A New Framework for Measuring Heterogeneity in Childbearing Strategies When Parents Want Sons and Daughters

Johannes Norling*

October 10, 2015

Abstract

In several countries, girls are more likely than boys to be aborted, to die in infancy, or to have younger siblings, all of which signal that parents want sons. However, standard techniques for measuring sex preferences fail to detect more subtle forms of sex preferences, especially when preferences are heterogeneous within a population. This paper introduces a new framework for estimating heterogeneity in population-level sex preferences using birth history records. This framework selects among many possible combinations of preferences over the sex and number of children to best match an observed population. Empirical estimates indicate that sex preferences are more widespread than previously reported and exhibit substantial heterogeneity within regions. In Africa, this heterogeneity is associated with agricultural traditions that favor men or women.

---

* Contact information: Department of Economics, University of Michigan, 611 Tappan Avenue, Ann Arbor, Michigan 48109; norling@umich.edu. Thank you to Barbara Anderson, Martha Bailey, Jacob Bastian, Hoyt Bleakley, Charlie Brown, Eric Chyn, Aaron Flaaen, Jeremy Fox, Andrew Goodman-Bacon, Morgan Henderson, Evan Herrnstadt, Max Kapustin, Justin Ladner, David Lam, Jeff Smith, Bryan Stuart, James Wang, and seminar participants at the University of Michigan and University of Cape Town for helpful comments and suggestions. This work was supported by the NICHD (T32 HD007339) as part of the University of Michigan Population Studies Center training program. I gratefully acknowledge the use of the services and facilities of the Population Studies Center (funded by NICHD Center Grant R24 HD041028). The opinions and conclusions expressed herein are solely mine and do not represent the opinions or policy of these funders or any agency of the federal government.

1. Introduction

Amartya Sen famously estimated that "more than 100 million women are missing" (Sen 1990). Due to abortion of girls, infanticide, and neglect, many of these women go missing at very young ages (Bongaarts and Guilmoto 2015). In addition to this direct selection against daughters, parents that want a son can simply continue bearing children until they have a son. Doing so concentrates girls in larger families in which resources are spread more thinly across many children (Jensen 2002). In China, India, and several other countries in Asia, son preference leaves girls outnumbered and disadvantaged from a young age (Guilmoto 2012, World Bank 2015).

Widespread son preference is less common in countries outside of Asia. Parents in the United States and several other countries tend to want a balance of sons and daughters, but in many countries in Africa and Latin America parents do not overwhelmingly want sons, daughters, or a balance of the two (Ben-Porath and Welch 1976, Arnold 1997, Angrist and Evans 1998, Bongaarts 2013). Standard techniques for measuring sex preferences cannot identify whether even shares of parents in these countries prefer sons and daughters or parents generally do not care about the sex of their children (Haughton and Haughton 1998). Such a distinction is crucial for understanding whether gender, which is fundamental to many areas of economic and social activity around the world (Alesina et al. 2013), also motivates childbearing decisions.

This paper introduces a revealed preference framework that, for the first time, measures heterogeneity in sex preferences in a population using a collection of birth histories. This new framework has three steps. First, I develop a model of childbearing in which parents care about the share of children that are boys and the total number of children. This model yields a set of possible childbearing strategies that govern the decision to have another child after each possible sequence of sons and daughters. Second, for a couple following each strategy, I calculate the likelihood that the couple stops childbearing after every possible sequence of sons and daughters. Third, I identify the combinations of strategies that best match an observed distribution of sequences of sons and daughters in completed families.

The new framework quantifies the range of preferences over sex and number of children that are consistent with an observed population. Crucially, this framework also measures the importance of the sex of children relative to the number of children. This relative importance determines how a couple weighs potentially competing objectives over the sex and number of their children. For example, a couple may want one child and prefer only sons, but whether the couple has a second child after a first-born daughter depends on the relative importance of having a small family versus having a son. If this couple would always stop after one child no matter its sex, then the couple's preference for sons has no bearing on its fertility decisions. Sex preferences matter only when the sex of previous children influences the decision to have another child.

Estimates using the new framework suggest that sex preferences are more widespread and heterogeneous than previously believed. Using several large-scale fertility surveys from Africa, Asia, and the Americas, this new framework estimates that, for at least 40 percent of couples in each of these regions, the sex of previous children influences the decision to have additional children. In Asia, son preference clearly dominates and at least half of parents want sons. In Africa and the Americas, many parents want sons and many want daughters.

Inheritance rules and other cultural characteristics contribute to widespread son preference in Asia (Das Gupta et al. 2003). Greater heterogeneity in sex preferences in Africa and the Americas suggests variation within these regions in the conditions that shape sex preferences. Africa comprises hundreds of ethnic groups that have a variety of male or female-favoring traditional practices and cultural norms (Murdock 1959, Murdock 1967). Estimates using the new framework identify pockets of son preference among ethnic groups in which land is traditionally inherited through the father's line, agriculture is traditionally performed primarily by men, and the plow is traditionally used in agriculture. Daughter preference is more common where inheritance of land follows matrilineal ties or women complete most agricultural tasks.

Standard approaches for measuring sex preferences during childbearing model sex preferences according to stopping rules, in which couples have a minimum number of children and then continue bearing children until they have a certain number of sons or daughters or reach a maximum number of total children (Seidl 1995). These rules permit sex preferences to only inflate fertility: without sex preferences, each couple would not exceed their minimum desired number of total children. For this reason, weakened sex preferences are thought to yield declines in overall fertility levels (Mutharayappa et al. 1997). The new framework in this paper more flexibly allows sex preferences to either inflate or deflate fertility levels: just as some couples may have additional children in order to have a desired son or daughter, other couples may stop childbearing early if they reach a particularly desirable combination of sons and daughters.

These two effects on fertility roughly offset in aggregate, and I estimate that eliminating all sex preferences in observed populations would change overall fertility levels by less than 0.2 children per couple. While sex preferences are widespread and govern many parents' childbearing decisions, these estimates suggest that they do not drive overall fertility levels.

This paper follows standard convention and refers to desire for sons or daughters as sex preferences. These preferences may result from a more fundamental optimization among the various tastes, incentives, and constraints that shape whether parents want sons or daughters. While I show that agricultural traditions in Africa are associated with son and daughter preference, I do not attempt to fully disentangle all components of sex preferences. Additionally, men and women may have substantially different sex preferences (Robitaille 2013, Ashraf et al. 2014), but large-scale birth history surveys are generally collected from women alone. This paper does not address inter-partner bargaining over childbearing decisions and refers to preferences as belonging to couples. Finally, it is not possible to recover an individual couple's preferences from their sequence of sons and daughters: when the sex of children is random, couples with the same preferences can have different sequences of sons and daughters, while couples with different preferences can have the same sequence (Haughton and Haughton 1998). Only the distribution of preferences across a group of couples is discernible.

2. Sex preferences and parity progression

Substantial research explores factors that influence the likelihood that a conceived child is male or female (Novitski and Sandler 1956, James 1971, Pickles et al. 1982, James 1990). This likelihood can vary by ancestry and environmental conditions, but there remains no widely agreed-upon and adopted method by which parents can influence the sex of a fetus, and in all populations the natural likelihood that each child is a boy is about 0.51 (Bongaarts 2013). However, amniocentesis and ultrasound technologies allow parents to identify the sex of a fetus and make sex-selective abortion possible. The share of births that are boys remained at or below 0.519 in all countries before 1980, but has since risen above 0.519 in Armenia, Azerbaijan, China, Georgia, India, Pakistan, South Korea, and Vietnam, suggesting substantial selective abortion of girls (World Bank 2015). Girls in many of these countries are also more likely than boys to die in childhood (Arnold 1997, Bongaarts and Guilmoto 2015). Because no country exhibits corresponding selection against boys, the study of sex preferences during childbearing is overwhelmingly the study of son preference.

Outside of Asia, the sex ratio at birth generally remains at the natural level. Particularly in Sub-Saharan Africa, abortion is heavily stigmatized and infanticide is historically less common than in Asia (Maharaj and Clelend 2006, Kumar et al. 2009). However, even where sex-selective abortion is rare, sex preferences may influence whether a couple continues childbearing after each son or daughter is born. Many surveys ask parents to report whether they want another child or are using contraception, and these decisions signal sex preferences when they vary according to the sex of previous children (Bongaarts 2013; Milazzo 2014). However, such surveys generally record these decisions over a limited time frame, such as since the birth of a parent's most recent child, and therefore do not fully address the sequential nature of childbearing for a cohort of parents. Many more surveys record the order and sex of all of a parent's live births. These birth history surveys offer the most comprehensive account of a parent's childbearing career. In this paper, I focus on inferring sex preferences using these sequences of sons and daughters.

Sex preferences shape the distribution of children across families. For example, if every birth is a boy with likelihood one-half and all couples have up to two children in order to have a son, then half of parents have a first-born son and stop, one-quarter have a daughter and then a son, and one-quarter have two daughters. Several characteristics of this distribution are commonly cited as signals that couples have sex preferences (Park 1983, Yamaguchi 1989, Seidl 1995, Clark 2000, Jensen 2002, Conley and Glauber 2006, Basu and de Jong 2010, Chaudhuri 2012). As given in column 1 of Table 1, no boy has a younger sibling while two-thirds of girls have a younger sibling. Boys have one-third of a sibling on average while every girl has one sibling. The overall sex ratio remains one boy for every girl, but three-quarters of last-born children are boys and the average share of sons per family is five-eighths. Parity progression ratios, which measure the share of parents with a particular combination of children that have another child, are greater for couples with one daughter than for couples with one son, suggesting that couples are generally not satisfied with having a daughter.

Of these statistics, parity progression ratios best measure sex preferences when not all couples have the same preference. In population 3 in Table 1, 60 percent of couples have up to two children in order to have a son, and the rest have up to three children in order to have a daughter. Boys and girls in this population are equally likely to have a younger sibling and the share of last-born children that are boys is one-half, both of which fail to indicate son or daughter preference. Boys have more siblings on average than do girls, suggesting that parents generally want daughters, but the average share of sons per family is greater than one-half, suggesting that parents generally want sons. Parity progression ratios better explain the mixture of preferences in the population: parents with one daughter are more likely than parents with one son to have a second child, while parents with two sons are most likely to have a third child, suggesting that son preference dominates among parents deciding to have a second child and daughter preference dominates among parents deciding to have a third child.

In actual populations, parity progression ratios exhibit substantial variation. For the main empirical analyses in this paper, I compile birth histories collected by the following seven large surveys: China's 1988 Fertility Survey (the "Two-per-thousand" survey); Demographic and Health Surveys; India's 1982 Rural Economic and Demographic Survey; Japanese General Social Surveys; United States Integrated Fertility Survey Series; Multiple Indicator Cluster Surveys; and World Fertility Surveys. As depicted in Figure 1 and listed in Appendix Table 1, these surveys offer substantial coverage in Africa, Asia, North America, and South America. Table 2 provides observed parity progression ratios by continent. On each continent, these exhibit variation among couples at the same parity. For example, 72.2 percent of couples in North America with two sons have a third child, which is 0.9 percentage points higher than couples with two daughters, 2.3 percentage points higher than couples with a daughter and then a son, and 2.8 percentage points higher than couples with a son and then a daughter.

The goal of this paper is to explain the childbearing behavior that drives this variation in observed parity progression ratios within populations. First, to gauge whether sampling error alone can explain this variation, I use a binomial test of the null hypothesis that all couples in a population with the same number of children are equally likely to have another child. For example, Appendix Table 2 demonstrates that, if all 1.5 million women observed in Asia were equally likely to have a second child, the likelihood is less than 0.001 that at most 78.6 percent of women with one son would have a second child and at least 79.9 percent of women with one daughter would have a second child. Sampling error alone similarly cannot explain observed variation in parity progression ratios in Asia after two children and after three children. In Africa and the Americas, sampling error could explain variation in parity progression after the first child but not after the second and third child. Appendix Figure 1 similarly indicates that, in most of the 94 countries in the compiled dataset, sampling error alone cannot account for all observed variation in parity progression ratios.

Because sampling error alone cannot explain all of the observed variation in parity progression ratios, at least some parents on each continent must make childbearing decisions that depend on the sex of previous children. As given in Table 2, at all parities, couples in Asia with only daughters are more likely to have another child than are couples with only sons. However, not all couples stop after having a son. For example, one-third of couples in Asia with three sons have another child. Similarly, in each of the other three continents, parity progression is greatest for couples with only sons or only daughters, but substantial shares of couples that already have a son and daughter go on to have another child.

This observed variation in parity progression ratios suggests that sex preferences in actual populations are not as homogeneous as in the simulated populations in Table 1. Table 2 does not clearly signal whether a minority of parents have strong sex preferences or a majority of parents have relatively weaker sex preferences. Especially outside of Asia, couples may prefer sons, daughters, a balance of sons and daughters, or they may not care about the sex of children. Some couples may want many children while others want just one or two. As noted by Ben-Porath and Welch (1976, page 292), "If populations are heterogeneous … observed patterns will be blurred." In the next section, I introduce a new approach for measuring this heterogeneity. Rather than calculate a single statistic from the distribution of sequences of children across families, the new approach specifies a set of possible preferences that parents can have and then identifies the combinations of preferences that best explain the entire distribution of sequences of children in completed families. While point identification of sex preferences is generally not possible, this new technique identifies meaningfully narrow bounds on preferences. These estimates indicate that, in Africa and the Americas, many parents prefer sons while many others prefer daughters.

