement rate and (c) male to female literacy rates. These are calculated among individuals ages 25-45.

Figure IV

Elderly Coresidence Rates with Sons and Daughters in India and China

Panel A: India's Religions *Panel B: China's Ethnic Groups*



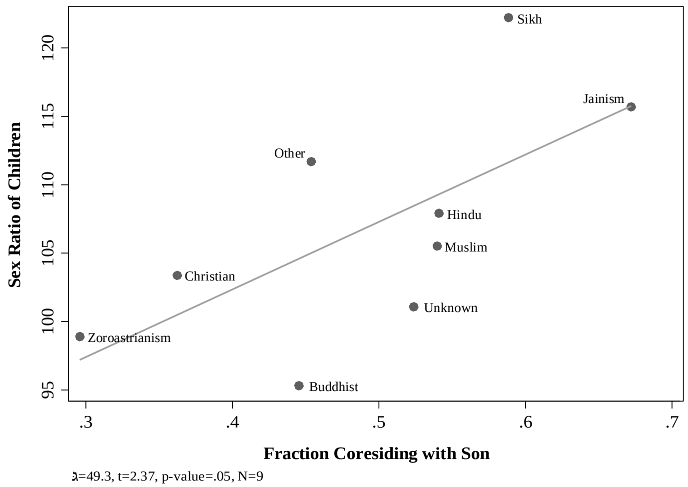
Notes: China 1990 IPUMS (10% sample), India IPUMS (1983, 1987, 1993, 1999, 2004). The y-axis is the fraction of men age 60+ living with an adult (age>=25) son or daughter.

Figure V

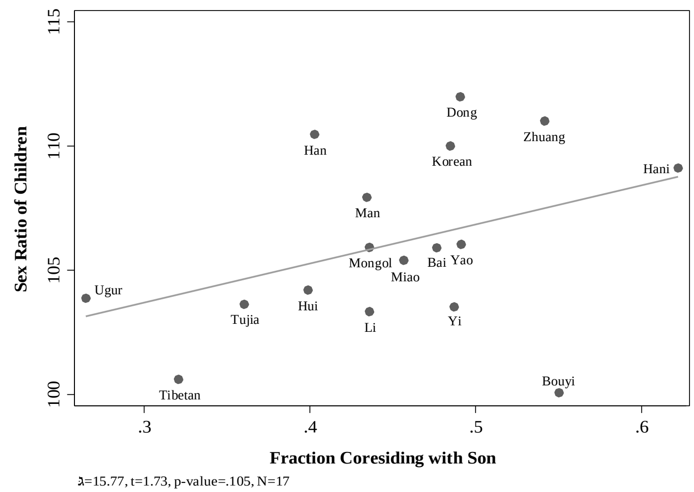
Coresidence and the Sex Ratio of Children Across India's Religions and China's Ethnic Groups

Panel A: Simple Correlation between Coresidence and the Sex Ratio

India, Religions

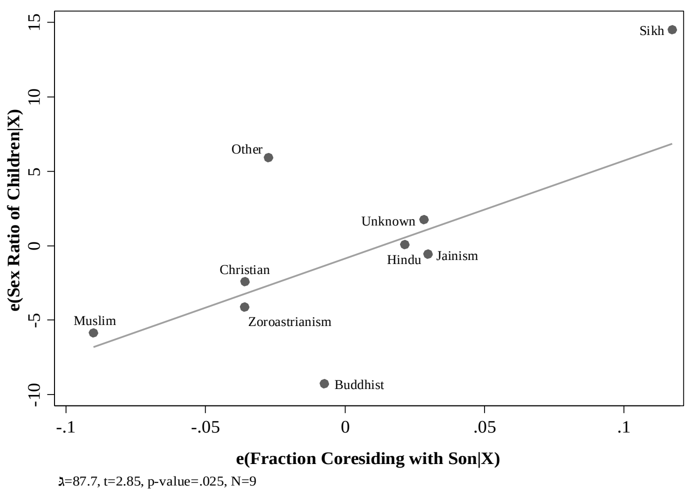


China, Ethnic Groups

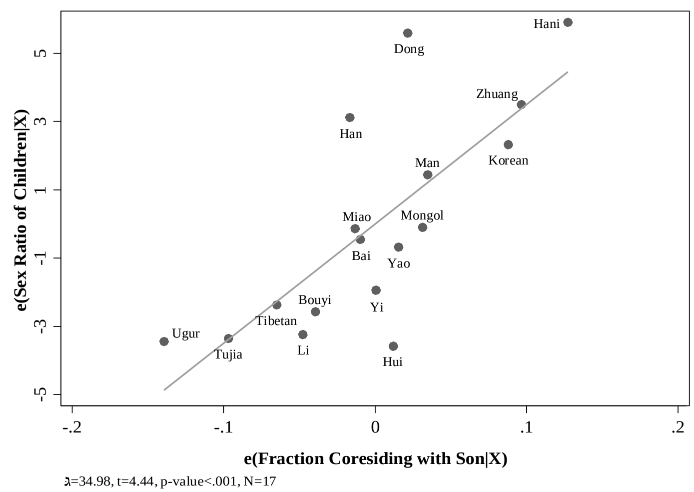


Panel B: Partial Correlation between Coresidence and the Sex Ratio

India, Religions



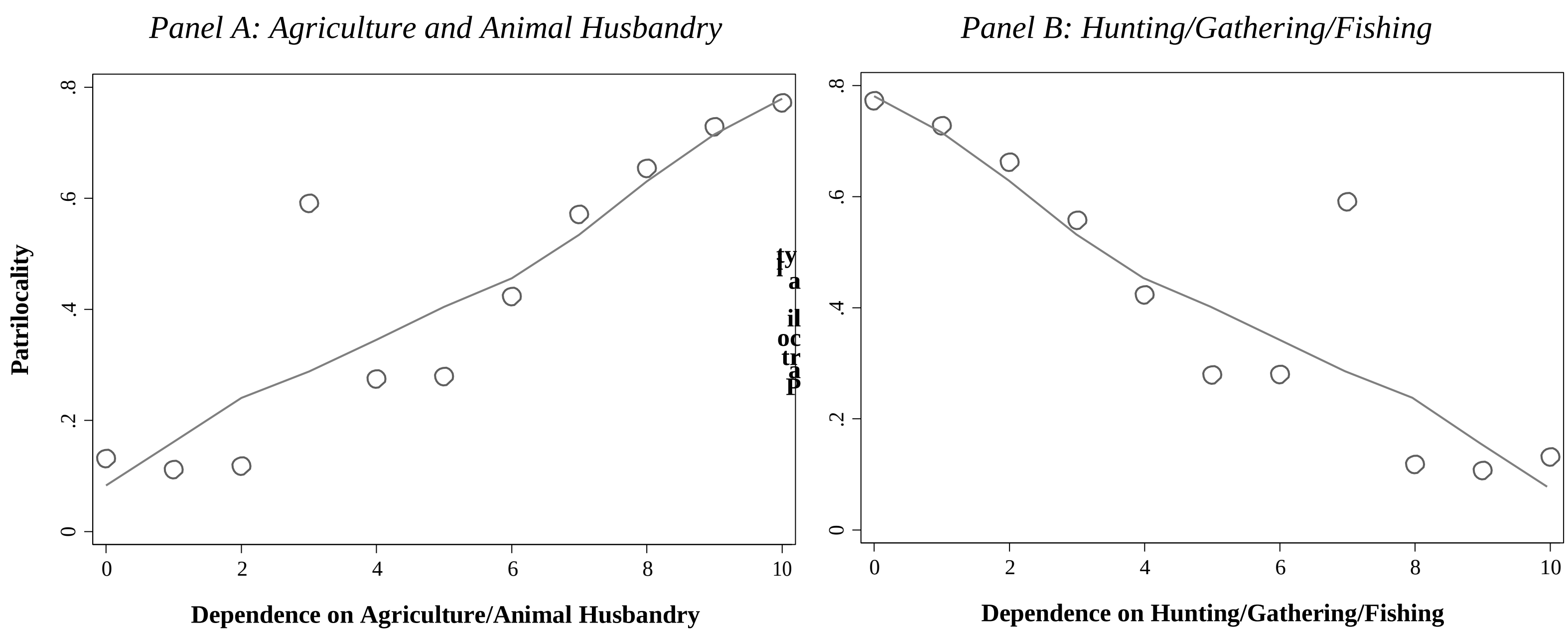
China, Ethnic Groups



Notes: Each plot reports the relationship between the fraction of men age 60+ living with an adult son (ages 25+) and the male fraction of children ages 0 to 5. Each observation for India is a religion and each observation for China is an ethnic group. In Panel B, we plot the partial correlation after controlling for controls presented in Table III: male to female education ratio, male to female employment ratio, and male to female literacy ratio.

