Fertility Decline and Missing Women*

Seema Jayachandran
Northwestern University

June 21, 2014

Abstract

India's male-biased sex ratio has worsened over the past several decades. In combination with the increased availability of prenatal sex-diagnostic technology, the declining fertility rate is a hypothesized factor. Suppose a couple strongly wants to have at least one son. At the natural sex ratio, they are less likely to have a son the fewer children they have, so a smaller desired family size will increase the likelihood they manipulate the sex composition of their children. This paper empirically measures the relationship between desired fertility and the sex ratio. Standard survey questions on fertility preferences ask the respondent her desired number of children of each sex, but people who want larger families have systematically stronger son preference, which generates bias. This paper instead elicits desired sex composition at specified, randomly determined, levels of total fertility. These data allow one to isolate the causal effect of family size on the desired sex ratio. I find that the desired sex ratio increases sharply as the fertility rate falls; fertility decline can explain roughly half of the increase in the sex ratio that has occurred in India over the past thirty years. In addition, factors such as female education that lead to more progressive attitudes could counterintuitively cause a more male-skewed sex ratio because while they reduce the desired sex ratio at any given family size, they also reduce desired family size.

*I thank Suanna Oh, Vrinda Kapur, and Niki Shrestha for excellent research assistance and Diva Dhar and Tarun Jain for allowing me to add questions to our joint survey. I also thank Rebecca Dizon-Ross, David Berger, and several seminar and conference participants for helpful comments. Financial support from the International Initiative for Impact Evaluation and National Science Foundation is gratefully acknowledged. Contact information: Department of Economics, Northwestern University, Evanston, IL 60208; seema@northwestern.edu.

1 Introduction

India's male-biased sex ratio has worsened over the past several decades despite gains for women and girls along other dimensions such as educational attainment, child marriage, and decision-making power in the household (Kishor and Gupta, 2009). One reason is the increased availability of prenatal sex-diagnostic technology, which has made sex-selective abortions possible. Another less obvious reason for the rising sex ratio is the decline in desired family size.[1] Suppose a couple strongly wants to have at least one son. If they wish to have six children, there is only a 1% chance they will be without a son, but if they wish to have only two children, there is a 24% chance.[2] At the natural sex ratio, they are less likely to have a son the fewer children they have, so as their desired family size decreases, the likelihood that they manipulate the sex of their children (through sex-selective abortion, infanticide, or neglect) might increase.

The time trends in many regions of the world are consistent with this idea that the desired number of children has fallen faster than the desired number of sons, putting upward pressure on the sex ratio. Figure 1 shows that the total fertility rate in India has been declining since 1960, while the sex ratio has been rising.[3] Previous scholars have conjectured that falling fertility helps explain time trends in the sex ratio in South Asia (Das Gupta and Bhat, 1997), East Asia (Park and Cho, 1995), Africa (Campbell and Campbell, 1997) and the Caucasus (Guilmoto, 2009), as well as the convergence in the sex ratio between south and north India (Basu, 1999).[4] Portner (2014) estimates a model of fertility in India, using the sex of births and length of birth spacing to infer abortions, and finds that more educated women and women in urban areas use sex-selective abortions more, which he interprets as due to their lower desired fertility. This paper's contribution is to directly estimate the causal relationship between family size

[1] I define the sex ratio as males to females. I use the terms *family size* and *fertility level* interchangeably to refer to total number of children.

[2] The probability is less than 25% because the natural sex ratio is slightly skewed toward males.

[3] The decline in actual fertility reflects a decline in desired fertility over this period, as shown in section 3.3. Other reasons for fertility decline are improved access to contraception which reduces unwanted births, and prenatal sex-diagnostics, which enable parents to select the gender of their children rather than going beyond their desired fertility level to achieve their desired number of sons.

[4] Das Gupta and Bhat (1997) name this the "intensification effect" while Guilmoto (2009) calls it the "fertility squeeze." Das Gupta and Bhat (1997) also conjecture another link between falling fertility and son preference ("parity effect") that could ameliorate the skewed child sex ratio (though not the sex ratio at birth), namely there should be fewer unwanted girls born as fertility rates fall, and unwanted girls have an especially high mortality rate.

and the desired sex ratio. The approach, as described below, uses survey responses on sex composition preferences, with family size specified by the surveyor and randomly determined.