3. A new framework for measuring heterogeneity in sex preferences

In this section, I introduce a three-step framework for measuring heterogeneity in sex preferences during childbearing. First, I define a set of possible childbearing strategies. Second, I calculate the likelihood that a couple following each strategy has each possible sequence of sons and daughters. Third, I identify the combinations of strategies in a population that best explain an observed distribution of sequences of sons and daughters.

A childbearing strategy is a rule governing the decision to have another child after every possible realization of sons and daughters. For example, a couple has one child and then has a second child only if the first-born child is a daughter. This strategy fits within a stopping rule model of childbearing in which a couple has a minimum number of children and then continues bearing children until it reaches either a target number of sons and daughters or a maximum number of total children (Sheps 1963; Keyfitz 1977; Seidl 1995; Jensen 2002; Basu and de Jong 2010). Stopping rules only permit sex preferences to inflate fertility above the minimum, which is the number of children the couple would have absent any son or daughter targets (McClelland 1979).

This section introduces an alternative model of childbearing in which preferences over the ideal share of children that are boys and ideal number of children form childbearing strategies. This model highlights the tradeoff between sex and number of children. While trying to improve upon an undesirable combination of sons and daughters, couples may have more than their ideal number of children. Alternatively, couples with a particularly desirable combination of sons and daughters may stop childbearing before they reach their ideal number of children. This model therefore more flexibly permits sex preferences to inflate or deflate fertility. Appendix Table 3 demonstrates that this model nests common stopping rules.

This new framework assumes that the likelihood that each birth is a boy is stochastic and the same across all couples. This assumption is clearly violated where sex-selective abortion is prevalent, and I exclude from the compiled dataset births to women who were of childbearing age after 1980 and live in a country in which the share of births that are boys has ever exceeded 0.519. Couples may also vary in their natural likelihoods of having a boy, but this variation is small (James 2009). For the main results, I assume that each child is a boy with likelihood 0.51, and I demonstrate the robustness of the empirical estimates to changes in this assumption.

3.1. Define a set of possible childbearing strategies

I assume that all couples have the same number of childbearing periods, T, and face the same likelihood that each child is a boy, l. In each period, a couple tries or does not try to have a child and then has a son, a daughter, or no child (there are no twin births). All couples have the same likelihood of conception when trying to get pregnant, p, and when not trying, q. Therefore, a couple that tries to get pregnant has a son with likelihood pl, a daughter with likelihood $p(1–l)$, and no child with likelihood $(1–p)$. A couple that does not try to get pregnant has a son with likelihood ql, a daughter with likelihood $q(1–l)$, and no child with likelihood $(1–q)$. Later I discuss sensitivity of the main empirical estimates to specifications of these parameters.

After the last childbearing period, the couple receives additively-separable utility from the share of children that are sons and the total number of children. The order in which sons and daughters are reached does not matter, just their total number. The couple has three preferences: ideal share of children that are sons, $r^*\in[0,1]$, ideal total number of children, $c^*\in[0,\mathrm{T}]$, and importance of the sex of children relative to the number of children, $\alpha\in[0,1]$. I consider a discrete subset of the infinite number of combinations of these preferences: $r^*\in\{0, 0.1, 0.2, \dots, 1\}$, $c^*\in\{0, 1, 2, \dots, T\}$, and $\alpha\in\{0, 0.1, 0.2, \dots, 1\}$.

A couple may not be able to simultaneously reach its ideal share of children that are boys and ideal number of children. The relative importance of the sex of children determines how the couple trades off these competing goals and captures the degree to which sex preferences affect childbearing decisions. A couple may consider only the number of children (α=0), only the sex of their children (α=1), or both (0<α<1) when deciding whether to have additional children. Sequential childbearing decisions are therefore fundamentally economic, balancing competing desires (the sex and number of children) under constraints (a finite number of childbearing periods and the randomness of the sex of each child).

A couple's childbearing decisions are governed by the following Bellman equation:

$$\mathrm{V}(s,d,T-t) = \mathrm{Max}\{EV_{\mathrm{try}}, EV_{\neg\mathrm{try}}\}, \text{ where} \quad (1)$$

$$EV_{\mathrm{try}} = plV(s+1,d,t+1) + p(1-l)V(s,d+1,t+1) + (1-p)V(s,d,t+1)$$

$$EV_{\neg\mathrm{try}} = qlV(s+1,d,t+1) + q(1-l)V(s,d+1,t+1) + (1-q)V(s,d,t+1)$$

$$V(s,d,T+1) = \begin{cases} -\alpha\left(\frac{s}{s+d} - r^*\right)^2 - (1-\alpha)(s+d-c^*)^2 & \text{if } s+d>0 \\ -\alpha \qquad\qquad\qquad - (1-\alpha){c^*}^2 & \text{if } s=d=0. \end{cases}$$

In childbearing period t, the couple tries or does not try to have a child in order to maximize expected utility at the end of its childbearing career (period T+1). (These two options – trying to become pregnant or not trying to become pregnant – approximate actual childbearing decisions that may involve a range of effort at becoming or avoiding becoming pregnant.) If trying and not trying to have a child yield the same expected utility, I assume the couple tries to have a child. The decision in each period depends on the number of sons, s, and daughters, d, to which the couple has already given birth. The couple receives bliss-point utility in period T+1 equal to the sum of the squared difference between the actual and ideal share of children that are boys and the squared difference between the actual and ideal total number of children. The importance of the sex of children relative to the number of children, α, weights these two terms.

3.2. Calculate the likelihood that a couple following each strategy has each sequence of children

Equation 1 is solved using backward induction. In period T, a couple tries to have a child if and only if the expected utility from trying to having a child weakly exceeds the expected utility from not trying to have a child. When facing the same choices in period T–1 the couple knows, for each possible outcome, what its decision will be in period T. Similar calculations govern decisions in all earlier periods. The set of decisions given all possible realizations of sons and daughters constitutes a childbearing strategy.

For example, assume that a couple faces two childbearing periods and an even chance that each child is a boy (T=2, l=½), has perfect control over conception (p=1, q=0), and cares only that all of their children are sons (α=1, r^*=1, c^* does not matter and can take any value). In the first period, this couple has a child. If this first-born child is a son, additional children could not raise and might reduce the share of children that are sons, so this couple does not have a second child. If the first-born child is a daughter, a second child cannot reduce and might raise the share of children that are sons, so this couple has a second child. Therefore, with likelihood l the couple has a son, with likelihood $(1–l)l$ the couple has a daughter and then a son, and with likelihood $(1–l)(1–l)$ the couple has two daughters. The couple has no chance of having any other sequence of children.

Imperfect control over conception is the only reason to defer childbearing. For example, assume that a couple faces two childbearing periods and an even chance that each child is a boy (T=2, l=½), can get pregnant while trying but faces some chance of accidentally becoming pregnant when not trying (p=1, 0<q<1), and wants one child regardless of its sex (α=0, c*=1, r* does not matter). If the couple tries to have a child in the first period, the couple succeeds but runs a risk of accidentally having a child in the second period, and with likelihood q the couple has two children. If the couple instead waits to try to have a child, with likelihood q the couple has a child in the first period, and then again with likelihood q the couple has a child in the second period, yielding a likelihood q^2 that the couple has two children. Because $q^2<q$, waiting to try to conceive lessens the risk of having two children. Knowing that it can always have a child if it wishes to, the couple defers trying to have a child until the second period.

Parents defer childbearing for a host of education, employment, and marriage reasons unrelated to the likelihood of conception. The age at which a woman begins making the decisions captured by model 1 may be later than the age at which she first becomes fecund, and the gap between these ages can vary across women. Assigning a single age at which every woman's childbearing career begins and dividing the remaining fecund years into T evenly-spaced childbearing periods, as in Wolpin (1984), therefore potentially invites substantial distortions: couples may defer childbearing after the first period for reasons unrelated to the likelihood of conception, but model 1 interprets this deferral as resulting from concern over the likelihood of conception. Women who begin childbearing at age 16, or 20, or 24 may in fact have very similar preferences over the number and sex of children.

To avoid misinterpreting the timing of births, I collapse the predicted distributions over sequences of sons and daughters as given in Appendix Table 4, ignoring the timing of each birth. I therefore ignore information about the spacing between births. Shorter spacing after the birth of daughters in some countries signals that parents prefer and are eager to have sons. In part because girls are breastfed for less time than boys, these shorter intervals can substantially reduce the chance that girls survive infancy (Jayachandran and Kuziemko 2011). By collapsing to just the sequence of children, I avoid misinterpreting the timing of childbearing, but at a cost of ignoring potentially meaningful information about the timing of births.

3.3. Identify the combinations of strategies that best explain an observed population

Given T childbearing periods, there are N=121(T+1) possible strategies (defined by each unique combination of preferences r*, c*, and α) and $M=2^{T+1}-1$ possible unique sequences of sons and daughters. **A** is an $M\times N$ matrix in which each element a_{mn} is the likelihood that a couple with combination of preferences *n* has order of children *m*. **D** is an *M*×1 matrix in which each element d_m is the share of couples with sequence of children *m* in an observed population. For example, consider a large population in which half of couples have a single son, one-quarter have a daughter and then a son, and one-quarter have two daughters. Assuming two childbearing periods, an even chance that each child is a son, and perfect control over conception (T=2, l=½, p=1, q=0), Table 3 defines **A** and **D**.

The goal is to determine the share of couples with each combination of preferences that best explains an observed distribution of sequences of children. If all couples follow the final strategy listed in **A** (r*=1, c*=2, α=1), then half of couples will have one son, one-quarter will have a daughter and then a son, and one-quarter will have two daughters. This distribution exactly matches the actual observed distribution of sequences of children, **D**. However, if all couples follow the next-to-last strategy listed in **A** (r*=1, c*=2, α=0.9), the resulting population will also exactly match the observed population. Several other strategies also exactly match the observed population, as does any large population composed of combinations of couples following any of these strategies.

Because there are an infinite number of such combinations of strategies that best match the observed population, the empirical challenge is to describe these combinations. In this section, I present an approach for estimating bounds on the average values of the three preferences (r*, c*, and α), and later I adapt this approach to estimate alternative summary measures. I use a two-stage procedure for calculating bounds on average values of preferences. The first stage identifies the smallest difference between the observed and predicted populations:

$$\begin{aligned} \underset{\mathbf{S}_{N\times 1}}{\text{Min}} \quad & \mathbf{W}_{1\times M}^{\mathrm{T}} |\mathbf{D}_{M\times 1} - \mathbf{A}_{M\times N}\mathbf{S}_{N\times 1}| \\ \text{s.t.} \quad & \mathbf{0}_{N\times 1} \leq \mathbf{S}_{N\times 1} \\ & \mathbf{1}_{1\times N}\mathbf{S}_{N\times 1} = 1. \end{aligned} \tag{2}$$

The candidate share, **S**, is an *N*×1 matrix in which each element s_n is the estimated share of couples with combination of preferences *n*. Minimum value function 2, hereafter referred to as equation 2, identifies a candidate share that minimizes the sum of absolute deviations between the observed population, **D**, and the predicted population, **AS**, subject to the constraints that the share of parents following each strategy are non-negative and sum to 1. I use the sum of absolute deviations (L_1) norm so that this minimized sum of absolute deviations can enter as a constraint into the second stage. Each deviation is weighted according to an M×1 matrix **W**. In all analyses in this paper I assign equal weight to all orders of children (**W**=**1**). Equation 2 identifies the smallest possible sum of absolute deviations between the observed and predicted populations, $e^{\min}=\mathbf{W}^{\mathrm{T}}|\mathbf{D}-\mathbf{AS}|$.

The second stage calculates the bounds on average values of each preference across all candidate shares that best match the observed population:

$$\begin{aligned} \underset{\mathbf{S}_{N\times 1}}{\text{Min}} \quad & \mathbf{F}^{\mathrm{T}}_{1\times N}\mathbf{S}_{N\times 1} \\ \text{s.t.} \quad & \mathbf{0}_{N\times 1} \leq \mathbf{S}_{N\times 1} \\ & \mathbf{1}_{1\times N}\mathbf{S}_{N\times 1} = 1 \\ & \mathbf{W}^{\mathrm{T}}_{1\times M}|\mathbf{D}_{M\times 1} - \mathbf{A}_{M\times N}\mathbf{S}_{N\times 1}| = e^{\min}. \end{aligned} \tag{3}$$

F is a N×1 matrix in which each element f_n is a characteristic of strategy n. For example, if each element of **F** is the ideal share of children that are boys in each strategy, then equation 3 chooses a candidate share that minimizes $\mathbf{F}^{\mathrm{T}}\mathbf{S}$, the average ideal share of children that are boys across all strategies. In addition to the same proportionality constraints as in equation 2, equation 3 requires that the sum of absolute deviations using this chosen candidate share must equal the minimized sum of absolute deviations from the first stage. I use the linear programming simplex algorithm to solve equations 2 and 3.

The minimized value of $\mathbf{F}^{\mathrm{T}}\mathbf{S}$ from equation 3 provides the minimum average ideal share of children that are sons across all combinations of strategies that most closely predict the observed population. Similarly, setting each element of **F** to equal the ideal number of children in each strategy or the relative importance of the sex of children permits calculations of the minimum average values of these preferences across all combinations of strategies that most closely predict the observed population. Finally, multiplying **F** by –1 and rerunning equation 3 yields the maximum average value of these preferences, $-\mathbf{F}^{\mathrm{T}}\mathbf{S}$.

Figure 2 provides the estimated bounds on the average values of the three preferences for each of the simulated populations in Table 1, calculated assuming three childbearing periods, a likelihood of one-half that each child is a boy, and perfect control over conception. Panel (a) indicates that the average ideal share of children that are boys is at least 0.6 in population 1, is at most 0.2 in population 2, and is between 0.1 and 0.75 in population 3. (Again, because of collinearity of strategies, it is possible to bound but not point identify preferences.) The distribution of sequences of children in completed families shape these estimated bounds. For example, half of couples in population 1 have just one son but no couples have just one daughter, suggesting that parents want most or all of their children to be sons. These estimates are consistent with parity progression ratios that suggest son preference in population 1, daughter preference in population 2, and no clear overall son or daughter preference in population 3.