Figure VI

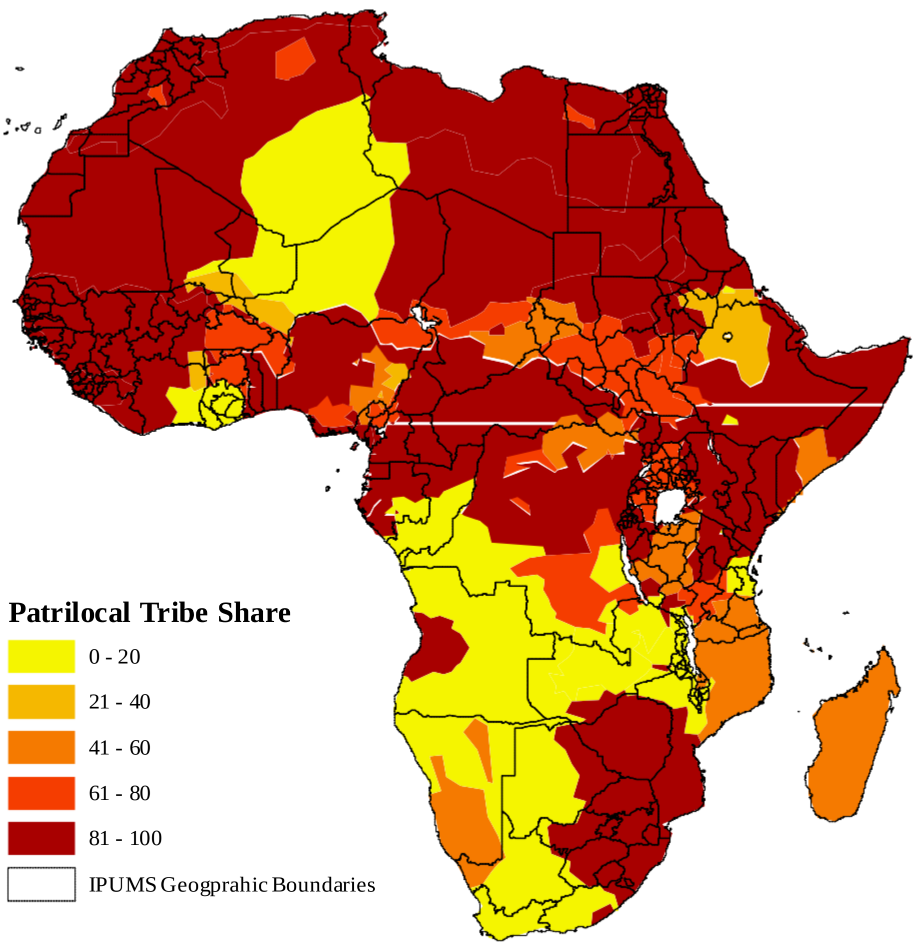
Coresidence Patterns and Subsistence Methods Across Societies



Notes : The figure plots the relationship between patrilocal norms and subsistence methods among 1,267 traditional socieites recorded in the Ethnographic Atlas (Murdock 1965). The figure is created using the lowess smoother in STATA (bandwidth .80). The index of dependence ranges from 0-9, with zero indicating 0-5% dependance and 9 indicating 85-100% dependance.

Figure VII

Residence Norms of Language Groups in Africa

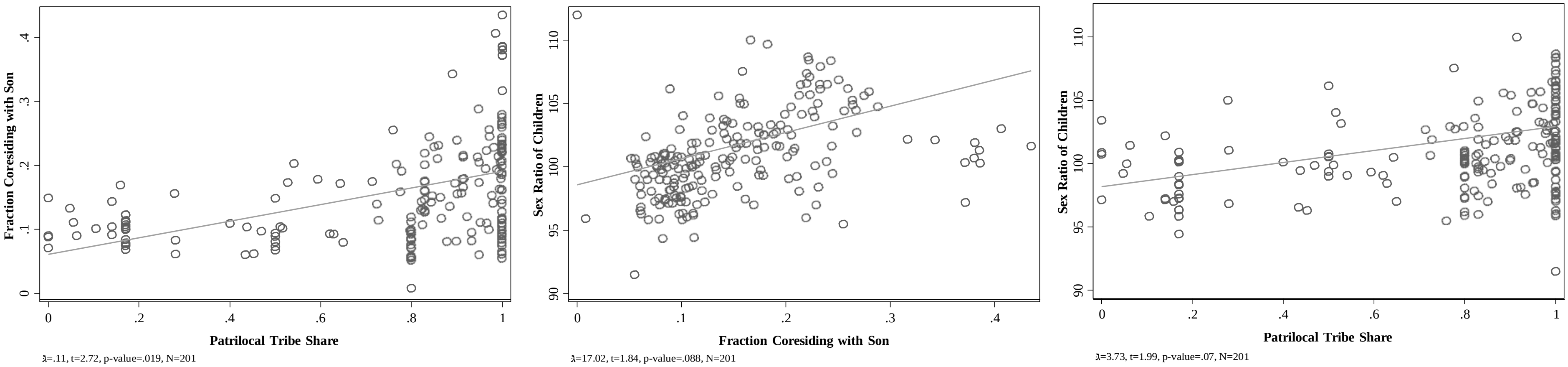


Notes : Each polygon represents a modern IPUMS geographic region, and the color shades refelct their residence customs, taken from Murdock (1959).

Figure VIII

Coresidence Patterns, Tribe Norms of Patrilocality and the Sex Ratio of Children Across Districts of Sub-Saharan Africa, IPUMS