There is an important distinction to note between son preference and how that preference manifests itself in the sex ratio. Families have both a preferred number of sons at any given fertility level (which I call *son preference*) as well as a preferred fertility level (which I call *family size preference*). Holding son preference fixed, the desired ratio of sons to daughters could change when family size preference changes. For example, this will occur if there are diminishing returns to having sons, which could arise because the desire to have at least one son is especially strong.[5] To preview the results, there is indeed a very strong preference for having at least one son in India, and the desired sex ratio falls as family size increases. At the extreme, if families have only one child, the desired population sex ratio is greater than 5 boys for every 1 girl.

By projecting how the sex ratio will change as desired family size changes, holding all else equal, I quantify one important channel for the changing sex ratio. The aim of the analysis is not to project how the sex ratio will evolve over time taking into account all factors. In addition to declining desired fertility, there might also be time trends in son preference (i.e., the desired number of sons at any given fertility level) or in the costs (financial or otherwise) of sex determination and sex-selective abortions.The main goal of this paper is to answer, what is the effect of changes in family size preferences on the desired sex ratio.[6]

The challenge in estimating this effect is to isolate exogenous variation in desired fertility. One approach has been to exploit a constraint on family size, such as the One Child Policy (OCP) in China, which creates variation in actual if not desired fertility.[7] Ebenstein (2010) uses cross-region variation in financial penalties for having extra children to show that OCP led to a more skewed sex ratio. However, OCP was not

[5]Because family size is an integer, the most preferred sex ratio necessarily varies with family size.

[6]There is a large literature on the opposite direction of causality: Son preference affects the gap between actual fertility and desired fertility if families use fertility-stopping behavior to achieve their desired number of sons (Das, 1987; Yamaguchi, 1989).

[7]A change in the ability to control total fertility, for example due to a family planning intervention, is another possibility. A prospective family planning intervention might have only a small effect on total fertility because many women already have some means of controlling their total fertility via sterilization. To my knowledge, there is no exogenous historical variation in sterilization one could analyze, especially after prenatal sex-determination became available in the mid-1980s and manipulation of family sex composition became easier. Moreover, most family planning campaigns have included general maternal and child health counseling and thus might not be valid instruments for total fertility.

gender-blind; exemptions were more forthcoming if the first child was a girl. A second concern is that financial penalties might be related to the degree of son preference in the region.

A related approach is to analyze policies that encourage lower fertility, such as financial schemes in India that jointly reward lower fertility and having relatively more girls. For example, the Devirupak program in the state of Haryana rewards parents if they have either fewer children or a larger fraction girls: The highest payout is given for having one girl (and then becoming sterilized) and a smaller payout is given for having either just one boy or having two girls. Anukriti (2013) finds that the policy reduced fertility and led to a more male-skewed sex ratio. However, because the policy was not gender-blind—it simultaneously incentivized lower fertility and a less male-skewed sex ratio—the analysis likely underestimates the effect of lower fertility on the sex ratio.[8]

In lieu of a natural experiment, a different approach is simply to use cross-sectional variation in desired family size and desired sex ratio to infer the causal relationship. Bhat and Zavier (2003) use data on women's ideal number of children by gender from the National Family Health Survey (NFHS), India's Demographic and Health Survey, to argue that when the ideal family size becomes small, the proportion of boys desired in fact decreases. However, such a cross-sectional analysis gives biased estimates if there is a systematic correlation between an individual's desired family size and degree of son preference. Intuitively, we might expect that those with more traditional values want larger families and also have stronger son preference (and indeed, my data show that such a correlation exists).[9] Thus, the estimates in Bhat and Zavier (2003) are likely driven by omitted variables.

This paper borrows from both of the approaches described above by, first, using exogenous variation in family size and, second, using survey questions on fertility preferences. The idea is simple: A hypothetical total fertility is specified to the survey respondent, and she is asked, given that total fertility, what is her preferred composition of boys and girls. By imposing the total number of children, one can characterize the

[8]Another possible effect of the program is that, analogous to extrinsic motivation crowding out intrinsic motivation, paying people to have daughters might exacerbate son preference.

[9]Basu and De Jong (2010) report that families that have a larger family size in India are more likely to exhibit son preference in their fertility stopping patterns, but this fact is partly (or mainly) due to son preference causing family size to be larger through stopping rules, rather than a positive correlation between son preference and family size preference.

respondents' sex ratio preferences at different exogenously determined fertility levels.[10] For comparison, respondents were also asked the traditional questions about desired family size and the sex composition of those children.