Panel (b) of Figure 2 indicates that the average ideal number of children could be between 0 and 1 in population 1, between 0 and 3 in population 2, and between 0 and 1.8 in population 3. This range is lowest in population 1 because, while there are three childbearing periods, no couple in population 1 has more than two children, suggesting that at least some couples are not willing to have a third child in order to have a son. In contrast, some couples in population 2 do have three children.

Panel (c) of Figure 2 indicates that, on average, the relative importance that couples place on the sex of children is between 0.8 and 0.9 in population 1, between 0.9 and 1 in population 2, and between 0.5 and 0.95 in population 3. While it is not possible to identify whether parents in population 3 generally prefer sons or daughters, these estimates demonstrate that, on average, parents in population 3 do care about the sex of their children. Within each population, the range of possible average relative importance of the sex of children shapes the width of the estimates in panels (a) and (b). For example, if all parents in population 2 care only about the sex of their children, then their preferences over ideal number of total children do not matter and can range anywhere from zero to three. If all parents instead care about both the sex and number of children, then the bounds on average ideal number of children narrow.

4. Main results: Estimates of preferences using the new framework

I solve equation 1 using the following quantities: eight childbearing periods, a likelihood of 0.51 that each child is a boy, a likelihood of 0.9 that a couple conceives when trying to get pregnant, and a likelihood of 0.1 that a couple conceives when not trying to get pregnant (T=8, l=0.51, p=0.9, q=0.1). Larsen (2005) reports that approximately 90% of women in northern

Tanzania are able to get pregnant when trying within two years. However, as with control over conception when not trying to get pregnant, this likelihood depends on many factors, such as access to contraception and the likelihood of survival of the fetus (Henry 1961, Di Renzo et al. 2007, Clifton 2010). These likelihoods likely vary across populations, across women within a population, and even over time for individual women. Later in this section, I demonstrate the sensitivity of the empirical estimates to the specification of these parameters.

I limit the sample introduced in section 2 to birth histories collected from women that gave birth eight or fewer times, were age 40 or above when observed, and do not have any twin births. For countries in which the share of births that are boys has ever exceeded 0.519 since 1980 (Armenia, Azerbaijan, China, India, Pakistan, South Korea, and Vietnam), I exclude women that reached age 40 after 1980. I calculate the distribution of sequences of children in each continent using weights that accompany each survey, with each survey's weight rescaled to have a mean value of one.

4.1. Estimated bounds on average values of preferences

The dark rectangles in panel (a) of Figure 3 provide bounds on the identified set of values of the average ideal share of children that are boys that best explain observed childbearing, calculated according to the procedure in section 3. The light rectangles provide 95-percent confidence intervals around the unknown true average value, calculated using subsampling following Romano and Shaikh (2008) with 1,000 subsamples. The coefficient of determination is provided next to each continent's name and is calculated according to McKean and Sievers (1987). These estimates suggest that the average ideal share of children that are boys is between 0.25 and 0.9 in Africa, 0.52 and 0.97 in Asia, 0.2 and 0.8 in North America, and 0.3 and 0.8 in South America. The estimates in panel (b) suggest that average ideal number of children is roughly five in Africa, just above four in Asia, and less than four in the Americas. (Because younger women that live in high-sex ratio countries in Asia are excluded from these estimates, women in Asia appear generally older than in other continents. The actual ideal number of children across all women in Asia today is likely much less than four.) The estimates in panel (c) suggest that the average relative importance of the sex of children is between about 0.2 and 0.8 in Africa and the Americas, and between 0.4 and 0.85 in Asia.

Variation in parity progression ratios, and corresponding variation in the observed distribution of sequences of children in completed families, drives these estimates. As discussed in section 2, this variation indicates that the sex of previous children is associated with the choice to have additional children. For this reason, this new framework estimates that average relative importance of the sex of children is greater than zero. As given in Table 2, at each parity, couples in Asia are substantially more likely to have another child when all previous children are girls compared to any other composition of previous children. This comparison is particularly stark when couples have three children: couples in Asia with three daughters are 9 percentage points more likely to have another child than are couples with three sons (76 percent compared to 67 percent). Because no other continent exhibits correspondingly stark parity progression ratios in favor of one sex, only in Asia does the new framework conclude that parents on average favor a particular sex.

4.2 Robustness of estimates to model parameters and sample definition

Figure 4 demonstrates the sensitivity of estimated bounds on average values of preferences in Africa to model specification and sample definition. The solid rectangles provide the identified set under the main specification, and the hollow rectangles under alternative specifications. As given in Figure 4a, as the number of childbearing periods permitted rises and women with greater numbers of children are included in the sample, the average ideal number of children rises and the bounds on sex preferences narrow. As the assumed likelihood that each birth is a boy rises, the upper and lower bounds on average ideal share of children that are boys fall slightly, but the bounds on the average ideal number of children and average relative importance of the sex of children remain roughly the same. For various specifications of the increment between each possible r^* and each possible α, sex preferences also remain roughly consistent. In all cases, the specification used in the main results best explain observed childbearing (coefficient of determination equal to 0.94).

Figure 4b demonstrates that estimated bounds on average preferences vary substantially across different assumed specifications of control over conception. The bounds are widest when the likelihood of conception when trying to get pregnant is low (0.8 or 0.7) or when the likelihood of conception when not trying to get pregnant is low (0). Estimated bounds on average values of preferences are narrower and remain roughly stable using larger values of these parameters and, as with the parameters compared in Figure 4a, the likelihoods of conception assumed when estimating the main results (0.9 when trying to get pregnant and 0.1 when not trying to get pregnant) best explain observed childbearing in Africa (coefficient of determination equal to 0.94). Figure 4b therefore suggests that actual couples face a small likelihood of infecundity and a small likelihood of accidentally becoming pregnant. Imperfect control over conception introduces some randomness into the observed sequences of sons and daughters. A model that assumes perfect control over conception cannot account for this randomness and struggles to explain observed childbearing. By allowing for imperfect control over conception, the framework presented in this paper better explains variation in sequences of sons and daughters in completed families.

Figure 4c demonstrates that the estimated bounds on average preferences remain stable as the minimum age at observation rises from 40 up to 44. Estimates are also stable across currently and previously married women, but widen considerably among never married women. Finally, the main estimates include all live births, and estimates vary only slightly excluding children that died before 1 month of age, before their first birthday, before their fifth birthday, or before their next sibling was born. However, the coefficient of determination rises slightly as deceased children are excluded.

4.3. Additional summary measures of sex preferences

Section 3.3 presents a two-stage optimization procedure for estimating bounds on average values across all couples of their ideal share of children that are boys, ideal number of children, and relative importance of the sex of children. This section adapts this procedure to consider four additional summary measures: *Prevalence*, equal to the share of couples that place any importance on the sex of children (α>0); *WantBalance*, equal to the share of couples want an even ratio of sons to daughters (r*=½); *WantSons*, equal to the share of couples that want more than half of their children to be sons (r*>½); and *WantDaughters*, equal to the share of couples that want more than half of their children to be daughters (r*<½). These summary measures are calculated using alternative specifications of **F** in the procedure in section 3.3. For example, to estimate the prevalence of sex preferences, each of the N elements of **F** equals 1 if the corresponding strategy n places any weight on the sex of children (α>0) and 0 otherwise.

The binomial test in section 2 demonstrates that at least some parents in each continent make childbearing decisions that depend on the sex of previous children. As presented in panel (a) of Figure 5, the estimated prevalence of sex preferences is at least 0.4 on all continents. These estimates suggest that the sex of previous children influences future childbearing decisions for at least 40 percent of parents. On all four continents, sex preferences are widespread and not concentrated among a small group of parents.

Panel (c) of Figure 5 indicates that at least half of couples in Asia prefer sons to daughters, confirming the findings in earlier studies that son preference dominates in Asia. Elsewhere, sex preferences are heterogeneous. As given in panel (b), at most about 60 percent of couples in Africa and the Americas want a balance of sons and daughters. As given in panels (c) and (d), the identified bounds on the shares of couples that want sons or daughters are wide, but the lower bounds are substantially greater than zero: at least 15 percent of couples in these continents prefer sons, and at least 10 percent prefer daughters. Outside of Asia, at least one-quarter of parents prefer sons to daughters or vice versa.

4.4. Comparison between estimated and reported sex preferences

Preferences estimated using the new technique can be compared with attitudinal surveys that elicit stated preferences. Many Demographic and Health Surveys ask women to report the number of sons and daughters they wanted when they began childbearing. Figure 6 compares average reported preferences, given by dark circles, and bounds on estimated preferences, given by hollow rectangles. As given in panel (a) of Figure 6, roughly half of parents on all continents report that they wanted a balance of sons and daughters. Estimated bounds on this share (which differ from those in Figure 5 because the sample in Figure 6 consists only of Demographic and Heath Survey respondents that report ideal number of children) suggest that less than half of parents actually want a balance of sons and daughters. Parents appear to generally overreport wanting a balance of sons and daughters. Similarly, as given in panel (b), parents appear to underreport the strength of preferences. Bounds on the strength of sex preferences are again estimated using the new framework, with each of the N elements of **F** set equal to the magnitude of the difference between the ideal share of children that are boys and one-half, $|r^*–½|$, for the corresponding strategy n. The estimated lower bound on the strength of preferences is at least

0.2 on all continents, substantially greater than the average reported magnitude of about 0.125 on all continents.

This comparison supports Bongaarts' (2013) assertion that such sensitive questions are particularly prone to rationalization bias, with parents likely to report that they wanted the children they actually had. Additionally, Figure 6 depicts preferences reported by women alone, while observed distributions of sequences of children result from childbearing decisions that reflect both women's and men's preferences. However, Robitaille (2013) demonstrates that reported son preference in India is generally just as strong among women as among men. The underreporting of the strength of preferences suggested by panel (b) of Figure 6 may therefore not entirely be an artifact of using only women's reported preferences. Finally, these attitudinal surveys do not address the sequential nature of childbearing. A couple's decision to stop having children depends not only its ideal composition of sons and daughters but also on the relative undesirability of other combinations of sons and daughters. Widely-administered attitudinal surveys do not measure the relative importance of the sex of children, α.

5. Traditional agricultural practices and sex preferences in Africa

Previous studies find evidence of substantial son preference in Democratic Republic of the Congo, Egypt, Nigeria and a few other countries in Africa (Bongaarts 2013, Milazzo 2014). However, most countries in Africa do not exhibit clear son preference or clear daughter preference (Arnold 1997). Panel (a) of Figure 3 confirms the general conclusion of previous work that Africa as a whole does not exhibit clear son or daughter preference, but Figure 5 suggests that there is substantial heterogeneity in Africa: many parents choose to have another child depending on the sex of previous children, with some parents favoring sons and others favoring daughters.

In this section, I measure sex preferences by ethnic group affiliation. Africa comprises hundreds of ethnic groups. The anthropologist George Murdock drew from thousands of reports and other documents to generate a map of historical ethnic group boundaries in Africa and a database of pre-colonial characteristics of many of these groups (Murdock 1959, 1967). Economists have used this database to study the formation of institutions in Africa and their relationship with economic development (Fenske 2009, Bolt 2010, Nunn and Wantchekon 2011, Michalopoulos and Papaioannou 2014). Alesina et al. (2013) use this database to show that, consistent with the Boserup hypothesis that plow-based cultivation requires greater upper body strength than hoe-based cultivation and therefore gives men an advantage in agricultural production, areas in which the plow was traditionally used in agriculture tend to have more male-favoring gender norms today.

While Alesina et al. (2013) do not address childbearing, others have examined the relationship between traditional cultural characteristics and sex preferences in Asia. Das Gupta et al. (2003) find that son preference is widespread where inheritance passes through patrilineal ties and a bride's family pay a dowry at marriage. While dowry is rarely paid in Africa, there is substantial variety, even across ethnic groups located in the same country, in whether inheritance of land traditionally follows patrilineal or matrilineal ties and whether agriculture is traditionally performed primarily by men or women. For example, among the Ewe of Ghana, Togo, and Benin, patrilineal heirs inherit land. Among the Ashanti of Ghana, matrilineal heirs inherit land. In Kenya, men traditionally perform most agricultural tasks in Boran regions, while women do so in Kamba areas (Murdock 1959, 1967).

Many Demographic and Health Surveys that record birth histories also record the location of each cluster of respondents. Figure 7 provides the distribution of these clusters by ethnic group characteristic, as determined using Murdock (1959, 1967). The top row presents the location of respondents living in areas where inheritance of land traditionally passed through patrilineal ties, agriculture was traditionally performed primarily by men, and the plow was traditionally used in agriculture. These characteristics all imply incentives for having sons. The bottom row provides the location of respondents living in areas with traditional characteristics that imply incentives for having daughters. While traditional use of the plow is concentrated in North Africa and Ethiopia, the other two characteristics exhibit substantial variation, particularly in West, Central, and East Africa.

These traditional agricultural practices are associated with contemporary sex preferences during childbearing. Figure 8 presents estimated bounds on the share of couples that want more sons than daughters minus the share that want more daughters than sons, with couples grouped by agricultural tradition. Again, these estimates are calculated using the procedure in section 3.3, with each of the N elements of $\mathbf{F}$ equal to 0 if the corresponding strategy n represents balance preference ($r^*=½$), 1 if the strategy represents son preference ($r^*>½$), and –1 if the strategy represents daughter preference ($r^*<½$). Where inheritance of land traditionally followed patrilineal ties, agriculture was primarily performed by men, or the plow was traditionally used in agriculture, couples today generally favor sons. Where inheritance of land followed matrilineal ties, agriculture was primarily performed by women, or the plow was not used in agriculture, couples generally favor daughters.