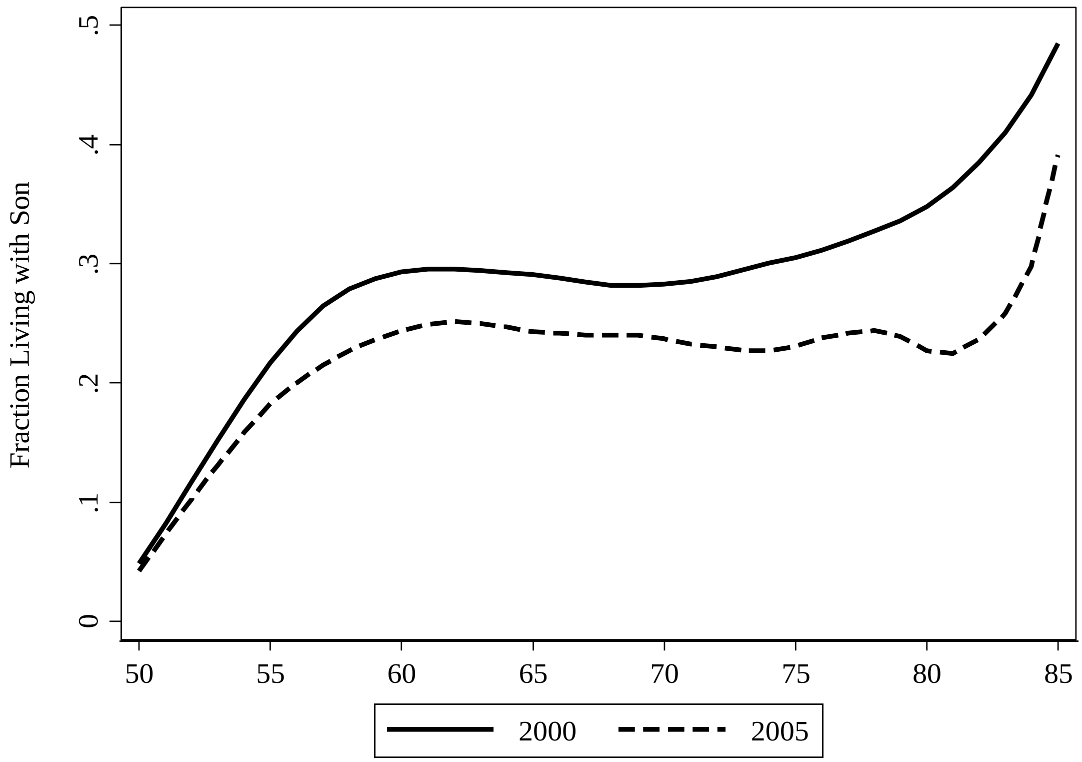
Panel A: Coresidence Patterns and Patrilocal Norms | *Panel B: Sex Ratios and Coresidence Patterns* | *Panel C: Sex Ratios and Coresidence Patterns*



Notes: Each observation (circle) is a district in Africa using the IPUMS geographic coding (GEOLEV1), which codes each country in terms of their largest sub-national geographic classification. In Panel A, we plot the correlation between the share of elderly men coresiding with an adult son against the patrilocal tribe share according to Murdock (1957). In Panel B, we plot the correlation between the sex ratio of children under 5 and the share of men age 60+ coresiding with adult sons (age>=25). In Panel C, we plot the correlation between the sex ratio of children under 5 and the patrilocal tribe share.

Figure IX

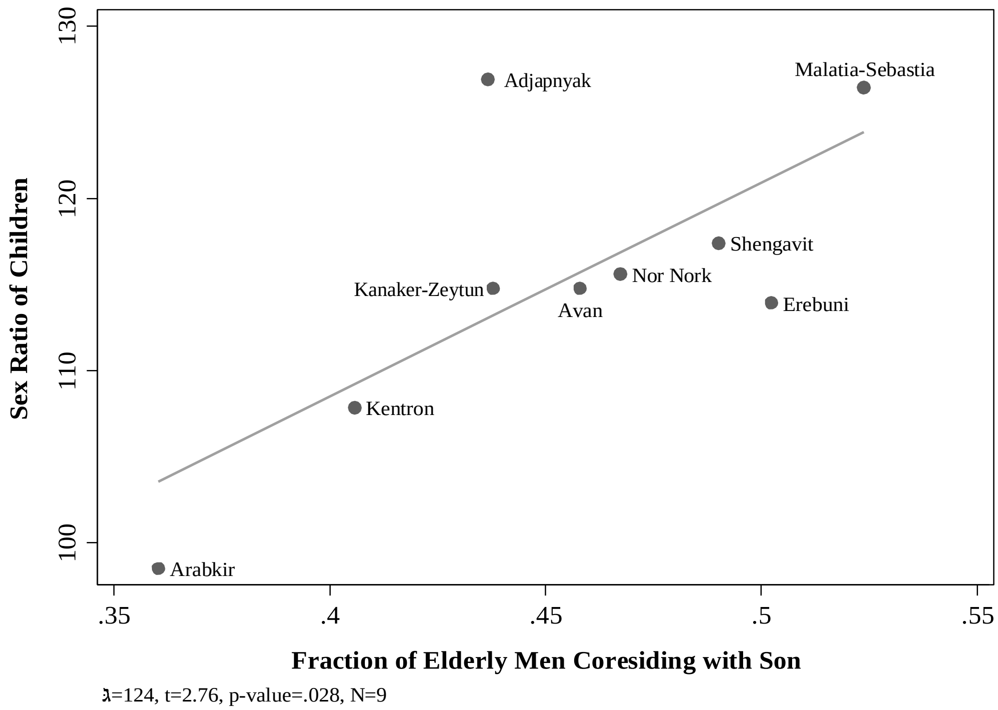
Age Profile of Elderly Coresidence with Sons in Korea



Notes : The figure plots the coresidence rate for men age 60+ for the years 2000 and 2005. The figure is created using the lowess smoother in STATA (bandwidth .30).

Figure X

Patterns in Elderly Coresidence and the Sex Ratio at Birth in Yerevan, Armenia



Notes : The plot reports the relationship between the fraction of elderly parents (ages 60+) living with an adult son (age>=25) and the male fraction of children ages 0 to 5 across districts of Yerevan, Armenia.

Web Appendix for

Patrilocality and Missing Women

Avraham Ebenstein
Hebrew University of Jerusalem

(Not for Publication)

April 2014

A1. Introduction

This appendix accompanies "Patrilocality and Missing Women" by Avraham Ebenstein. Section A2 provides further details on the data used in the papers as well as their sources. Section A4 and A5 report additional figures and tables that were discussed in the body of the paper, but were omitted due to space considerations. Section A5 provides further details on the algorithm used for calculation of coresidence rates.

A2. Data and Their Sources

A2.1. DHS and IPUMS: The primary data sources used for our analysis were obtained from from the *Demographic and Health Surveys* (DHS) collected by USAID and the *Integrated Public Use Microdata Series* (IPUMS), *International* provided by the Minnesota Population Center, University of Minnesota. The DHS is composed of multiple datasets for each country. In order to calculate coresidence rates, we used the Household Recode (HR). Each observation includes a complete roster of the family, and records for each household member the following variables: sex, age, relationship to head, and marital status. DHS also provides more specific data on each adult male (*MR)* ages 15-64 and adult females (*IR)* ages 15-49. These data sets are used to generate our calculations of gender equity (employment, education, literacy). Some countries only collected female recodes and not male recodes, and other countries only have household recodes, leading to slightly smaller sample sizes for our regressions with measures of gender equity. Certain countries were also missing education, literacy, and/or employment information and thus were excluded from the analysis with controls for gender equity. The IPUMS samples generally contained information on the necessary variables, though literacy was not available for countries. The samples used in the analysis, and the number of observations, are listed in Table A1 for DHS and Table A2 for IPUMS.

A2.2. Korean Population Data: The analysis on Korea focused on Korean census data from 2000, and a population survey from 2005. The sample sizes are 883,350 and 294,348 respectively. These data contained employment variables, enabling us to analyze the relationship between the sex ratio of children and parental occupation by year. Coresidence rates by year were calculated using the household roster.

A2.3. Armenian Population Data: The analysis on Armenia focused on the Armenian census of 2001. The sample contained 134,257 children ages 0-18, from an original data set of 326,560 observations, and was made available by IPUMS. These data contained information on education, enabling us to analyze the relationship between the sex ratio of children and father's education before and after the collapse of the Soviet Union.

A2.4. Sub-Saharan Africa and Patrilocal Tribes: The analysis of Sub-Saharan Africa is carried out by using the shapefile of Africa which is a digitized version of Murdock (1959), prepared by Nathan Nunn and made available at the Harvard Digital Library website.[1] This records 843 native groups in Africa, which can be aggregated into 105 culture groups. In the text, I produce a figure of African patrilocal norms based on these rates, and calculate the fraction of tribes that were patrilocal in each IPUMS region by calculating the weighted average of the patrilocal tribe share by area. This is used to assign to each IPUMS "GEOLEV" region an estimate of the patrilocal tribe share, which I then can use in regressions with modern-day sex ratios and coresidence rates. The listing of groups and their norms is reported in Figure A1.