These questions were fielded to men and women in four districts in Haryana, a state in north India. The survey sample comprises parents of secondary school students. The fertility-preference questions asked the respondents about the fertility outcomes they desired for their child rather than themselves, thus avoiding the problems associated with retrospective questions.

Haryana has the most male-biased sex ratio in India, with a child sex ratio (0 to 6-year-olds) of 1.20 based on the 2011 Census, but it is in fact typical of north India in terms of son preference, as shown in Appendix Table 1. Where it differs is that is is wealthier and has lower fertility than the rest of the region.[11] This paper's thesis is that Haryana's low fertility and high sex ratio are related: Due to its low fertility, its son preference translates into a worse sex ratio than seen elsewhere. Thus, Haryana may be a harbinger of how the sex ratio will evolve in the rest of north India as fertility falls.

This essential distinction between son preference and how it manifests in the sex ratio is not just specific to India. Figure 2 shows that across 53 low- and middle-income countries, higher GDP per capita is associated with weaker son preference, specifically a lower desire to have more sons than daughters. But, the child sex ratio is *more* male-skewed in richer countries. Economic development appears to mitigate son preference but not the problem of missing girls. Meanwhile, desired fertility falls sharply with GDP per capita.[12] What is also striking in the figure is that India is not much of an outlier in terms of wanting more sons than daughters, but it is an outlier in terms of its low desired fertility and its high sex ratio. Thus, both the worldwide patterns and the ways

[10]To prevent respondents from anchoring on their first answer, each respondent was asked the sex composition question for only one, randomly chosen, fertility level.

[11]Appendix Table 1 compares Haryana to the other "Hindi belt" states in north India using round 3 of the NFHS. Haryana is wealthier than the other Hindi belt states, as measured by electrification, access to piped water, average education, and a broad-based wealth index. Actual and desired fertility are lower in Haryana. Meanwhile, son preference is similar (slightly lower) in Haryana compared to the rest of the region; son preference is measured as the proportion who want more sons than daughters or the gender gap in schooling.

[12]The data are from Demographic and Health Surveys (DHS) of ever-married women age 15 to 49. All three univariate relationships are statistically significant at the 5% level. Note that the magnitude of the negative relationship between GDP per capita and wanting more sons than daughters might be underestimated because, as shown later, wanting more sons than daughters is more common when family size is smaller (so when GDP per capita is higher).

in which India is anomalous are consistent with the ideas put forth in this paper.

The main result of the paper is that the desired sex ratio increases sharply as the fertility rate falls. When the family size specified to the respondent is 3 children, the desired sex ratio is 1.12, while with 2 children, it rises to 1.20 and with 1 child it rises to 5.64. Thus son preference is far from homothetic; it is characterized by a strong desire to have at least one son and a considerably more moderate desire to have at least two sons. These results point to the importance of cultural rather than economic roots of son preference. The primacy of the eldest son in Hinduism – from his role in funeral rituals to family inheritance – would give rise to this strong preference for the first son and then a preference for gender balance more or less thereafter. If son preference was due to contemporaneous economic factors, for example the lower labor market potential of women, then it would likely look much more homothetic.

A second result is that the desired sex ratio is actually below 1 when family size is 4 or higher. Not only are there diminishing returns to having sons, there is a crossover: once a couple has 1 or 2 sons, they prefer to have a higher proportion of daughters than sons. This crossover has not been noted in the literature before, to my knowledge.

These data enable one to estimate how much of the sex ratio trend in India in recent decades can be explained by the declining fertility rate. I show that roughly one half of the increase in the sex ratio over the past thirty years can be accounted for by the shift toward smaller family sizes.

The fertility preference data also demonstrate why comparisons using observational variation in desired family size are misleading: Desired family size and son preference are positively correlated. The families that want small families have atypically low son preference. One implication is that if, in future years, desired fertility decreases most among those who currently intend to have a large family, this will further exacerbate the problem of missing women: Those with the strongest son preference will develop a need to manipulate the sex of their children as they adjust downward their desired family size.