Variation in the distribution of sequences of children in completed families drives these estimates. For example, among ethnic groups in which inheritance traditionally passes through patrilineal ties, 5.8 percent of couples stop after having two daughters, while 7.1 percent stop after having two sons. Among ethnic groups in which inheritance follows matrilineal ties, the shares are reversed, with 7.1 percent of couples stopping after two daughters but just 6.4 percent stopping after two sons. While the identified sets for each pair of estimates in Figure 8 overlap and it is not possible to statistically rule out that the direction of sex preferences are the same in each pair, these estimates suggest that agricultural traditions are associated with sex preferences today. Diversity in incentives for having sons and daughters crosses country borders and is consistent with variation in whether parents want sons or daughters.

6. Effect of sex preferences on fertility levels

High fertility has several adverse associations. Having many children raises the likelihood that a woman will die during childbirth, a risk that is exacerbated by high likelihoods of maternal death per childbirth in many countries that also have high fertility rates (Stanton et al. 2000). Parents with many children tend to invest less in each child’s education, leading to poorer labor market achievement for these children in adulthood (Lloyd and Brandon 1994, Pop-Eleches 2006). Rapid population growth unaccompanied by technological development can also place Malthusian pressure on food resources (Galor and Weil 2000, Hansen and Prescott 2002).

Mutharayappa et al. (1997), Bhat and Zavier (2003), and others propose that weakening sex preferences could foster declines in fertility by giving parents less incentive to have many children. This conclusion stems from the stopping rule model of sex preferences, in which parents exceed their desired minimum number of children only to have a target number of sons or daughters. By lowering or removing these targets, overall fertility should fall. Freedman and Coombs (1974) criticize this assumption on the grounds that it does not allow parents to stop early upon reaching a particular combination of sons and daughters.

In section 3.1 of this paper I introduce a more flexible bliss-point model of sex preferences in which parents may fall short of or exceed their ideal number of total children. The procedure in section 3.3 can be adapted to bound the expected change in children per couple as preferences change. For example, among couples following the strategy represented by the final column of **A** in Table 3 ($r^*=1$, $c^*=2$, $\alpha=1$), half have one child and half have two children. If these couples no longer cared about the sex of their children ($\alpha=0$), they would all have their ideal number of children ($c^*=2$). Therefore, half of these couples would have one additional child, and half would have the same number of children as before. Similar calculations for other strategies show that, if all couples no longer cared about the sex of their children, some would have additional children, some fewer, and some the same number. To calculate the expected average change in children per couple if all couples did not consider the sex of children when making childbearing decisions, each of the N elements of **F** is set equal to the expected change in fertility if a couple following strategy n retains the same ideal share of children that are boys and ideal number of total children (r^* and c^*) but cares only about their total number of children ($\alpha=0$).

Panels (a) and (b) of Figure 9 provide bounds on empirical estimates of the expected change in overall fertility under two separate counterfactual scenarios: all couples want a balance of sons and daughters ($r^*=½$), and all couples care only about their total number of children ($\alpha=0$). In each scenario, the average number of children per couple would change by at most 0.2. For example, if all couples in Africa want a balance of sons and daughters, then fertility will rise by at most 0.05 children per couple. While this small aggregate change comprises larger changes for individual couples, these estimates suggest that compelling parents to move toward balanced or weakened sex preferences would only slightly change overall fertility levels.

Panel (c) of Figure 9 provides estimated bounds on the change in children per couple if all couples want one fewer child than before ($c^*–1$). On all continents, fertility would fall by between 0.6 and 0.75 children per couple. The magnitude of this decline is less than one because some couples already want zero children ($c^*=0$). Imperfect control over conception also tempers this decline: as parents want fewer children, they spend more childbearing periods trying not to get pregnant, but during these periods they still run a risk of conception. The relative importance of the sex of children also tempers this decline: among parents that care mostly or entirely about the sex of their children, a desire for one fewer child will have little effect on childbearing decisions. However, this decline is still substantial. Together, the three panels of Figure 9 suggest that sex preferences alone do not shape overall fertility levels. Policies aimed at reducing each couple's ideal total number of children would more effectively lower aggregate fertility.

7. Conclusion

Son preference is widespread in many countries in Asia. While parents in much of the rest of the world do not overwhelmingly want sons or daughters, parity progression ratios and the sequences of sons and daughters across completed families exhibit substantial variation within regions. This paper introduces a new framework for measuring the heterogeneity in childbearing strategies that generates this variation. This framework identifies the combinations of preferences that best explain observed sequences of sons and daughters. I estimate that at least 40 percent of parents in Africa, Asia, and the Americas make childbearing decisions that depend on the sex of previous children. Substantial shares of parents do not want a balance of sons and daughters, and the direction of these preferences identifiably favors sons only in Asia. Together, these findings suggest that sex preferences are widespread and heterogeneous.

These findings provide a richer account of the role of gender in childbearing decisions. While differential treatment of and opportunities for boys and girls is widespread (Arnold 1997), the study of sex preferences during childbearing has focused largely on Asia and a handful of countries outside of Asia where parents overwhelmingly want sons or want a balance of sons and daughters. This paper reveals that, in much of the world, sex preferences are present but varied. Some parents want sons, others daughters, and others a balance of sons and daughters. By calculating bounds on sex preferences – bounds that are meaningfully narrow – this paper connects the study of sex preferences with a growing literature on partial identification in economics, reviewed by Tamer (2010).

Heterogeneity in sex preferences within a population implies heterogeneity in the underlying tastes, incentives, and constraints that shape sex preferences. I show that variation in agricultural traditions in Africa are associated with whether parents tend to prefer sons or daughters today. These traditions follow ethnic group boundaries that cross country borders. While household surveys are generally conducted at the country level, these findings indicate that alternative groupings of parents can identify greater homogeneity in preferences.

Finally, this paper suggests that, while widespread, sex preferences do not shape overall fertility levels. Sex preferences lead some parents to have more than their ideal total number of children and others to stop childbearing early if they reach a particularly desirable combination of sons and daughters. While individual couples may substantially exceed or stop short of their ideal number of children, in aggregate these effects balance out, and weakening or eliminating sex preferences entirely would only slightly change overall fertility levels. Factors that reduce the number of overall children that parents want to have, such as reduced infant mortality or improved economic opportunities for women, may offer more effective policy levers for reducing fertility.

Works cited

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

Angrist, Joshua, and William Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review*, 88(3): 450-477.

Arnold, Fred. 1997. *Gender Preferences for Children*. DHS Comparative Studies No. 23. Calverton, Maryland: Macro International Inc.

Ashraf, Nava, Erica Field, and Jean Lee. 2014. "Household Bargaining and Excess Fertility: An Experimental Study in Zambia." *American Economic Review*, 104(7): 2210-2237.

Basu, Deepankar, and Robert de Jong. 2010. "Son Targeting Fertility Behavior: Some Consequences and Determinants." *Demography,* 47(2): 521-536.

Ben-Porath, Yoram, and Finis Welch. 1976. "Do Sex Preferences Really Matter?" *Quarterly Journal of Economics*, 90(2): 285-307.

Bhat, Mari, and Francis Zavier. 2003. "Fertility Decline and Gender Bias in Northern India." *Demography,* 40(4): 637-657.

Bolt, Jutta. 2010. "Explaining Long-Run Economic Development in Africa: Do Initial Conditions Matter?" Ph.D. dissertation, University of Groningen

Bongaarts, John. 2013. "The Implementation of Preferences for Male Offspring." *Population and Development Review*, 39(2): 185-208.

Bongaarts, John, and Christophe Z. Guilmoto. 2015. "How Many More Missing Women? Excess Female Mortality and Prenatal Sex Selection, 1970–2050." *Population and Development Review,* 41(2): 241-269.

Chaudhuri, Sanjukta. 2012. "The Desire for Sons and Excess Fertility: A Household-Level Analysis of Parity Progression in India." *International Perspectives on Sexual and Reproductive Health*, 38(4): 178-186.

Clark, Shelley. 2000. "Son Preference and Sex Composition of Children: Evidence from India." *Demography,* 37(1): 95-108.

Clifton, V. L. 2010. "Review: Sex and the Human Placenta: Mediating Differential Strategies of Fetal Growth and Survival." *Placenta*, 31(3): S33-S39.

Conley, Dalton, and Rebecca Glauber. 2006. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility." *Journal of Human Resources*, 41(4): 722-737.

Das Gupta, Monica, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung, and Bae Hwa-Ok. 2003. "Why is Son Preferences 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.

Di Renzo, Gian Carlo, Alessia Rosati, Roberta Donati Sarti, Laura Cruciani, and Antonio Massimo Cutuli. 2007. "Does Fetal Sex Affect Pregnancy Outcome?" *Gender Medicine*, 4(1): 19-30.

Fenske, James. 2009. "Does Land Abundance Explain African Institutions?" Yale University, Department of Economics, Working Paper No. 74.

Freedman, Ronald, and Lolagene C. Coombs. 1974. *Cross-Cultural Comparisons: Data on Two Factors in Fertility Behavior*. New York: Population Council.

Galor, Oded, and David N. Weil. 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review*, 90(4): 806-828.

Guilmoto, Christophe Z. 2012. *Sex Imbalances at Birth: Current Trends, Consequences, and Policy Implications*. Bangkok, Thailand: UNFPA Asia and the Pacific Regional Office.

Hansen, Gary D., and Edward C. Prescott. 2002. "Malthus to Solow." *American Economic Review*, 92(4): 1205-1217.

Haughton, Jonathan, and Dominique Haughton. 1998. "Are Simple Tests of Son Preference Useful? An Evaluation Using Data from Vietnam." *Journal of Population Economics*, 11(4): 495-516.

Henry, Louis. 1961. "Some Data on Natural Fertility." *Biodemography and Social Biology*, 8(2): 81-91.

James, William H. 1971. "Cycle Day of Insemination, Coital Rate, and Sex Ratio." *Lancet*, 1(7690): 112-114.

James, William H. 1990. "On the Magnitude of Variation in the Human Sex Ratio at Birth." *Current Anthropology*, 31(4): 419-420.

James, William H. 2009. "Variation in the Probability of a Male Birth Within and Between Sibships." *Human Biology*, 81(1): 13-22.

Jayachandran, Seema, and Ilyana Kuziemko. 2011. "Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India." *Quarterly Journal of Economics*, 126(3): 1-54.

Jensen, Robert. 2002. "Equal Treatment, Unequal Outcomes? Generating Sex Inequality through Fertility Behavior." Unpublished document, J. F. Kennedy School of Government, Harvard University, Cambridge MA.

Keyfitz, Nathan. 1977. *Introduction to the Mathematics of Population with Revisions*. Reading, Massachusetts: Addison-Wesley.

Kumar, Anuradha, Leila Hessini, and Ellen M.H. Mitchell. 2009. "Conceptualising Abortion Stigma." *Culture, Health, & Sexuality*, 11(6): 625-639.

Larsen, Ulla. 2005. "Research on Infertility: Which Definition Should We Use?" *Fertility and Sterility*, 83(4): 846-852.

Lloyd, Cynthia B, and Anastasia J. Gage-Brandon. 1994. "High Fertility and Children's Schooling in Ghana: Sex Differences in Parental Contributions and Educational Outcomes." *Population Studies*, 48(2): 293-306.

Maharaj, Pranitha, and John Cleland. 2006. "Condoms Become the Norm in the Sexual Culture of College Students in Durban, South Africa." *Reproductive Health Matters*, 14(28): 104-112.

McClelland, Gary. 1979. "Determining the Impact of Sex Preferences on Fertility: A Consideration of Parity Progression Ratio, Dominance, and Stopping Rule Measures." *Demography*, 16(3): 377-388.

McKean, Joseph W., and Gerald L. Sievers. 1987. "Coefficients of Determination for Least Absolute Deviation Analysis." *Statistics & Probability Letters*, 5(1): 49-54.

Michalopoulos, Stelios, and Elias Papaioannou. 2014. "National Institutions and Subnational Development in Africa." *Quarterly Journal of Economics*, 129(1): 151-213.

Milazzo, Annamaria. 2014. "Son Preference, Fertility and Family Structure: Evidence from Reproductive Behavior among Nigerian Women." World Bank Policy Research Working Paper 6869.

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

Murdock, George. 1967. "Ethnographic Atlas: A Summary." *Ethnology*, 6(2): 109-236.

Mutharayappa, Rangamuthia, Minja Kim Choe, Fred Arnold, and T. K. Roy. 1997. "Son Preference and Its Effect on Fertility in India." National Family Health Survey Subject Reports, Number 3.

Novitski, E., and L. Sandler. 1956. "The Relationship between Parental Age, Birth Order, and the Secondary Sex Ratio in Humans." *Annals of Human Genetics*, 21(2): 123-131.

Nunn, Nathan, and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review,* 101(7): 3221-3252

Park, Chain Bin. 1983. "Preference for Sons, Family Size, and Sex Ratio: An Empirical Study in Korea." *Demography,* 20(3): 333-352.

Pickles, A. R., R. Crouchley, and R. B. Davies. 1982. "New Methods for the Analysis of Sex Ratio Data Indepdendent of the Effects of Family Limitation." *Annals of Human Genetics,* 46(1): 75-81.