A2.5. Years since Neolithic Transition: The number of thousand years elapsed, until the year 2000, since the majority of the population residing within a country's modern national borders began practicing sedentary agriculture as the primary mode of subsistence. This measure, reported by Louis Putterman (2008), is compiled using a wide variety of both regional and country-specific archaeological studies as well as more general encyclopedic works on the transition from hunting and gathering to agriculture during the Neolithic Transition. The reader is referred to www.econ.brown.edu/fac/Louis%5FPutterman/agricultural%20data%20page.htm for a detailed description of the primary and secondary data sources employed by the author in the construction of this variable.

A3. Historical Origins of Patrilocality and its Relationship to "Missing Girls"

In the manuscript, I present a conceptual framework which generates the prediction that the norms responsible for "missing girls" are intimately connected to land quality and the need for

[1] http://worldmap.harvard.edu/data/geonode:Murdock_EA_2011_vkZ

sons to inherit one's land. In this section, I present evidence supporting the predictions of the framework that are referenced in the text but not included due to space considerations.

A3.1. *Summary Statistics describing the* Ancestors of the "Missing Girls"

In Table A3, I present summary statistics for the ethnic groups that were the ancestors of groups currently engaging in sex selection against girls. These data are taken from the Standard Cross-Cultural Sample (SCCS), a subset of 186 ethnic groups taken from the Ethnographic Atlas that are meant to be representative of the overall sample. The seven cultures that are directly related to the groups currently engaging in sex selection are the Chinese, Lolos, Vietnamese, Koreans, Abkhaz, Armenians, and Punjabis. I compare these groups to all the remaining groups on the following dimensions: residence norms after marriage, subsistence methods, and fixity of settlement. The table indicates that the groups which currently engage in sex selection had several common features. These were groups that all engaged in patrilocality, relied on agriculture and animal husbandry, and had greater fixity of settlement. Interestingly, however, these groups did not have significantly lower participation of women in agriculture. Insofar as gender norms exhibit persistence (Alesina et al. 2013), these results may help explain why gender equity is not correlated with the phenomenon of "missing girls." While the need for a son to defend and inherit land is clear, women in these societies have historically contributed substantially to subsistence.

A3.2. Patrilocality, Intensive Agriculture, and Land Inheritance Norms

I examine the relationship between patrilocality and subsistence methods in Table A4. As a visual preview of the results, I present in Figure A3 the sample means of patrilocality stratified by whether the group is classified as practicing no agriculture, casual agriculture, extensive agriculture, or intensive agriculture. The figure indicates that the share of groups that are patrilocal is rising monotonically with respect to their reliance on agriculture.

In the regressions, I predict the probability that a given ethnic group will practice patrilocality using their subsistence methods as explanatory variables including no additional controls (column 1), including continent fixed effects (column 2), or including country fixed effects (column 3). As shown in Table VII, I find in Panel A that groups engaging in intensive agriculture are significantly more likely to have patrilocal norms. I measure the intensity of agriculture by classifying the members of the ethnographic atlas in four groups: 0 (no agriculture), 1 (casual agriculture), 2 (shifting agriculture), and 3 (intensive agriculture). Roughly half of the sample practices patrilocality, so the coefficient in column 1 indicates that moving from no agriculture (index=0) to intensive agriculture (index=3), the probability of practicing patrilocality increases 48.9 percentage points (or roughly the sample mean). Even with country fixed effects, I estimate that groups are 3.87 percentage points more likely to be patrilocal for each unit increase in the agricultural intensity index. In Panel B, I examine the relationship between an ethnic group's reliance on more basic subsistence methods, with the prior that groups engaging in hunting/gathering/fishing will have little inheritance, and will be unlikely to need a son to protect their land and inherit their property. Indeed, we observe that in all specifications, reliance on these basic methods is negatively associated with the adoption of patrilocal norms, and this result is robust to the inclusion of either continent or country fixed effects.

Finally, in Panel C, I examine the role of animal husbandry in inducing the adoption of patrilocality. I find, as anticipated, that greater reliance on animal husbandry significantly increases the propensity to have patrilocal residence norms. This is logical and consistent with the paper's thesis that patrilocal norms emerge from a desire to bequeath land to sons. The use of animals in agriculture allows for higher land productivity. Cattle are specifically well suited for agriculture, since cattle can be used for more productive plowing and also generates fertilizer, which can increase the land's productivity. Therefore, the presence of cattle will be associated with very strong norms towards patrilocal residence. To further demonstrate this point, I include a regressor which is whether the group's subsistence does not involve cattle. As predicted, after controlling for total reliance on animal husbandry, groups which do not use bovine animals in their subsistence are significantly less likely to be patrilocal.

The findings presented in Table A4 are further reinforced by Figure A4, which reports land inheritance norms among groups in the Ethnographic Atlas stratified by whether they are patrilocal. The results indicate that male-dominated land inheritance customs are intimately linked with patrilocality: *every* society which participates in primogeniture or other "male only" inheritance systems are patrilocal. Furthermore, patrilocal societies are more likely to exhibit male preference in land inheritance, whereas non-patrilocal societies are more balanced. This is consistent with a thesis that the customs of coresidence norms for sons is related to inheritance norms, rather than a bias against girls of a more general variety.

A3.3. Connecting the Past to the Present: Patrilocality and Modern Sex Ratios

In this section, I examine whether patrilocal norm are correlated with the modern phenomenon of ‘missing girls.’ I examine this using two very different data sources. The first is provided by Putterman (2008), who catalogues the timing of each country’s Neothilic Transition. A visual preview of these data are provided in Figure A4 which demonstrates that certain parts of the globe adopted agriculture first, presumably because the land was amenable to this type of subsistence. The visual similarity to the areas of the world that have high sex ratios is noticeable, with India, China, and the Caucuses all having early transitions to agriculture.

In Table A5, I present cross-country regressions analyzing the relationship between coresidence with sons and the sex ratio of children. We are restricted to the countries which are available in Putterman (2008), which generates a sample of 107 countries and 103 for which we can observe the reduced form. As shown in Figure A5, coresidence rates and sex ratios are higher among countries with long traditions of agriculture. The countries singled out as having high sex ratios – India, China, Armenia, Azerbaijan – all had Neolithic transitions over 9,000 years ago. Presumably, an earlier Neolithic transition resulted from having land that was more conducive to agriculture, and made a further advance to intensive agriculture more likely. The pattern in the figure is consistent with an interpretation that intensive agriculture leads to stronger patrilocal norms, and that the strength of these customs is related to the suitability of land for this form of subsistence. While myriad other factors are not accounted for in these regressions, the fact that the countries in the Caucuses share a similar agricultural history and similar demographic experience with China and India is compelling evidence in favor of this interpretation.

As shown in Table A5, I find using OLS that increasing the fraction of elderly who live with a son by 10 percentage points is associated with an increase the sex ratio at birth by 1.24 boys per 100 girls. However, once I instrument for today’s coresidence rate with the years since the Neolithic revolution, we observe that a similar increase in coresidence rates would increase the sex ratio at birth by 3.77 boys. It is worth noting that this instrument is unlikely to pass the exclusion restriction, since the timing of the Neolithic revolution reflected a region’s integration with its neighbors and other factors that could have affected the path of history (Diamond 1997). However, the results are consistent with the body of evidence that the adoption of agriculture was intimately related to the norm of living with sons, and the timing of adoption is correlated today with both coresidence rates and the 'missing women' phenomenon.