Finally, I examine how female education affects the sex ratio. Rather than holding son preference fixed as the earlier analyses did, I allow it to also change, with the goal of illustrating the offsetting effects that progressive forces have on the sex ratio. On the one hand, female education leads weaker son preference, that is a desire for fewer sons at any given family size. On the other hand, it leads to a decline in desired fertility—which, as I have argued, will lead to a higher desired sex ratio, all else equal. Thus, combined,

it is ambiguous how increased female education, and progressive forces more generally, will affect the desired sex ratio. I show that for female education, the two offsetting forces cancel out; the net effect on the desired sex ratio is essentially zero with the point estimate suggesting that, if anything, this modernizing force, counterintuitively, causes individuals to want a more male-skewed sex composition of children.

2 Data

2.1 Sample

The data for the analysis were collected in September 2013 to January 2014 as part of a baseline survey conducted to evaluate a secondary-school-based gender sensitization program in four districts of Haryana, India. The four study districts—Jhajjar, Panipat, Rohtak, and Sonipat—are adjacent to New Delhi and have lower fertility and a more skewed sex ratio than average for Haryana.

Many of the particular features of the sampling strategy were for the purpose of the ongoing randomized evaluation, such as a maximum of one school per village to minimize spillovers and oversampling grade 6 girls for whom we expect the largest program impacts. The sampling strategy was, first, to select 314 government secondary schools (from among the roughly 350 in the 4 districts), excluding schools with low enrollment or high attrition between grades 6 to 8 and including at most one school per village. Second, within these schools, on average 45 students were selected to be interviewed; since in some schools, there were fewer than 45 students eligible to be surveyed, in other schools, the sample size was higher. Male and female students in grades 6 and 7 were included in the sample; grade 6 girls comprise 33.3% of the sample, and grade 6 boys, grade 7 girls, and grade 7 boys each comprise 22.2% of the sample. To be eligible for the survey, one of the student's parent needed to provide informed consent, and the student needed to provide informed assent to be surveyed. The surveys of students were conducted in the schools.

Then, for a random 40% of the surveyed students, one of their parents was surveyed; the reason for surveying only a subset of parents was a budgetary constraint. For this subsample, surveyors visited the household, and either the mother or father was randomly chosen to be surveyed. If the parent was not home or available to be interviewed during the three attempts the survey team made, a replacement household was randomly chosen.

In total, attempts were made to interview 3587 mothers and 3503 fathers, with a response rate of 89.6% for mothers and 70.2% for fathers. The lower response rate for fathers is not surprising as men were more likely away from home working when the survey was conducted. The final sample comprises 3215 mothers and 2460 fathers.[13]

Descriptive statistics for the sample are presented in Table 1. Female respondents are 35 years old on average, and male respondents, 40 years old. The illiteracy rate is 39% among females and 16% among males. The sample is 95% Hindu, and about 18% of respondents belong to a scheduled caste. The average number of children the respondents have is 3.5 and the percent sons among their children is 54%.[14]

2.2 Elicitation of fertility preferences

The survey collected data from the parents on family background, gender attitudes (for example, tolerance for gender-based violence), and fertility preferences, the last of which is the focus of this paper. The fertility preference questions are prospective questions about the fertility of the adult respondent's child (or the child's future wife). The purpose of asking prospective rather than retrospective questions about the respondent's own fertility, as is done in the Demographic and Health Surveys (DHS), was to avoid the bias created by answering questions about fertility preferences retrospectively (Westoff and Ryder, 1977; Rosenzweig and Wolpin, 1993). In addition, these forward-looking questions are more relevant when projecting future fertility patterns (as long as parents' preferences influence their children's fertility).[15]

Importantly, the survey questions differ from the standard ones in that a total fertility level is specified, and the parent simply gives the gender mix: "Suppose your

[13]An additional 701 households were interviewed in which the child lived with only one parent. Because some of the analysis will compare results for mothers and fathers and it would not have been possible to randomly select which parent to interview in cases where the child lives with one parent, the analysis is restricted to parents of students who live with both parents. The results are nearly identical when including the additional 701 respondents and are available from the author.

[14]Percent sons is calculated excluding the sampled student; because the sampling frame includes more girls than boys, percent sons is mechanically lower (51%) if the sample student is included. Two percent of parents have a missing value for the percent sons variable because they have only one child.