Pop-Eleches, Cristian. 2006. "The Impact of an Abortion Ban on Socioeconomic Outcomes of Children: Evidence from Romania." *Journal of Political Economy,* 114(4): 744-773.

Robitaille, Marie-Claire. 2013. "Determinants of Stated Son Preference in India: Are Men and Women Different?" *Journal of Development Studies,* 49(5): 657-669.

Romano, Joseph, and Azeem Shaikh. 2008. "Inference for Identifiable Parameters in Partially Identified Econometric Models." *Journal of Statistical Planning and Inference,* 138(9): 2786-2807.

Seidl, Christian. 1995. "The Desire for a Son is the Father of Many Daughters." *Journal of Population Economics,* 8(2): 185-203.

Sen, Amartya. 1990. "More than 100 Million Women are Missing." *The New York Review of Books,* 37(20): 61-66.

Sheps, Midel. 1963. "Effects on Family Size and Sex Ratio of Preferences Regarding the Sex of Children." *Population Studies,* 17(1): 66-72.

Stanton, Cynthia, Noureddine Abderrahim, and Kenneth Hill. 2000. "An Assessment of DHS Maternal Mortality Indicators." *Studies in Family Planning,* 31(2): 111-123.

Tamer, Elie. 2010. “Partial Identification in Econometrics.” *Annual Review of Economics*, 2(1): 167-195.

Wolpin, Kenneth. 1984. “An Estimable Dynamic Stochastic Model of Fertility and Child Mortality.” *Journal of Political Economy*, 92(5): 852-874.

World Bank. 2015. “World DataBank.” Available at <http://databank.worldbank.org/data/home.aspx>, downloaded April 3, 2015.

Yamaguchi, Kazuo. 1989. “A Formal Theory for Male-Preferring Stopping Rules of Childbearing: sex Differences in Birth Order and in the Number of Siblings.” *Demography*, 26(3): 451-465.

Table 1. Standard measures of sex preferences in simulated populations

	Population 1 All couples stop after 1st son and have up to 2 children	**Population 2** All couples stop after 1st daughter and have up to 3 children	**Population 3** 60% of couples from population 1, 40% from population 2
Share of boys with a younger sibling	0.00	0.86	0.38
Share of girls with a younger sibling	0.67	0.00	0.38
Average number of siblings of boys	0.33	1.71	0.94
Average number of siblings of girls	1.00	0.57	0.81
Share of all children that are boys	0.50	0.50	0.50
Share of last-born children that are boys	0.75	0.13	0.50
Average share of sons per family	0.63	0.33	0.51
Parity progression ratios			
1st child *2nd child* *3rd child*			
Boy – –	0	1	0.4
Girl – –	1	0	0.6
Boy Boy –	–	1	1
Boy Girl –	–	0	0
Girl Boy –	0	–	0
Girl Girl –	0	–	0

Notes: The distribution of sequences of children across in completed families is calculated assuming a likelihood of one-half that each birth is a boy. In population 1, one-half of parents have a first-born son and stop, one-quarter have a daughter and then a son, and one-quarter have two daughters. In population 2, one-half of parents have a first-born daughter and stop, one-quarter have a son and then a daughter, one-eighth have three sons, and one-eighth have two sons and then a daughter. In population 3, 30 percent of couples have one son, 20 percent have one daughter, 10 percent have one son and then one daughter, 15 percent have one daughter and then one son, 15 percent have two daughters, 5 percent have three sons, and 5 percent have two sons and then a daughter. Among parents that start with each given sequence of sons and daughters, parity progression ratios measure the share that have at least one more child.

Table 2. Parity progression ratios in observed populations

1st child	2nd child	3rd child	Africa	Asia	North America	South America
Boy	–	–	0.817	0.786	0.779	0.761
Girl	–	–	0.816	0.799	0.779	0.759
Boy	Boy	–	0.795	0.688	0.722	0.696
Boy	Girl	–	0.783	0.690	0.699	0.669
Girl	Boy	–	0.784	0.674	0.694	0.675
Girl	Girl	–	0.795	0.765	0.713	0.698
Boy	Boy	Boy	0.776	0.669	0.676	0.66
Boy	Boy	Girl	0.759	0.641	0.633	0.635
Boy	Girl	Boy	0.759	0.634	0.649	0.646
Boy	Girl	Girl	0.769	0.696	0.655	0.654
Girl	Boy	Boy	0.762	0.633	0.654	0.654
Girl	Boy	Girl	0.761	0.696	0.651	0.652
Girl	Girl	Boy	0.765	0.639	0.642	0.643
Girl	Girl	Girl	0.778	0.757	0.660	0.666

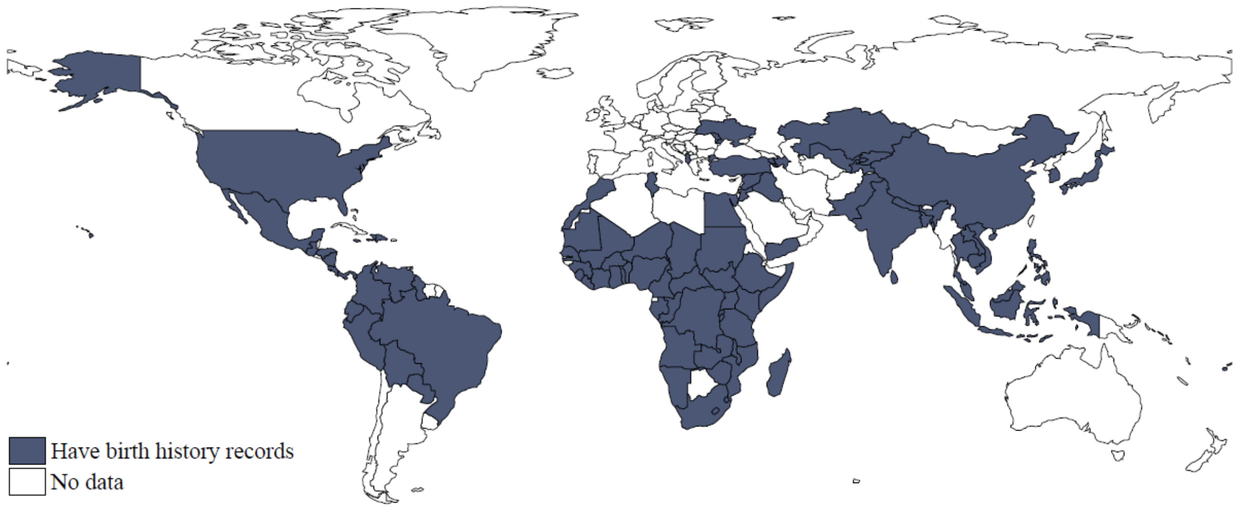
Notes: Among parents that start with each given sequence of sons and daughters, parity progression ratios measure the share that have at least one more child. Source: See Appendix Table 1.

Table 3. Structure of matrices used in section 3

		Share in population		Likelihood of each sequence of children given each combination of preferences						
			Ideal share of children that are boys, r^*:	0.0	0.0	···	0.5	···	1.0	1.0
1st	2nd		Ideal number of children, c^*:	0	0	···	1	···	2	2
child	child		Relative importance of the sex of children, α:	0.0	0.1	···	0.5	···	0.9	1.0
–	–	0		1	1	···	0	···	0	0
Boy	–	½		0	0	···	½	···	½	½
Girl	–	0		0	0	···	½	···	0	0
Boy	Boy	0		0	0	···	0	···	0	0
Boy	Girl	0		0	0	···	0	···	0	0
Girl	Boy	¼		0	0	···	0	···	¼	¼
Girl	Girl	¼	**D**	0	0	···	0	···	¼	¼ **A**

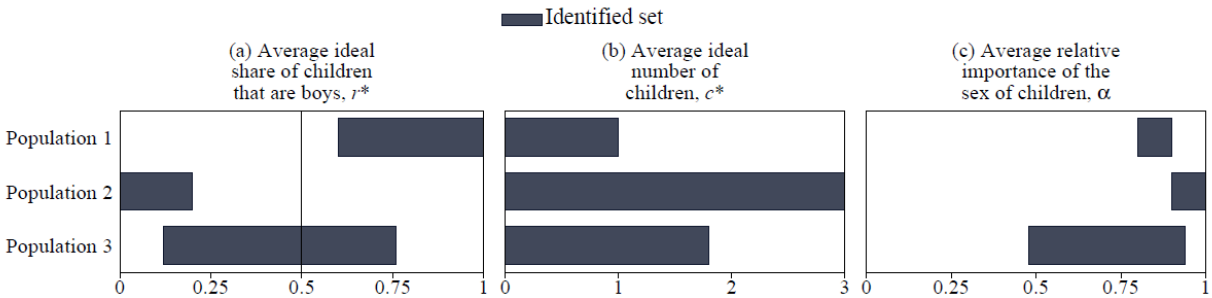
Notes: See section 3.3.

Figure 1. Countries represented in final dataset



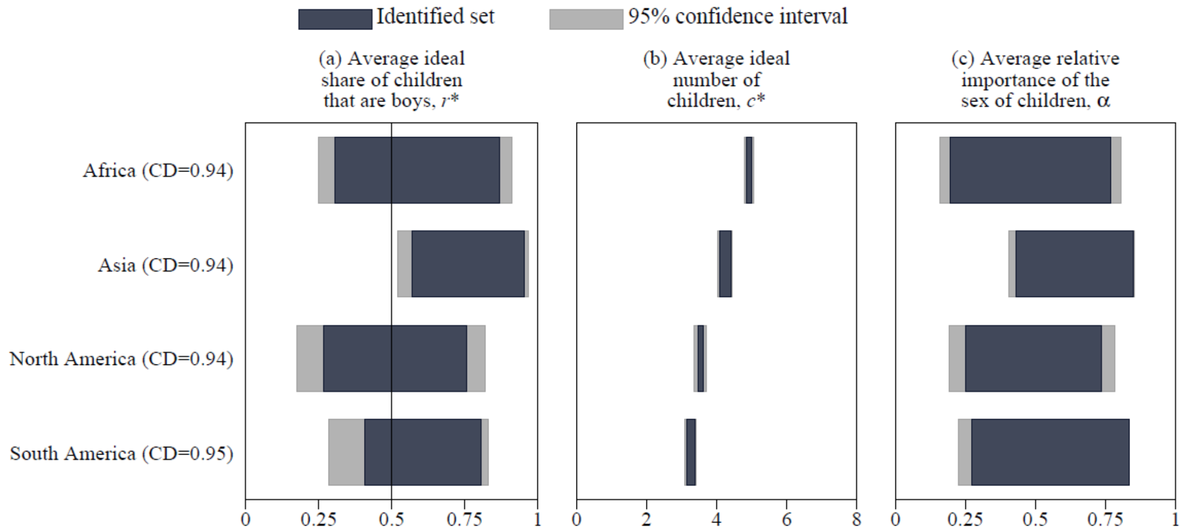
Source: See Appendix Table 1.

Figure 2. Bounds on average values of preferences in simulated populations



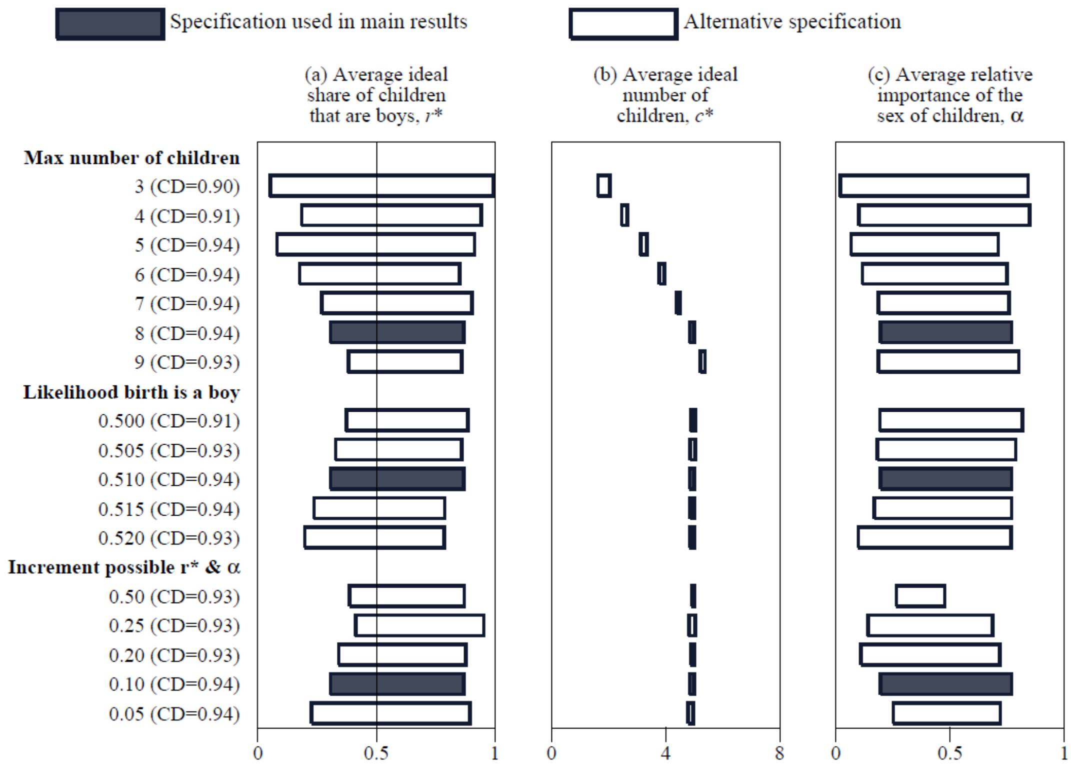
Notes: Populations defined in Table 1. The solid rectangles provide estimated bounds on the average value of preferences that best explain the observed populations, calculated according to the procedure given in section 3, assuming three childbearing periods, a one-half chance that each birth is a boy, and perfect control over conception. For example, in population 1, the average ideal share of children that are boys could be anywhere between 0.6 and 1.