A4. Conditional Sex Ratios Worldwide

The paper contends that skewed sex ratios are driven by parents seeking a son to care for them in old age. If this were the case, one would expect to observe normal sex ratios among parents who already have a son. As shown in Table A6, this is indeed the case. Using IPUMS and DHS data, I calculated sex ratios conditioned on the sex of the existing children and the results indicate that in countries with high sex ratios, the sex ratios are normal following sons.

A5. Calculating Coresidence Rates

In order to calculate the coresidence rate of men above the age of 65 we matched adult males with their adult children using the IPUMS constructed family interrelationship variable POPLOC. POPLOC locates the father of each individual and provides the line/person number of the father. Likewise, IPUMS provides a MOMLOC variable to do a similar analysis for the coresidence of mothers with their children. For non-IPUMS datasets such as DHS and Korea, we constructed a POPLOC variable using relationship codes and is explained in further detail below.

We match the children with their respective fathers by replacing their person number with their father's person number (using POPLOC) and then merge family members by this identifying number. At this point, a roster of each father and child is created where each observation corresponds to a distinct child with information about the father and child's age and sex. Using this information we are able to restrict our definition of coresidence to a minimum father's age, maximum child's age, or a child's marital status. Once a dummy is created for a male child we calculate the mean male-fraction in the family which ranges from 0 (coresiding with only daughters, no sons) to 1 (coresiding with only sons, no daughters). Using the total values of coresiding fathers with only sons, only daughters, and both sons and daughters we can calculate the fraction of all men over 65 in these categories. While we chose fathers over 60 coresiding with only sons as our independent variable, appendix figures 1-2 show various permutations such as mothers coresiding with daughters and the ratio of all-male coresidence to all-female coresidence among older fathers.

Creating POPLOC in the DHS

The most important variable used to generate a poploc variable is the household relationship variable (hv101), which defines each household member's relationship to the head. The relationship variable takes values such as Head, Spouse, Son/Daughter, Parent, and In-Law; any value corresponding to non-relatives or an infrequent visitor will be ignored since we are creating blood-family relationships. Since the dataset is "wide", containing the information for each household member by creating new variables for every member (i.e. hv101_01, hv101_02,etc. for member 1, 2, etc.) we begin by reshaping the data set so that each household member gets his/her own row of data specified by a variable called pernum (person number). Each household member will have his/her own information (pernum, age, sex), in addition to the other household member's information.

Relationships

Consider the household in Table 1 below. We see that the only potential father in the family is the head of the household since there are three kids classified as son/daughter and a spouse. As such, we look for any children classified with the relationship of son/daughter and assign them a poploc of 1 since the father is in the head position.

Table 1

Pernum	Relationship	Age	Sex	Poploc
1	Head	40	M	0
2	Spouse	38	F	0
3	Son	20	M	1
4	Daughter	15	F	1
5	Son	8	M	1

In Table 2, the head of the household is not the father but rather the mother of the family. Since the mother is the head of the household, we can deduce that the spouse is the father of the sons/daughters of the head. Thus, we assign a poploc to the children that corresponds to the pernum of the spouse of the household (the father). Since the spouse is not constrained to have a pernum value of 1, we find that the father is the fifth in the family line and, as such, all the children of the head of the household are assigned a poploc equal to 5, pointing to the male spouse.

Table 2

Pernum	Relationship	Age	Sex	Poploc
1	Head	38	F	0
2	Son	20	M	5
3	Daughter	15	F	5
4	Son	8	M	5
5	Spouse	40	M	0

The next case dealt with is when the head of the household, or the spouse of the head, is living with a parent. In this case, we must look for members of the family with a relationship of Parent or Parent-in-Law and assign poploc's to their children (the head or spouse). In Table 3, both the head and spouse have parents in the house. Since the relationship variable describes relationships to the head of the household, the spouse's parents will be categorized as Parent-In-Law.

Pernum	Relationship	Age	Sex	Poploc
1	Head	30	M	3
2	Spouse	31	F	5
3	Parent	63	M	0
4	Parent	64	F	0
5	Parent-In-Law	59	M	0

The male parent with pernum equal to 3 must be the father of the head so we assign the head of the household a poploc equal to 3. Note that while poploc is sex-specific, we could create momloc which would locate the mother by simply looking for female parents and in-laws (i.e. pernum 4). Similarly, the male parent-in-law is the father of the spouse and she is assigned a poploc equal to 5.

This program does not generate pointer variables for grandchildren in multi-generational families since it is not possible to determine which son of the head fathered the grand-child. In any case, for these multi-generational families, most poplocs that would be generated would be for kids and/or parents that are too young to be considered co-residing (children younger than 25 years old and parents under 60). Note also that if DHS data were collected in a group that practices polygamy, since there are multiple spouses, it would not be possible to determine which spouse is related to which in-law. The code does not adjust for this possibility and matches each spouse to an in-law.

Table A1

Datasets Used from Demographic and Health Surveys (DHS)

No.	Country	Year	HR (1)	HR (2)	IR	MR	No.	Country	Year	HR (1)	HR (2)	IR	MR
1	Albania	2008	7,999	31,085	7,584	3,013	37	Lesotho	2009	9,391	44,343	7,624	3,317
2	Armenia	2010	6,700	23,625	5,922	1,584	38	Morocco	2003	11,513	63,202	16,798	NA
3	Angola	2010	8,030	40,435	8,589	NA	39	Moldova	2005	11,095	32,110	7,440	2,508
4	Azerbaijan	2006	7,180	30,614	8,444	2,558	40	Madagascar	2011	8,094	40,216	17,375	8,586
5	Bangladesh	2011	17,141	83,307	17,842	3,997	41	Mali	2006	12,998	72,121	14,583	4,207
6	Burkina Faso	2010	14,424	79,769	17,087	7,307	42	Mauritania	2000	6,149	35,169	7,728	2,191
7	Benin	2006	17,511	88,545	17,794	5,321	43	Maldives	2009	6,443	40,904	7,131	1,727
8	Bolivia	2008	19,564	77,688	16,939	6,054	44	Malawi	2010	24,825	118,603	23,020	7,175
9	Brazil	1996	13,283	56,669	12,612	2,949	45	Mozambique	2011	13,919	62,374	13,745	4,035
10	Burundi	2010	8,596	42,320	9,389	4,280	46	Nicaragua	2001	11,328	60,822	13,634	2,912
11	Congo, Dem. Rep.	2007	8,886	47,677	9,995	4,757	47	Nigeria	2010	5,895	29,988	33,385	15,486
12	Central African Republic	1994	5,551	26,960	5,884	1,729	48	Niger	2006	7,660	45,856	9,223	3,549
13	Congo, Rep.	2009	7,096	29,648	7,051	3,146	49	Namibia	2006	9,200	41,933	9,804	3,915
14	Cote d'Ivoire	2011	4,368	23,862	3,040	886	50	Nepal	2011	10,826	49,559	12,674	4,121
15	Cameroon	2011	14,214	70,588	15,426	7,191	51	Peru	2003	46,073	187,037	28,951	2,487
16	Colombia	2009	51,447	204,091	53,521	NA	52	Philippines	2008	12,469	60,705	13,633	4,766
17	Dominican Republic	2007	32,431	123,652	27,195	27,975	53	Pakistan	1990	7,193	50,076	6,611	1,354
18	Egypt, Arab Rep.	2008	18,968	90,962	16,527	NA	54	Paraguay	1990	5,683	29,523	5,827	
19	Eritrea	2002	9,389	45,070	8,754	NA	55	Rwanda	2010	12,540	56,464	13,633	6,329
20	Ethiopia	2003	16,702	77,491	16,515	14,110	56	Sierra Leone	2008	7,284	41,306	7,374	3,280
21	Gabon	2012	9,755	40,762	8,422	5,654	57	Senegal	2010	7,902	64,921	15,688	4,929
22	Ghana	2008	11,778	46,192	4,916	4,568	58	Sao Tome and Principe	2008	3,536	13,414	2,615	2,296
23	Guinea	2005	6,282	36,907	7,954	3,174	59	Swaziland	2006	4,843	21,724	4,987	4,156
24	Guatemala	2098	5,587	30,655	6,021	NA	60	Chad	2004	5,369	28,756	6,085	1,887
25	Guyana	2009	5,632	22,812	4,996	3,522	61	Togo	1998	7,517	42,048	8,569	3,819
26	Honduras	2011	21,362	100,203	22,757	7,120	62	Timor-Leste	2009	11,463	67,383	13,137	4,076
27	Haiti	2012	13,181	59,451	14,287	9,493	63	Turkey	2003	10,836	47,478	8,576	1,971
28	India	2005	109,041	529,540	124,385	74,369	64	Tanzania	2011	10,040	52,924	10,139	2,527
29	Indonesia	2007	40,701	178,573	32,895	8,758	65	Ukraine	2007	13,379	34,120	6,841	3,178
30	Jordan	2009	13,577	72,541	10,109	NA	66	Uganda	2011	11,340	53,615	8,674	2,295
31	Kenya	2008	9,057	38,411	8,444	3,465	67	Uzbekistan	1996	3,703	19,148	4,415	NA
32	Cambodia	2010	15,667	76,728	18,754	8,239	68	Vietnam	2005	6,337	26,813	5,665	NA
33	Kazakhstan	1999	5,844	19,991	4,800	1,440	69	Yemen, Rep.	1991	12,836	83,889	6,010	NA
34	Comoros	1996	2,252	14,081	3,050	795	70	South Africa	1998	12,247	52,608	11,735	NA
35	Kyrgyz Republic	1997	3,672	16,908	3,848	NA	71	Zambia	2007	7,164	35,399	7,146	6,500
36	Liberia	2011	4,162	19,214	7,092	6,009	72	Zimbabwe	2010	9,756	41,793	9,171	7,480