[15]For the first 22% of students surveyed, the desired fertility questions were also included in the student survey, referring to the student's future fertility. The Haryana Department of Education then requested that these questions be removed from surveys conducted in schools, as they deemed 11 to 13 years old too young for these questions. For the student subsample asked the questions (3387 respondents), the results are similar to those for parents. There is a negative effect of family size on the percent sons desired; the average percent sons desired is 76% when family size is 1 and declines monotonically to 50% by a family size of 5.

son/daughter [the specific grade 6 or 7 child we surveyed] was going to have N children, how many of them would you want to be boys and how many would you want to be girls?"[16] In principle, each respondent could have been asked about several values of N but to avoid anchoring, each parent was asked the question for only one value of N randomly chosen, with equal likelihood, from the integers between 1 and 5. Given the randomized design, cross-person comparisons should accurately measure within-person preferences at different fertility levels. Appendix Table 2 compares the subsamples assigned different values of N. Characteristics are balanced across the groups, with the exception that the percent sons among the respondent's children is marginally different. The empirical results are robust to adjusting for baseline characteristics.

The standard DHS fertility questions were also asked (about the respondent's child), in which the respondent specified both the desired number and sex composition: "How many children do you want your son/daughter to have? How many of these children would you like to be boys, how many would you like to be girls, and for how many does the gender not matter?" (In the survey, these questions were asked before the question that randomly specified the total fertility.)

3 Results

3.1 Negative effect of family size on percent sons desired

Figure 3 presents the main result of the paper graphically. Along the horizontal axis are the five randomly assigned specified family sizes (number of children). The vertical axis plots the average percent sons that are desired by parents. When the family size is 1, the average is over 80% and declines sharply as family size increases. The lower the family size, the more the desired sex composition is skewed toward sons.

Because the responses are similar for mothers and fathers (mothers have slightly higher son preference), the remainder of the analysis focuses on pooled results for mothers and fathers. Table 2 shows the full distribution of responses. The first column is the subsample asked about a family size of 1. The vast majority of respondents, 84.9%, would want this one child to be a son. The bottom rows of the table aggregate these

[16]The questions do not ask for gender preference by the birth order. There does not seem to be a strong preference for sons to come at earlier birth order in India, as evidenced by the fact that the sex ratio is not very skewed for first births and is most skewed for last births (due to both sex selective abortions and stopping rules).

responses and report the average percent sons desired for the population (84.9%), and the corresponding sex ratio (5.6).

The second column is the subsample asked about a family size of 2. The most common preference is one boy, one girl (84.6% of respondents), with 12.4% preferring two boys and 3.1% preferring two girls. These responses correspond to a desired percent sons of 54.6% and sex ratio of 1.20. For family size of 3, the responses correspond to 52.8% sons, or a sex ratio of 1.12. While the change in percent sons seen between family size 1 and family size 2 is especially large, the changes are meaningful in magnitude at all family sizes: Reducing family size from 3 to 2 increases the desired population sex ratio from 1.12 to 1.20.

The pattern continues as family size increases to 4 and 5: The larger the family size, the lower the desired percent sons. One quite striking result is that the average percent sons is below 50% for family size of 4 and 5. Respondents appear to strongly want to have 1 or 2 sons but then prefer that the additional children are girls. For example, at a family size of 5, 18.1% of respondents have a preference for 4 or more girls, while only 3.3% have a preference for 4 or more boys. Note that this preference for having more daughters than sons when family size is large is important for understanding the dynamics of the sex ratio, but the broader normative interpretation is unclear; it could be the case that parents want daughters so they can help care for siblings and perform household chores.

Table 3 tests the statistical significance of the negative relationship between family size and the percent sons desired. In an ordinary least squares regression, the percent sons decreases by 8.5 percentage points for every additional child, with a p-value < 0.01.[17] In column 2, estimating the coefficients separately for each family size, the monotonic decline is statistically significant at the 10% level or lower at each increment. The results are nearly identical adjusting for differences in baseline characteristics, as shown in columns 3 and 4.[18]

[17]The analysis pools parents' responses about either their son or their daughter (whoever was surveyed for the student sample). The results are very similar for responses about sons and daughters. When the regression reported in Table 3, column 1, is estimated separately for the subsample of parents asked about their daughters, the coefficient on *Randomly-specified family size* is -0.086; for the subsample asked about their sons, the coefficient is -0.083.