Figure 3. Estimated bounds on average values of preferences, by continent



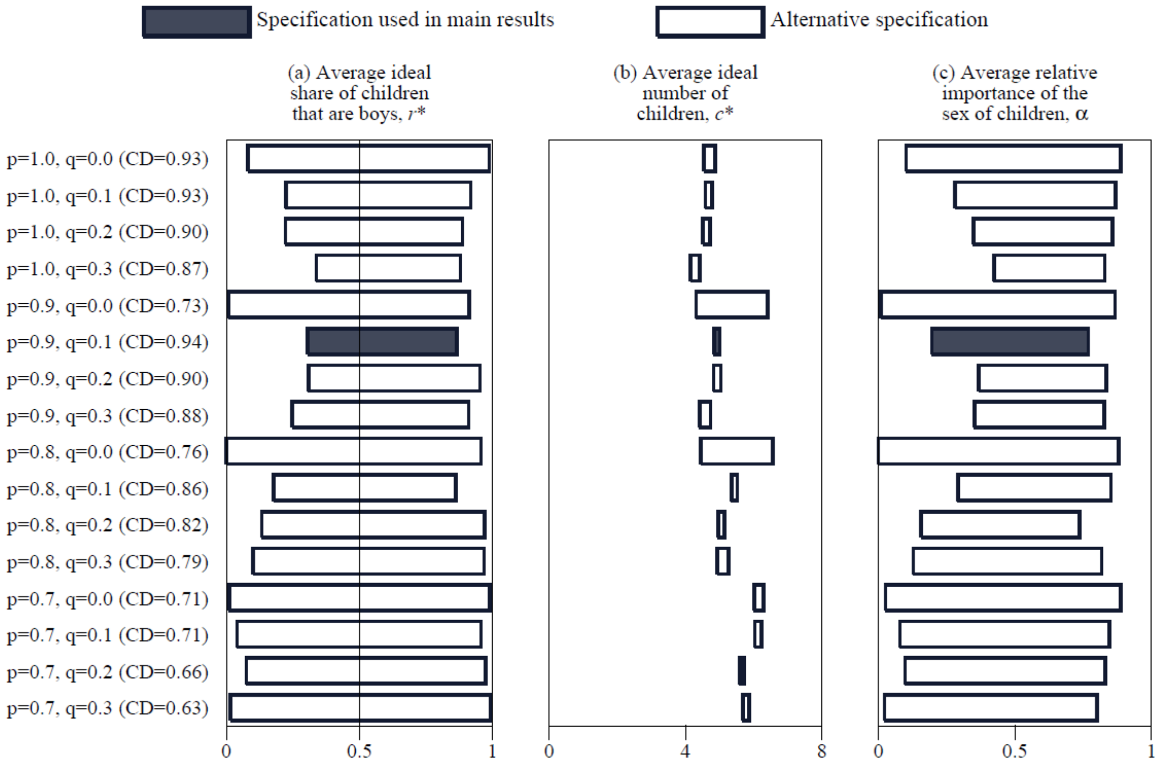
Notes: The dark rectangles provide bounds on average values of preferences, calculated according section 3 with parameter values assumed as given in section 4. The light rectangles provide 95-percent confidence intervals around the unknown true value, calculated from 1,000 subsamples drawn according to Romano and Shaikh (2008). Coefficients of determination (CD) calculated according to McKean and Sievers (1987). For example, in Africa the average ideal share of children that are boys is between 0.25 and 0.9. Sample defined in Appendix Table 1, restricted to women that gave birth eight or fewer times, were age 40 or above when observed, and do not have any twin births.

Figure 4a. Robustness of preferences in Africa to choice of model parameters



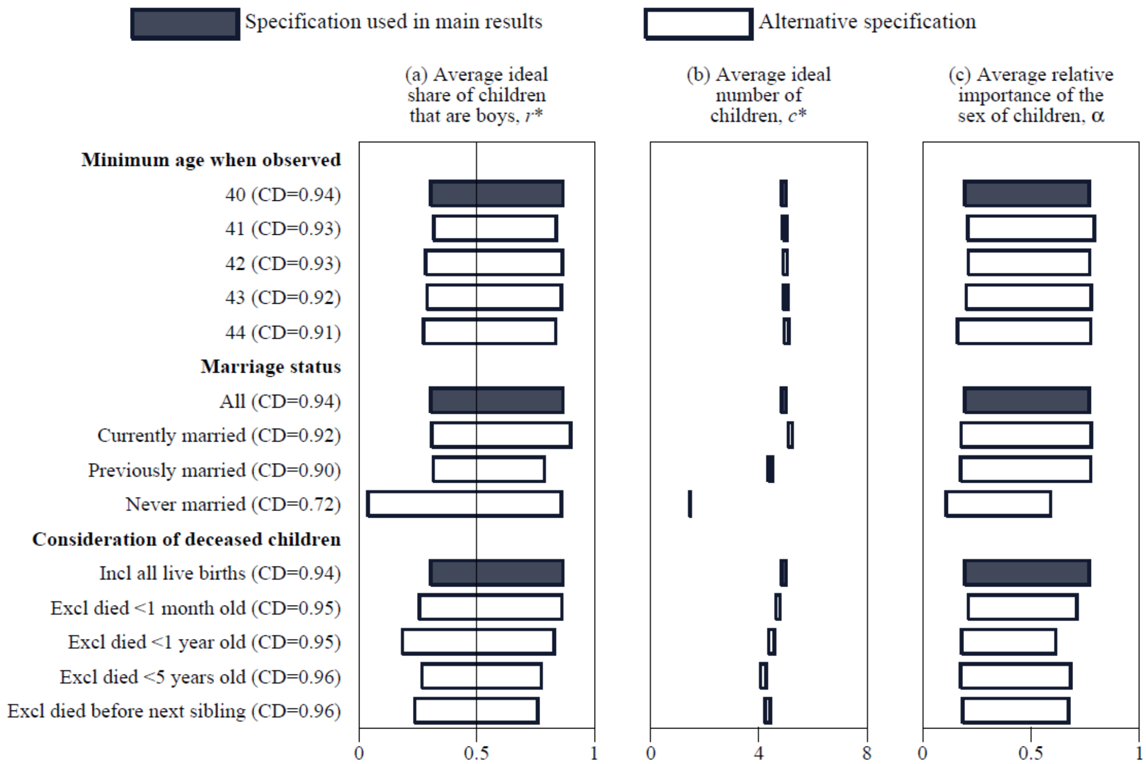
Notes: This figure compares the estimated bounds on average values of preferences in Africa (solid rectangles) with estimated bounds under alternative assumptions (hollow rectangles). Calculations are performed according to section 3, with the main parameter assumptions provided in section 4. Coefficients of determination (CD) calculated according to McKean and Sievers (1987). The first set of comparisons varies the number of childbearing periods, *T*, which is also the maximum number of children permitted in the sample. The second set of comparisons varies the increment between every possible ideal share of children that are sons, *r**, and every possible relative importance of the sex of children, *α*. The final set of comparisons varies the weight placed on each deviation between the actual and predicted populations in equations 2 and 3. Sample defined as in Figure 3.

Figure 4b. Robustness of preferences in Africa to assumed control over conception



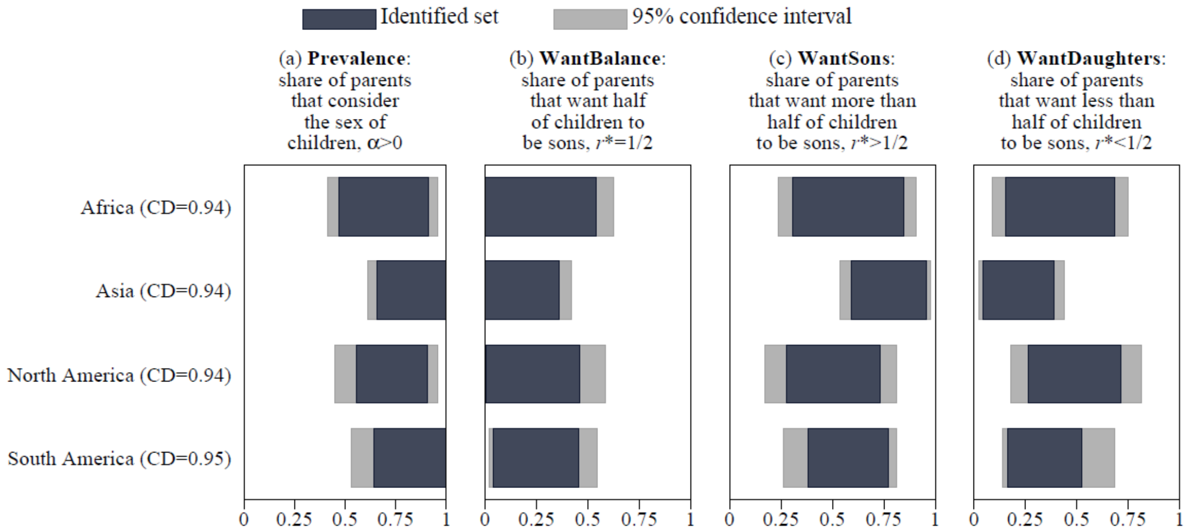
Notes: This figure compares the estimated bounds on average values of preferences in Africa (solid rectangles) with estimated bounds under alternative assumptions (hollow rectangles). Calculations are performed according to section 3, with the main parameter assumptions provided in section 4. Coefficients of determination (CD) calculated according to McKean and Sievers (1987). *p* is the assumed likelihood of conception when trying to get pregnant, and *q* is the assumed likelihood of conception when not trying to get pregnant. Sample defined as in Figure 3.

Figure 4c. Robustness of preferences in Africa to changes in sample definition



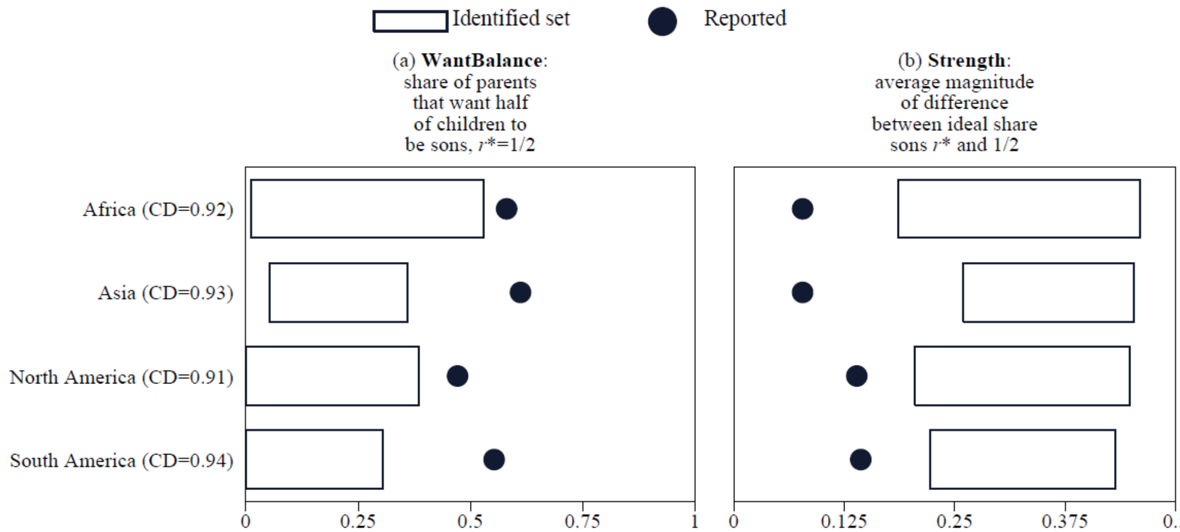
Notes: This figure compares the estimated bounds on average values of preferences in Africa (solid rectangles) with estimated bounds under alternative assumptions (hollow rectangles). Calculations are performed according to section 3, with the main parameter assumptions provided in section 4. Coefficients of determination (CD) calculated according to McKean and Sievers (1987). The first set of comparisons varies the minimum age at which women's birth histories are included in the analysis sample. The second set varies the marriage status of women whose birth histories are included in the analysis sample. The final set varies how deceased children are included in the analysis sample. Sample defined as in Figure 3.