Notes : The sample consists of the most recent DHS survey for each country. An "NA" designation refers to samples where individual recodes were not collected. The column headed HR(1) refers to the number of observations on the Household Roster and HR(2) refers to the total number of individuals in these households.The column headed "IR" is the roster of women ages 15-49, and "MR" is the roster of men ages 15-64.

Table A2

Datasets Used from International Public Use Microsamples (IPUMS)

No.	Country	Year	N	No.	Country	Year	N
1	Argentina	2010	3966245	38	Kyrgyz Republic	2009	564986
2	Armenia	2001	326560	39	Malawi	2008	1341977
3	Austria	2001	803471	40	Malaysia	2000	435300
4	Bangladesh	2011	7205720	41	Mali	1998	991330
5	Belarus	1999	990706	42	Mexico	2010	11938402
6	Bolivia	2001	827692	43	Mongolia	2000	243725
7	Brazil	2010	9693058	44	Morocco	2004	1482720
8	Burkina Faso	2006	1417824	45	Nepal	2001	2583245
9	Cambodia	2008	1340121	46	Netherlands	2001	189725
10	Cameroon	2005	1772359	47	Nicaragua	2005	515485
11	Canada	2001	801055	48	Pakistan	1998	13102024
12	Chile	2002	1513914	49	Panama	2010	341118
13	China	1990	11835947	50	Peru	2007	2745895
14	Colombia	2005	4006168	51	Philippines	2000	7417810
15	Costa Rica	2000	381500	52	Portugal	2001	517026
16	Cuba	2002	1118767	53	Puerto Rico	2005	35416
17	Ecuador	2010	1448233	54	Romania	2002	2137967
18	Egypt, Arab Rep.	2006	7282434	55	Rwanda	2002	843392
19	El Salvador	2007	574364	56	Senegal	2002	994562
20	Fiji	2007	84323	57	Sierra Leone	2004	494298
21	France	2006	19816137	58	Slovenia	2002	179632
22	Germany	1987	3160224	59	South Africa	2007	949105
23	Ghana	2000	1894133	60	South Sudan	2008	542765
24	Greece	2001	1028884	61	Spain	2001	2039274
25	Guinea	1996	729071	62	St. Lucia	1991	13382
26	Haiti	2003	599228	63	Sudan	2008	5066530
27	Hungary	2001	23603049	64	Switzerland	2000	364086
28	India	2004	1299674	65	Tanzania	2002	3732735
29	Indonesia	2010	23603049	66	Thailand	2000	604519
30	Iran, Islamic Rep.	2006	1299674	67	Turkey	2000	3444456
31	Iraq	1997	1944278	68	Uganda	2002	2497449
32	Ireland	2006	440314	69	United Kingdom	2001	1843525
33	Israel	1995	556365	70	United States	2010	3061692
34	Italy	2001	2990739	71	Uruguay	2006	256866
35	Jamaica	2001	205179	72	Venezuela, RB	2001	2306489
36	Jordan	2004	510646	73	Vietnam	2009	14177590
37	Kenya	2009	3841935	74	West Bank and Gaza	2007	227067

Notes : The results in the paper use the most recent IPUMS survey available for each country.

Table A3

Ancestors of the "Missing Girls" and their Characteristics

	(1) High Sex Ratio Groups	(2) All Other Groups	(3) Difference
Patrilocality	1.00	0.62	0.38***
Forms of Subsistence			
Agriculture	63.35	41.10	22.25***
Animal Husbandry	27.65	16.74	10.91*
Gathering	3.65	12.85	-9.20***
Hunting	3.64	17.24	-13.60***
Fishing	8.78	18.31	-9.53***
Politcal & Social Characteristics			
Frequently Attacked (1=yes)	0.72	0.35	0.37**
Size of Community	6.71	3.57	3.14***
Fixity of Residence	5.00	3.72	1.28***
Class or Caste System (1=yes)	1.00	0.35	0.65***
Main Crop Type			
Cereals (1=yes)	1.00	0.47	0.53***
Roots (1=yes)	0.00	0.21	-0.21***
Female Contribution to Subsistence			
Murdock Measure	28.75	34.82	-6.07
Number of Cultures	7	179	

Notes : The data are taken from the Standard Cross-Cultural Sample (Murdock and White 1969). The seven cultures in the column of those with high sex ratios are the Chinese, Lolos, Vietnamese, Koreans, Abkhaz, Armenians, and Punjabis. The size of communities is classified by categories ranging from 1 (fewer than 50) to 8 (cities of greater than 50,000). Fixity of settlement ranges from 1-5, with 1 representing fully nomadic tribes and 5 representing groups with fixed settlement.

Table A4

Relationship between Subsistence Methods and Patrilocality

	(1)	(2)	(3)
	Dependent Variable: Patrilocal (1=yes)		
	Pooled OLS	Continent Fixed Effects	Country Fixed Effects
Panel A: Intensive Agriculture and Patrilocality			
Intensity of Agriculture (index 0-3)	0.163***	0.0796***	0.0387**
	(0.01)	(0.01)	(0.02)
Observations	1,065	1,065	1,065
R Squared	0.13	0.30	0.50
Panel B: Hunting, Gathering, Fishing and Patrilocality			
Percent Reliance on Hunting, Gathering, Fishing (0-100%)	-0.700***	-0.473***	-0.227***
	(0.03)	(0.05)	(0.06)
Observations	1,266	1,266	1,266
R Squared	0.21	0.29	0.47
Panel C: Animal Husbandry and Patrilocality			
Percent Reliance on Animal Husbandry (0-100%)	0.835***	0.602***	0.416***
	(0.08)	(0.08)	(0.10)
Absence of Bovine Animals (1=yes)	-0.222***	-0.150***	-0.0830**
	(0.03)	(0.03)	(0.04)
Observations	1,158	1,158	1,158
R Squared	0.20	0.31	0.49

Notes: Sample is composed of 1,267 ethnic groups contained in the Ethnographic Atlas (Murdock 1965). The dependent variable in all regressions is a dummy variable for the ethnic group practicing patrilocality. The independent variables are listed in the left-most colulmn. Intensity of agriculture is defined as [0=none, 1=casual, 2=extensive, 3=intensive]. The other measures of subsistence vary between 0 and 1. An indicator is coded for the absence of bovine animals using V40 of the Ethnographic Atlas. Standard errors are reported in parentheses below the coefficients and are heteroskedastic-consistent. * significant at 10% ** significant at 5%. *** significant at 1%.