[18]The survey question did not allow respondents to say they were indifferent about sex composition. It is unlikely that the systematic patterns seen would arise from respondents being indifferent and giving arbitrary answers when forced to choose. The standard DHS-type question analyzed below did allow for indifference (to be consistent with the standard DHS question), asking respondents how many of the

These patterns shed light on the root cause of son preference. Son preference is not homothetic; as family size grows, the desired percent sons is not constant, but instead falls sharply. To first approximation, families want one son and if that preference is satisfied, want close to an equal number of sons and daughters. If son preference were due to contemporaneous economic factors, for example the lower labor market opportunities of women, then parents would be weighing the same considerations with each birth, and we would expect preferences to be close to homothetic. Instead, the patterns are consistent with cultural factors looming large. The eldest son plays an important role in the patrilineal and patrilocal kinship system of Hinduism through which parents live with their eldest son and the eldest son inherits family property; Hinduism also decrees that, for salvation, a male heir of the deceased must perform their funeral rites such as lighting the pyre (Dyson and Moore, 1983; Das Gupta, 1987; Arnold et al., 1998).

3.2 Correlation between family size preference and son preference

The standard fertility preference questions jointly ask about family size preference and son preference. For any level of son preference, the desired proportion of sons will vary with family size, so one cannot use such data to obtain an unbiased measure of how son preference is correlated with family size preference, or the effect of declining family size on the sex ratio, as the analysis in this subsection illustrates.

Table 4 shows that respondents with a preference for a larger family size also have stronger son preference. Each row is a subsample that reports a different desired family size. The modal response is 2 children.[19] Each column is a subsample asked about sex composition for a different randomly-specified family size (which more often than not differed from their desired family size). Each cell reports the average percent sons desired at the randomly-specified family size for the subsample defined by the row and column. First, for each row, the pattern seen earlier is present: The desired proportion sons is

total desired children they wanted to be boys, how many they wanted to be girls, and for how many they did not care about gender. The results in Table 3 are similar restricting the sample to the 83% of respondents who expressed strict gender preferences over each child's gender in the standard DHS question; the coefficient in column 1 remains -0.085.

[19]Because only 9 respondents report a desired family size larger than 4, they are grouped with those desiring a family size of 4. In addition, 10% of respondents are excluded from this analysis because they did not give a numerical answer to the question, either responding "Up to God," responding "Up to spouse/family," saying they did not know, or giving an inconsistent answer in which the number of desired children by gender did not add up to their desired number of children.

higher the smaller the family size is. Second, as can be seen by looking within columns, the desired proportion of sons at each randomly-specified family size is increasing in the desired family size. This implies that son preference and family size preference are positively correlated in the population. Given that the pattern is seen at each randomly-specified family size, it also holds in aggregate: The average desired percentage of sons at the randomly-specified family size—the unbiased measure of son preference—is increasing in desired family size. Figure 4 shows graphically that son preference is stronger among those who want larger families.

However, if we instead use respondents' answers to the standard questions and calculate their desired proportion of sons when they jointly choose family size and number of sons, the pattern is different; the average percent sons desired at the *desired* family size exhibits a non-monotonic pattern, as seen in Figure 4.[20] The pattern reflects the combination of the negative effect of family size on the percent sons desired (true effect) and the fact that individuals who prefer a larger family also have stronger son preference (bias).

Table 5 shows the relationship between desired family size and son preference in a regression framework. In columns 1 and 2, son preference is measured as the percent sons desired based on the fertility questions introduced in this study, which specify the family size to the respondent. Column 1 estimates the average linear effect; when desired family size increases by 1, the average percent sons desired increases by 6.6 percentage points. Column 2 estimates separate coefficients for each value of desired family size.

Columns 3 and 4 use the approach from the previous literature (Bhat and Zavier, 2003) in which the standard fertility questions are used, and the goal is measure how declining fertility affects the gender composition in the population. The outcome is the percent sons desired based on the standard question where the respondent specifies his or her desired family size. In column 3, the average linear effect of desired family size is estimated. The coefficient is negative and significant, but at -0.023 is much smaller in magnitude than was seen in Table 3, column 1 (coefficient of -0.085), where the measure of son preference was independent of the respondent's family size preference. Because of the positive correlation between family size preference and son preference, the standard fertility questions underestimate the negative effect of family size on the desired sex

[20]Following the DHS, the individual answered how many of the desired children ideally would be boys, girls, or no gender preference. The no-gender-preference children are counted as 0.5 sons and 0.5 daughters.

ratio. Column 4 shows that underlying the average negative effect is a non-monotonic pattern, as was seen in Figure 4. The coefficient for family size 2 is negative, but the coefficient for family size 3 is less negative, or in other words, the percent sons desired declines between family size 1 and 2, but increases between family size 2 and 3 (and then decreases again between family size 3 and 4). The non-monotonic pattern is due to the estimates combining the true negative effect of family size on the percent sons desired and the positive cross-sectional correlation between son preference and family size preference.