Figure 5. Additional summary measures of sex preferences, by continent



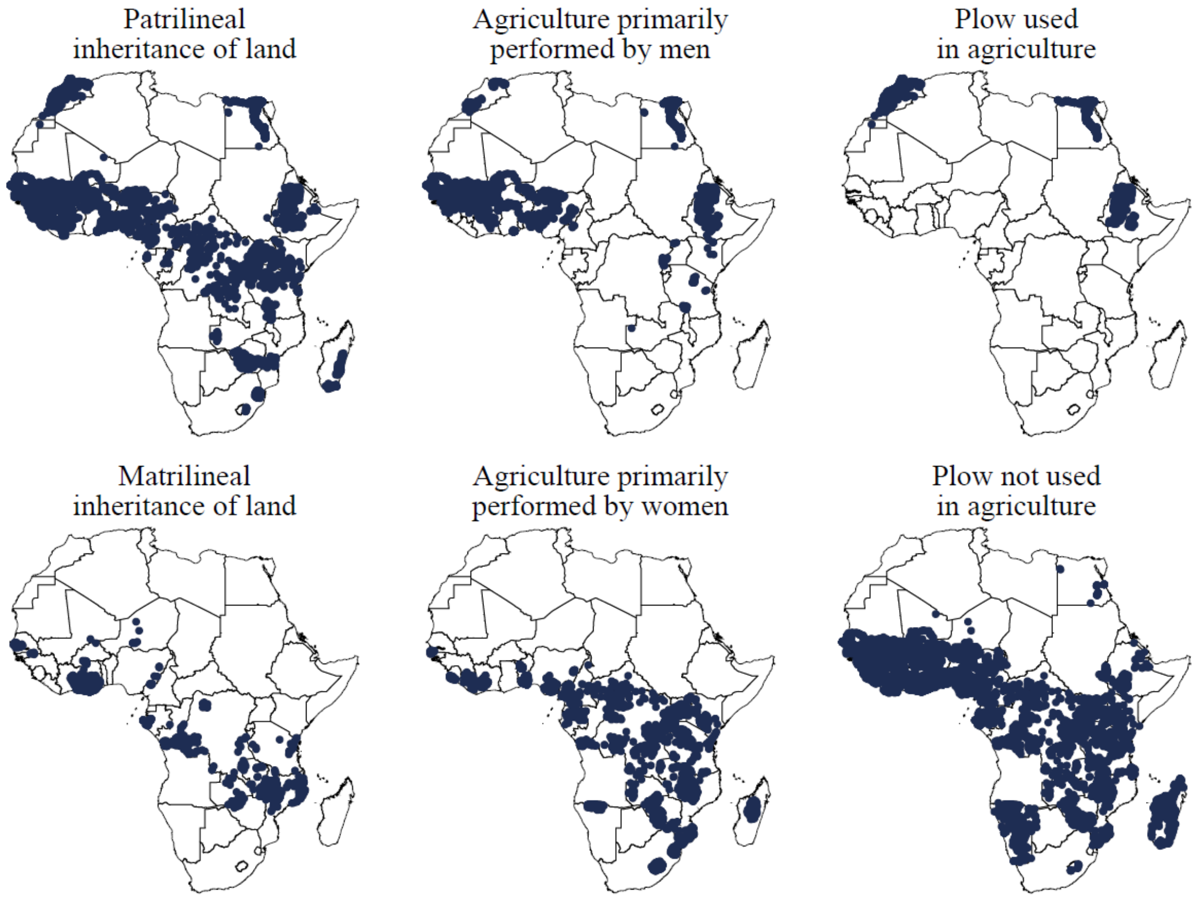
Notes: The dark rectangles provide bounds on summary measures of preferences, defined in section 4.3 and calculated according to section 3. The light rectangles provide 95-percent confidence intervals around the unknown true value, calculated from 1,000 subsamples drawn according to Romano and Shaikh (2008). Coefficients of determination (CD) calculated according to McKean and Sievers (1987). For example, in Africa the share of parents that make childbearing decisions that depend on the sex of previous children is between 0.4 and 0.95. Sample defined as in Figure 3.

Figure 6. Comparison between estimated and reported preferences, by continent



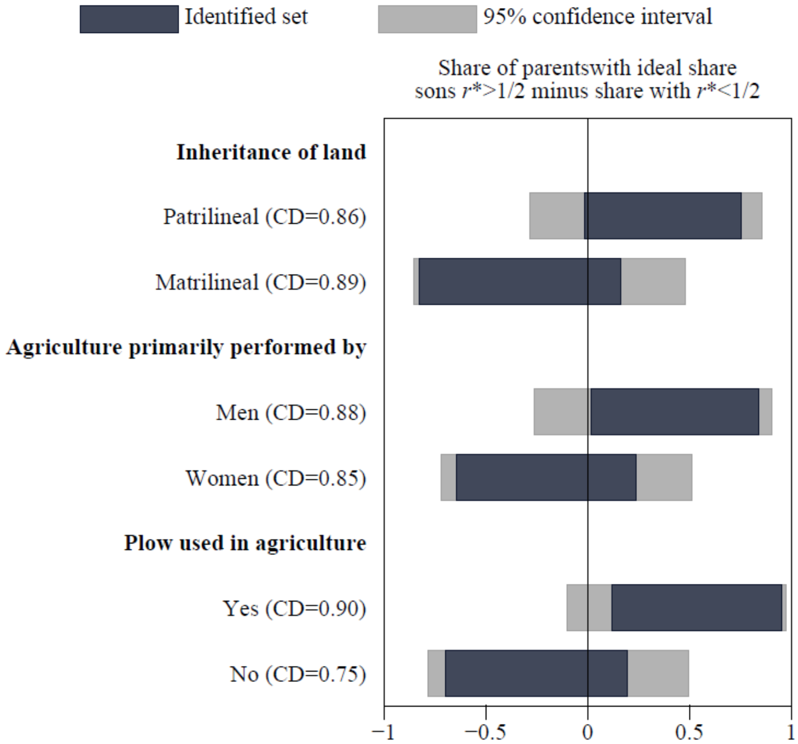
Notes: The hollow rectangles provide bounds on summary measures of preferences, defined in sections 4.3 and 4.4 and calculated according to section 3. The solid circles provide average reported values of these summary measures, as recorded in Demographic and Health Surveys. The surveys ask women to report the number of total children, sons, and daughters they wanted at the start of their childbearing career. Coefficients of determination (CD) calculated according to McKean and Sievers (1987). For example, 60 percent of parents in Africa report wanting a balance of sons and daughters, while estimated values of this share are between 2 and 52 percent. Sample defined as in Figure 3, restricted to include only women who report wanting eight or fewer children and whose total number of desired children equals the sum of their desired number of sons and desired number of daughters.

Figure 7. Respondent locations by agricultural traditions in Africa



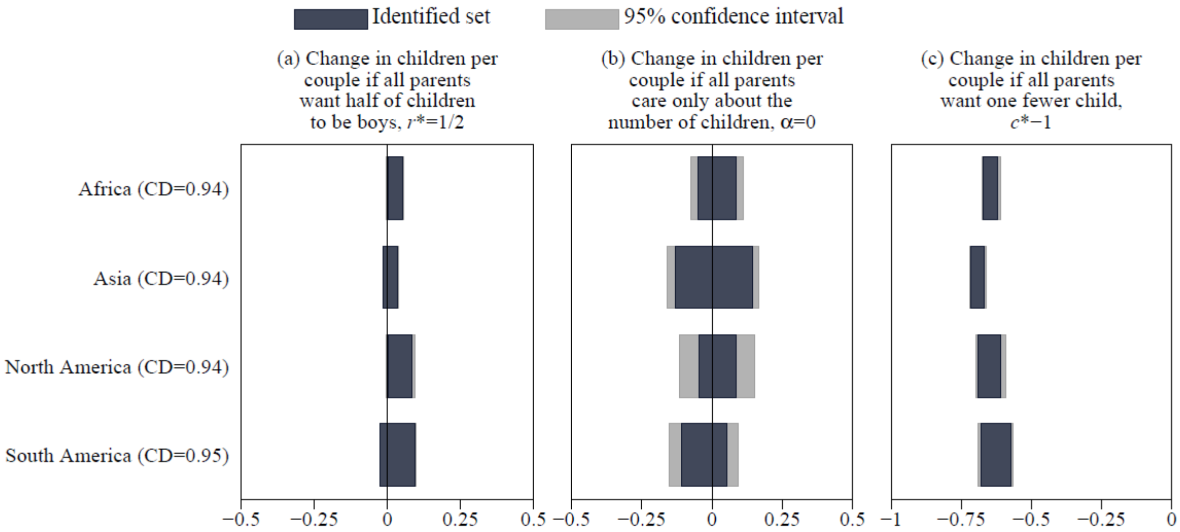
Notes: This figure provides the location of birth history survey respondents in Africa, divided by traditional agricultural practices. Ethnic group characteristics are taken from George Murdock's Ethnographic Atlas (1967). Not all characteristics are recorded for all ethnic groups. Of the 533 ethnic groups in Africa, inheritance of land traditionally follows the patrilineal line in 307 groups, the matrilineal line in 67 groups, neither in 32 groups, and is not recorded in 127 groups. Agriculture is traditionally primarily performed by men in 76 groups, by women in 146 groups, my men and women equally in 102 groups, and is not recorded in 209 groups. The plow was historically used in agriculture in 32 groups, was not used in 455 groups, and was not recorded in 46 groups. Ethnic group boundaries follow Murdock's Tribal Map of Africa (1959). Nunn and Wantchekon (2011) generously make available an electronic shapefile of the Tribal Map, and Fenske (2009) generously provides a crosswalk between the Tribal Map and Ethnographic Atlas. Birth histories are drawn from all Demographic and Health Surveys that record respondent latitude and longitude. These locations are recorded with up to five kilometers of imprecision, so I exclude respondents that live within five kilometers of an ethnic group boundary..

Figure 8. Direction of sex preferences by ethnic group characteristic in Africa



Notes: The dark rectangles provide bounds on the direction of sex preferences, defined in section 5 and calculated according to section 3. The light rectangles provide 95-percent confidence intervals around the unknown true value, calculated from 1,000 subsamples drawn according to Romano and Shaikh (2008). Coefficients of determination (CD) calculated according to McKean and Sievers (1987). Estimates are calculated according to traditional agricultural practices as given in Figure 7. For example, in areas where the plow was traditionally used in agriculture, the identified set for the direction of preferences is mostly positive, suggesting son preference. Sample defined as in Figure 7.

Figure 9. Estimated change in children per couple as preferences change



Notes: The dark rectangles provide bounds on the estimated change in overall fertility levels as preferences change, defined in section 6 and calculated according to section 3. The light rectangles provide 95-percent confidence intervals around the unknown true value, calculated from 1,000 subsamples drawn according to Romano and Shaikh (2008). Coefficients of determination (CD) calculated according to McKean and Sievers (1987). For example, if all parents in the sample in Africa desired half of their children to be boys, overall fertility would rise by at most 0.05 children per woman. Sample defined as in Figure 3.

Appendix Table 1. Birth history sources

Country	Sources	Women
Africa		
Angola	DHS(2011)	8,589
Benin	DHS(1996,2001,2006,2011) WFS(1981)	50,121
Burkina Faso	DHS(1992,1998,2003,2010)	42,363
Burundi	DHS(1987,2010)	13,359
Central African Rep.	DHS(1994)	5,884
Cameroon	DHS(1991,1998,2004,2011) WFS(1978)	43,673
Chad	DHS(1996,2004)	13,539
Comoros	DHS(1996,2012)	8,379
Congo	DHS(2005)	7,051
Congo, Dem. Rep.	DHS(2007,2013)	28,822
Cote d'Ivoire	DHS(1994,1998,2011) WFS(1980)	26,963
Egypt	DHS(1988,1992,1995,2000,2003,2005,2008) WFS(1980)	103,072
Ethiopia	DHS(1992,1997,2003)	45,952
Gabon	DHS(2000,2012)	14,605
Ghana	DHS(1988,1993,1998,2003,2008) MICS(2011) WFS(1979)	41,225
Guinea	DHS(1999,2005,2012)	23,849
Kenya	DHS(1988,1993,1998,2003,2008) WFS(1977)	47,307
Lesotho	DHS(2004,2009) WFS(1977)	18,322
Liberia	DHS(1986,2006,2008,2013)	25,967
Madagascar	DHS(1992,1997,2003,2008)	38,643
Malawi	DHS(1992,2000,2004,2010) MICS(2006,2013)	103,253
Mali	DHS(1987,1995,2001,2006,2012)	50,760
Mauritania	MICS(2011) WFS(1981)	16,209
Morocco	DHS(1987,1992,2003) WFS(1980)	37,836
Mozambique	DHS(1997,2003,2011)	34,942
Namibia	DHS(1992,2000,2006,2013)	31,156
Niger	DHS(1992,1998,2006,2012)	34,463
Nigeria	DHS(1990,1999,2003,2008,2010,2013)	104,888
Rwanda	DHS(1992,2000,2005,2007,2010)	49,277
Sao Tome & Principe	DHS(2008)	2,615
Senegal	DHS(1986,1992,1997,2005,2008,2010,2012) WFS(1978)	81,670
Sierra Leone	DHS(2008,2013)	24,032
Somalia	MICS(2006)	6,639
South Africa	DHS(1998)	11,735
South Sudan	MICS(2010)	8,051
Sudan	MICS(2010) WFS(1978)	15,287
Swaziland	DHS(2006) MICS(2010)	9,672
Tanzania	DHS(1991,1996,1999,2004,2009)	41,855
Togo	DHS(1988,1998)	11,929
Tunisia	DHS(1988) MICS(2011) WFS(1978)	18,513
Uganda	DHS(1988,1995,2000,2006,2009,2011)	40,385
Zambia	DHS(1992,1996,2001,2007)	29,885
Zimbabwe	DHS(1988,1994,1999,2005,2010) MICS(2009,2014)	60,062

Sources: China Fertility Survey (CFS, a 10 percent sample of the "Two-per-thousand" survey); Demographic and Health Survey (DHS); India Rural Economic and Demographic Survey (REDS); Japanese General Social Survey (JGSS); United States Integrated Fertility Survey Series (IFSS); Multiple Indicator Cluster Surveys (MICS); and World Fertility Survey (WFS). Only surveys that meet the following conditions are included: surveys that have national coverage; surveys that collect complete birth histories; surveys in which less than 1 percent of women report a total number of children that differs from that given in their birth history; and surveys in which less than 1 percent of women report births out of order. The sample excludes birth histories that do not report the order and sex of each of a woman's children. The final column of this table provides the number of women included in the final compiled dataset.