Table A5

Cross-Country Relationship between Coresidence Rates the Sex Ratios of Children Using Time Since the Neolithic Transition as an Instrument

	(1)	(2)	(3)	(4)
	Dependent Variable: Sex Ratio of Children		Dep Var: Coresidence Rate	Dep Var: Sex Ratio of Children
	OLS	Reduced Form	First Stage	2SLS
Share of Elderly Living With Son	12.4***			37.7***
	(4.16)			(9.31)
Log of Years since Neolithic Transition		2.87***	0.080***	
		(0.61)	(0.02)	
Observations	107	103	98	98
R Squared	0.131	0.245	0.202	

Notes : Each column represents a separate regression. The sex ratio of children and the coresidence rates are calculated from the most recent IPUMS/DHS survey for each country as described in Table 1. The instrument for the coresidence rate is the log years since the neolithic transition. Standard errors are reported in parentheses below the coefficients and are heteroskedastic-consistent. Data on the timing of the neolithic by country are taken from Putterman (2008). * significant at 10% ** significant at 5%. *** significant at 1%.

Table A6

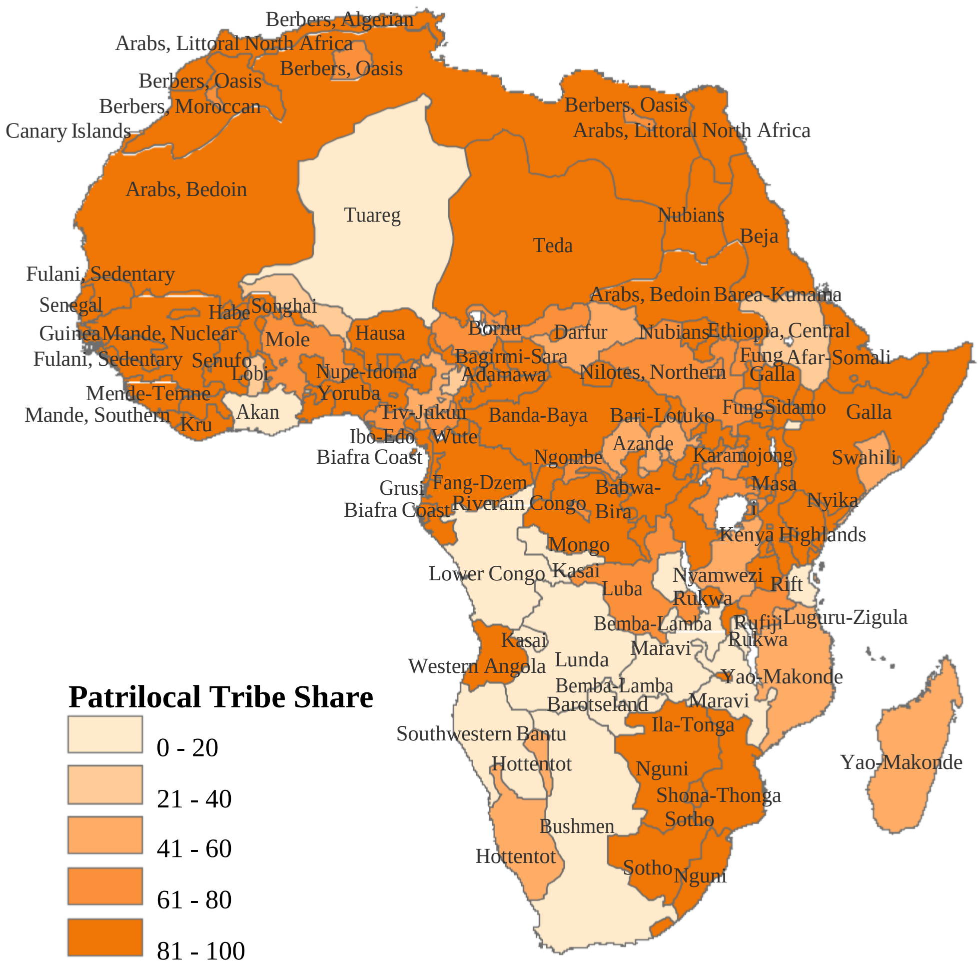
Sex Ratios of Next Birth by Sex of Existing Children in Select Countries

		Asia			Middle East & N. Africa			Sub-Saharan Africa			OECD		
Parity	Sex Combination	India	Armenia	Azer-baijan	Iran	Turkey	Palestine	Malawi	Ghana	Rwanda	USA	Germany	Israel
Overall	Any	108	113	112	109	105	104	99	99	101	104	105	104
1st	None	104	113	114	109	108	114	100	99	97	104	106	107
2nd	No Boys	113	105	123	105	105	109	102	89	98	100	104	103
	At least one boy	106	101	99	107	107	109	94	98	105	105	105	108
3rd	No Boys	120	193	142	109	105	108	103	100	103	109	109	109
	At least one boy	111	110	98	108	107	105	97	119	111	103	104	103

Notes: These calculations are for the most recent IPUMS/DHS survey for each country.

Figure A1

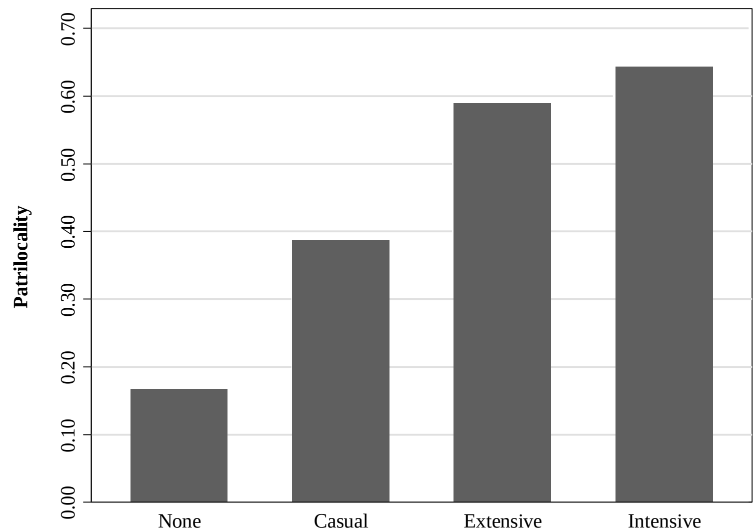
Residence Norms of Language Groups in Africa



Notes : Each polygon represnts the ancestral homeland of an ethnic group in Sudan, and the color shades refelct their residence customs, taken from Murdock (1959).

Figure A2

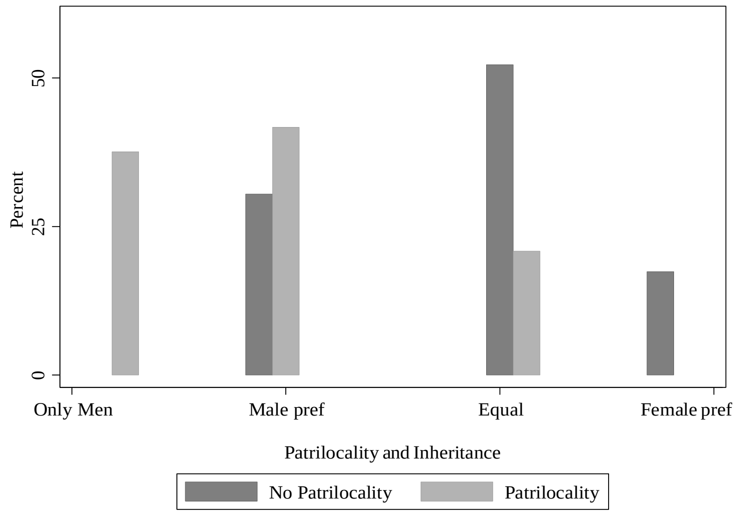
Coresidence Patterns and Intensity of Agriculture



Notes : The figure plots the relationship between patrilocal norms and subsistence methods among 1,267 traditional socieites recorded in the Ethnographic Atlas (Murdock 1965).

Figure A3

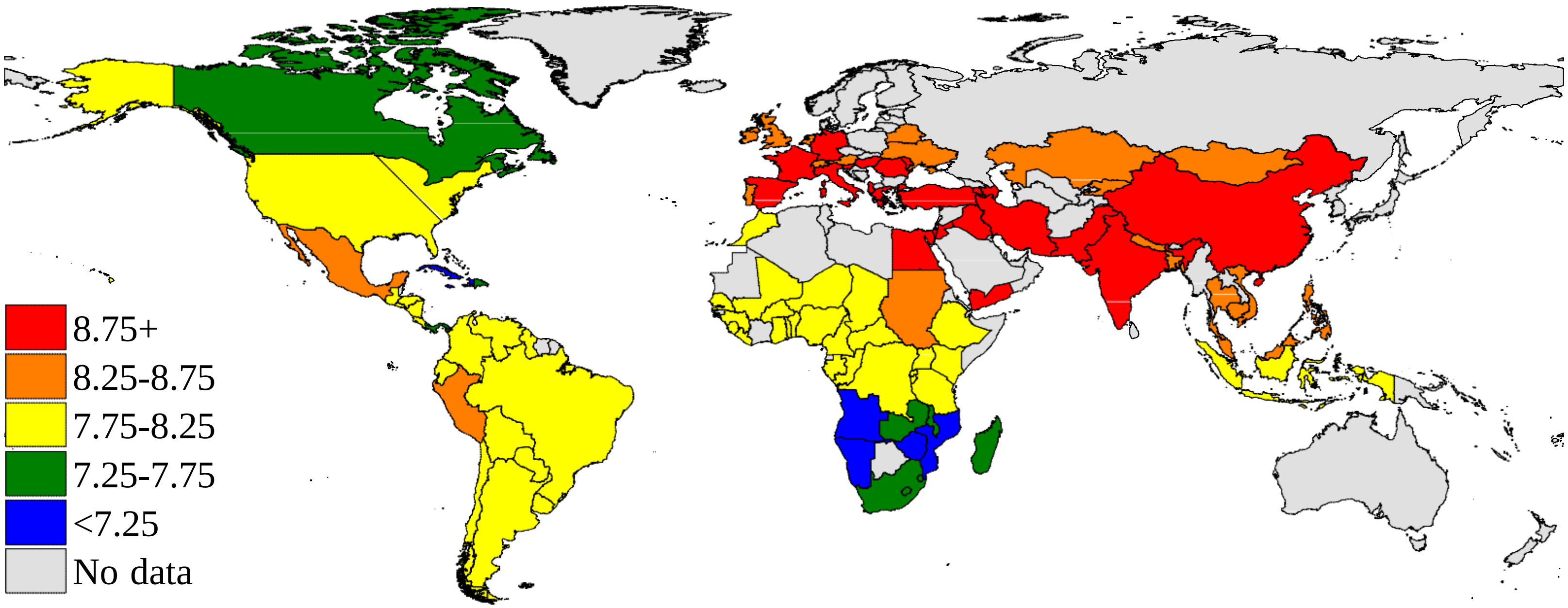
Patrilocality and Male Domination in Land Inheritance



Notes : The figure plots a histogram of the frequency of inheritance practices by whether residence after marriage was patrilocal.

Figure A4

Time Since Neolithic Transition in Log Years

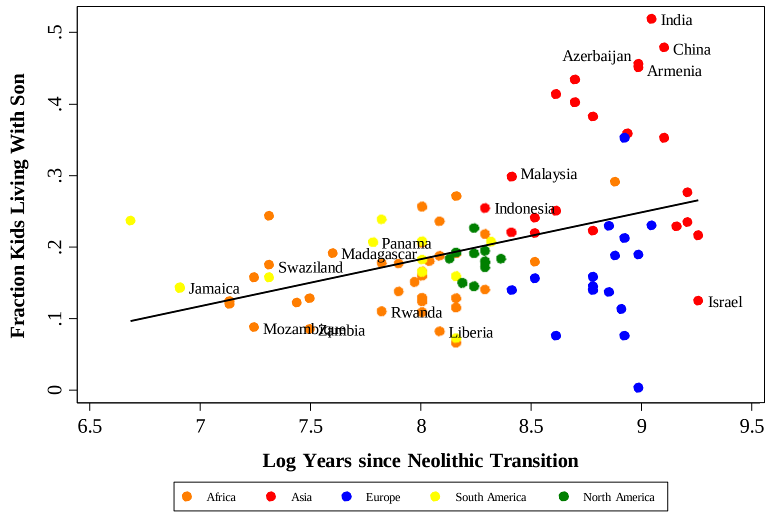


Note : Data are taken from Putterman (2008)

Figure A5

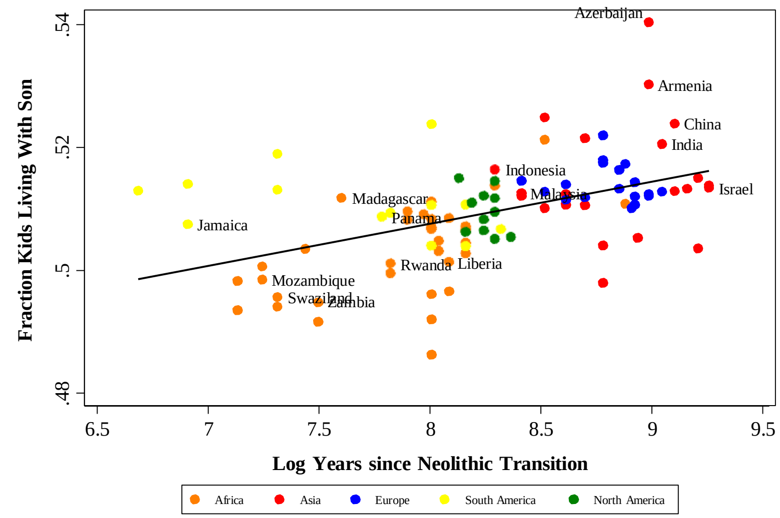
Coresidence Rates, the Male Fraction of Births, and Log Years since the Neolithic Transition across Countries

Panel A: Coresidence Rates and the Neolithic Transition



λ=.07, t=3.8, p-value<.001, N=107

Panel B: Male Fraction of Births and the Neolithic Transition



λ=.01, t=4.46, p-value<.001, N=107

Notes: Panel A plots the relationship between the fraction of men age 60+ living with an adult son (age>=25) and the log years since the neolithic transition. Panel B plots the relationship between the male fraction of children, ages 0 to 5, and the log years since the neolithic transition. The coresidence patterns and male fraction of children are taken from the most recent IPUMS and DHS data. Data on the timing of the neolithic by country are taken from Ashraf and Galor (2011).