Appendix Table 1 (continued). Birth history sources

Country	Sources	Women
Asia		
Armenia	DHS(2000,2005,2010)	18,918
Azerbaijan	DHS(2006)	8,444
Bangladesh	DHS(1993,1996,1999,2004,2007,2011) WFS(1975)	76,080
Cambodia	DHS(2000,2005,2010)	50,928
China	CFS(1988)	461,393
India	DHS(1992,1998,2005) REDS(1982)	310,083
Indonesia	DHS(1987,1991,1994,1997,2002,2007) WFS(1976)	163,301
Iraq	MICS(2006,2011)	82,380
Japan	JGSS(2000,2001,2002,2005,2006,2008,2010)	12,906
Jordan	DHS(1990,1997,2002,2007,2009,2012)	50,352
Kazakhstan	DHS(1995,1999)	8,571
Korea	WFS(1974)	5,430
Kyrgyzstan	DHS(1997,2012)	12,056
Laos	MICS(2011)	22,476
Malaysia	WFS(1974)	6,317
Maldives	DHS(2009)	7,131
Nepal	DHS(1996,2001,2006,2011) MICS(2014) WFS(1976)	60,724
Pakistan	DHS(1990,2006,2012) WFS(1975)	35,144
Philippines	DHS(1993,1998,2003,2008,2013) WFS(1978)	81,662
Sri Lanka	WFS(1975)	6,809
Syria	WFS(1978)	4,487
Tajikistan	DHS(2012)	9,656
Thailand	DHS(1987)	6,774
Timor-Leste	DHS(2009)	13,137
Turkey	DHS(1993,1998,2003) WFS(1978)	27,601
Uzbekistan	DHS(1996)	4,415
Vietnam	DHS(1997,2002)	11,329
Yemen	DHS(1991) MICS(2006) WFS(1979)	12,357
Europe		
Albania	DHS(2008)	7,584
Moldova	DHS(2005) MICS(2012)	13,438
Ukraine	DHS(2007)	6,841
North America		
Costa Rica	WFS(1976)	3,935
Dominican Rep.	DHS(1986,1991,1996,1999,2002,2007,2013) WFS(1975)	87,735
Guatemala	DHS(1987,1995,1998)	23,584
Haiti	DHS(1994,2000,2005,2012) WFS(1977)	43,909
Honduras	DHS(2005,2011)	42,705
Jamaica	WFS(1975)	3,095
Mexico	DHS(1987) WFS(1976)	16,620
Nicaragua	DHS(1997,2001)	26,694
Panama	WFS(1975)	3,701
Trinidad & Tobago	DHS(1987) WFS(1977)	8,787
United States	IFSS(1970,1973,1976,1982,1988,1995,2002)	59,851
Oceania		
Fiji	WFS(1974)	4,928
South America		
Bolivia	DHS(1989,1993,1998,2003,2008)	62,306
Brazil	DHS(1986,1991,1996)	24,725
Colombia	DHS(1986,1990,1995,2000,2004,2009) WFS(1976)	136,941
Ecuador	DHS(1987) WFS(1979)	11,510
Guyana	DHS(2009) WFS(1975)	9,635
Paraguay	DHS(1990) WFS(1979)	10,448
Peru	DHS(1986,1991,1996,2000,2003,2009,2010,2011,2012) WFS(1977)	218,527
Venezuela	WFS(1977)	4,361

Appendix Table 2. Binomial test of equal parity progression ratios in Asia

1st child	2nd child	3rd child	(1) Number of respondents that reach given sequence	(2) Number of respondents that progress beyond given sequence	(3) Parity progression ratio	(4) Parity progression ratio if even across parity	(5) Binomial test p-value: n=(1) k=(2) p=(4)	(6) Minimum p-value, by parity
Boy	–	–	673,900	530,020	0.786	0.793	0.000	0.000
Girl	–	–	626,359	500,709	0.799		0.000	
Boy	Boy	–	272,723	187,652	0.688	0.703	0.000	0.000
Boy	Girl	–	257,297	177,526	0.690		0.000	
Girl	Boy	–	259,683	175,154	0.674		0.000	
Girl	Girl	–	241,026	184,269	0.765		0.000	
Boy	Boy	Boy	96,438	64,472	0.669	0.670	0.224	0.000
Boy	Boy	Girl	91,214	58,463	0.641		0.000	
Boy	Girl	Boy	90,955	57,706	0.634		0.000	
Boy	Girl	Girl	86,571	60,264	0.696		0.000	
Girl	Boy	Boy	90,359	57,222	0.633		0.000	
Girl	Boy	Girl	84,795	59,042	0.696		0.000	
Girl	Girl	Boy	96,528	61,679	0.639		0.000	
Girl	Girl	Girl	87,741	66,409	0.757		0.000	

Notes: This table demonstrates how the binomial tests discussed in section 2 are calculated. In the compiled dataset described in Appendix Table 1, there are 1,570,861 women living in Asia that provide birth histories. Each row of this table represents a unique sequence of sons and daughters. Column 1 provides the number of women that reach each given sequence, column 2 the number that progress beyond each sequence, and column 3 the corresponding parity progression ratios. Column 4 presents the parity progression ratio if all women with the same number of total children had been equally likely to have another child. For example, of the 1,300,259 women that have one child, 1,030,729, or 79.3 percent, have a second child. Column 5 presents the one-tailed p-values from a binomial test in which the number of trials is given by column 1, the number of successes is given by column 2, and the assumed true probability of success is given by column 4. For example, if the 673,900 women that have a boy all have another child with likelihood 0.793, 534,208 would be expected to have a second child and the likelihood is 0.000 that 530,020 or fewer will in fact have a second child. Column 6 identifies the smallest p-value by parity. For all parities, these smallest p-values are less than 0.001, suggesting that sampling error alone cannot account for the variation in parity progression ratios.

Appendix Table 3. Correspondence between stopping rules and bliss-point utility framework

				Son preference	Daughter preference	Balance preference	No preference
			Stopping rule				
			Min. children:	1	1	1	2
			Max. children:	2	2	3	2
			Stop after:	1st son	1st daughter	1 son and 1 daughter	
			Bliss-point preferences				
			Childbearing periods, T	2	2	3	2
			Ideal share of children that are boys, r^*:	1	0	0.5	0
			Ideal number of children, c^*:	2	2	2	2
Observed sequence of children			*Relative importance of The sex of children, α:*	1	1	1	0
1st child	2nd child	3rd child					
Boy	–	–		½			
Girl	–	–			½		
Boy	Boy	–			¼		¼
Boy	Girl	–			¼	¼	¼
Girl	Boy	–		¼		¼	¼
Girl	Girl	–		¼			¼
Boy	Boy	Boy				⅛	
Boy	Boy	Girl				⅛	
Boy	Girl	Boy					
Boy	Girl	Girl					
Girl	Boy	Boy					
Girl	Boy	Girl					
Girl	Girl	Boy				⅛	
Girl	Girl	Girl				⅛	

Notes: This table demonstrates that the bliss-point utility model 1 explains several standard stopping rules. Each column reports a childbearing strategy, and each row presents the likelihood that a parent following that strategy has the indicated sequence of children under the assumption that each child is a son with likelihood one-half and parents have perfect control over conception. For example, a couple following the son-preferential stopping rule in column 1 has a son with likelihood one-half, a daughter and then a son with likelihood one-quarter, and two daughters with likelihood one-quarter. A couple following the bliss-point preferences strategy in column 1 has the same expected distribution over possible sequences of children. These likelihoods for the bliss-point preference strategies are calculated according to section 3.2.

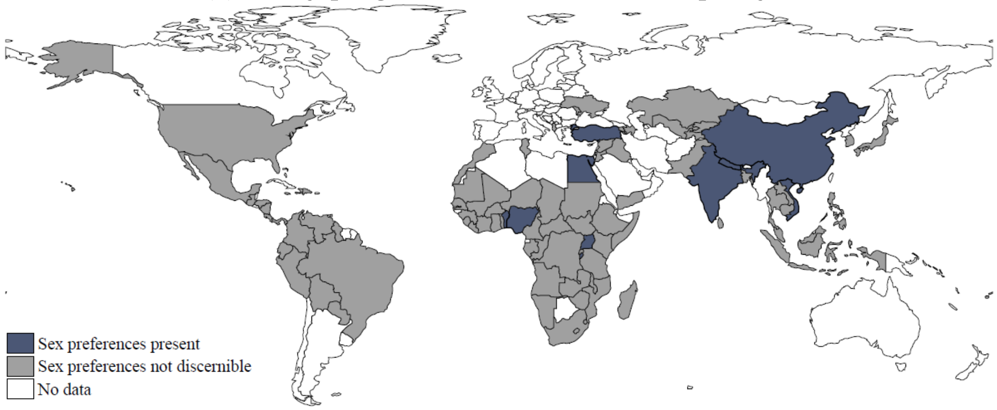
Appendix Table 4. Collapse of childbearing into sequence of sons and daughters

(a) **Original**				(b) **Collapsed by sequence of sons and daughters**		
Period #1	Period #2	Likelihood		Child #1	Child #2	Likelihood
No child	No child	0	→	–	–	0
No child Son	Son No child	$(1–q)l$ $ql(1–q)$	→	Son	–	$(1–q)l + ql(1–q)$
No child Daughter	Daughter No child	$(1–q)(1–l)$ $q(1–l)(1–q)$	→	Daughter	–	$(1–q)(1–l) + q(1–l)(1–q)$
Son	Son	$qlql$	→	Son	Son	$qlql$
Son	Daughter	$qlq(1–l)$	→	Son	Daughter	$qlq(1–l)$
Daughter	Son	$q(1–l)ql$	→	Daughter	Son	$q(1–l)ql$
Daughter	Daughter	$q(1–l)q(1–l)$	→	Daughter	Daughter	$q(1–l)q(1–l)$

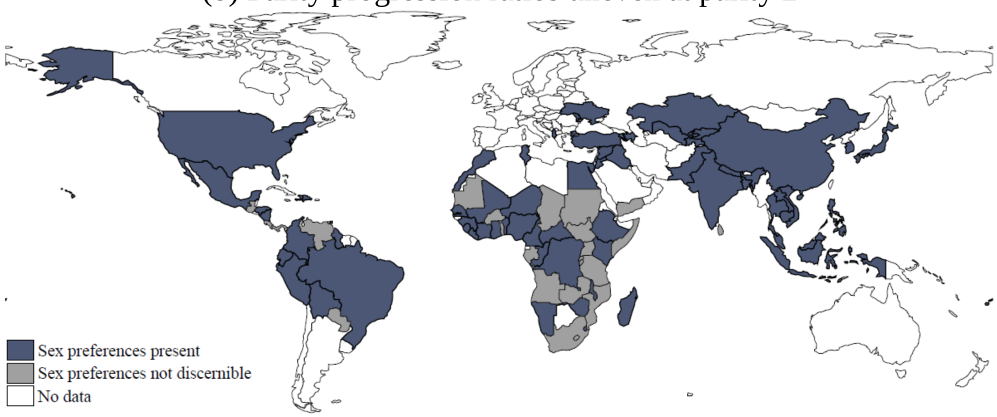
Notes: Panel (a) presents the likelihood that a couple has each possible sequence of childbearing periods with no child, a son, and a daughter, given that the couple wants one child regardless of its sex (α=0, c^*=1, r^* does not matter), has perfect control over conception when trying to get pregnant (p=1), can accidentally conceive when not trying to get pregnant (q>0), and each child is a boy with likelihood l. The couple's childbearing decisions are governed by equation 1 and discussed in section 3.2. The couple does not try to get pregnant in the first period, and then only tries to get pregnant in the second period if the first period did not yield a child. Panel (b) collapses these likelihoods by each unique sequence of sons and daughters, ignoring the timing of each birth.

Appendix Figure 1. Countries in which parity progression ratios indicate sex preferences

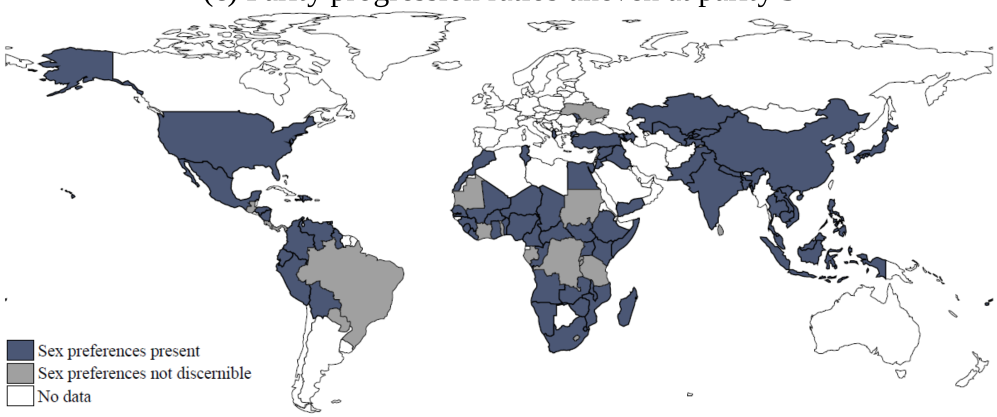
(a) Parity progression ratios uneven at parity 1



(b) Parity progression ratios uneven at parity 2



(c) Parity progression ratios uneven at parity 3



Notes: Using survey data described in Appendix Table 1, this map indicates countries for which the minimum p-value from the binomial test described in Appendix Table 1 is less than 0.05. For these countries, highlighted in dark blue, sampling error alone cannot explain observed variation in parity progression ratios, suggesting that the sex of previous children influences whether parents have another child. For countries highlighted in gray, sampling error alone could explain are discernible at the five-percent level of significance. Panels (a), (b), and (c) perform this test for parity progression ratios among parents that have one child, two children, and three children.