3.3 How much of the sex ratio trend can fertility decline explain?

The results presented show that the desired sex ratio increases when family size is smaller. This fact suggests that the declining total fertility rate (TFR) over the past decades in India may have been an important contributor to the rising sex ratio. That is, the time trend in TFR shown in Figure 1 is one cause of the time trend in the sex ratio shown.

I next quantify how much of the time trend in the sex ratio could be explained by falling TFR. Of course, several other factors besides falling TFR also contribute to changes in the sex ratio (e.g., changing access to ultrasound, changing son preference); the goal here is to hold these other factors fixed and assess whether the contribution of fertility decline to the sex ratio trend is large or small.

To do so, I undertake a straightforward exercise in which I combine the average desired percent sons at different fertility levels shown in Table 2 and annual data on desired total fertility. Note that actual TFR likely exceeds desired TFR, both because families also use stopping rules to achieve their desired number of sons and because of incomplete access to contraception. I construct desired total fertility from rounds 1 to 3 of the NFHS. Specifically, I calculate the distribution of desired number of children by year among women who are age 25 to 34 in that year (i.e., in their prime childbearing years). The average value of desired TFR is shown in Figure 5 and, as expected, is lower than actual TFR, which was shown in Figure 1. Because the earliest NFHS was conducted in 1992, there are no data on desired TFR for 1971 or earlier, so I restrict the time period for this exercise to 1981 to 2011.

I assume that some fraction of the population deviates from the natural sex ratio to

attain their desired sex ratio or, equivalently, their desired proportion of boys; I denote this proportion θ. I assume that the remainder of the population does not manipulate the gender of their children and has the natural proportion of sons or uses stopping rules to achieve their desired number of sons.[21]

I use 1.02 as the natural sex ratio (equivalent to the proportion of sons being 50.74%). Following the literature, I am assuming the natural sex ratio at birth in India is the observed value for Sub-Saharan Africa, which is 1.033 (Sen, 1992; Garenne, 2011). The natural sex ratio for 0 to 6 year olds (the age range for which sex ratio data are available for India) is lower than this because the natural rate of infant and child mortality is higher for boys than girls. The sex ratio of children age five years and younger is 1.017 in Rwanda and 1.021 in Kenya, so I use 1.02 as the best estimate of the natural sex ratio among 0 to 6 year-olds in India (Republic of Kenya, 2009; Goverment of Rwanda, 2009).

For each year from 1981 to 2011, I use the distribution of desired TFR (from the NFHS) and the desired proportion of boys for different levels of TFR (from the Haryana parent survey) to construct the average desired proportion of boys.[22] Suppose that in a certain year 50% of NFHS mothers want 4 children, 30% want 3 children, and the remainder want 2 children. Then the average percent sons desired in that year is 0.5 $\times$ the sample average percent sons desired at a randomly-specified family size of 4 + 0.3 $\times$ the sample average percent sons desired at a randomly-specified family size of 3 + 0.2 $\times$ the sample average percent sons desired at a randomly-specified family size of 2.[23]

To determine θ, the fraction of the population that manipulates its sex ratio, I calculate the value of θ that allows one to match the *level* of the sex ratio in 1981. In other words, there is a unique θ that yields the actual 1981 sex ratio of 1.0395 given

[21]Stopping rules do not affect the population sex ratio because each birth is a random draw at the natural sex ratio; stopping rules do generate a cross-family correlation between family size and the within-family sex ratio (Clark, 2000; Jensen, 2003).

[22]Because the highest family size for which I have fertility preference data is 5, I assume that for desired family size larger than 5, the desired percent sons is the same as it is for a family size of 5.

[23]The data on desired TFR and the child sex ratio are for all of India, while the data on desired sex composition are from the sample of parents in Haryana. It is unclear if using preference data for Haryana will overstate or understate how much fertility decline has caused the sex ratio to rise for India. While Haryana's level of son preference is higher than average for India (because south India has lower son preference), what is relevant for the projection is the slope, i.e., how fast the desired number of sons falls as desired total fertility falls, and this slope could be higher or lower in Haryana. Also, because θ is set by matching the 1981 sex ratio, when the level of son preference is higher, the proportion of people manipulating child gender is assumed to be lower, and the net effect on the projected trend in the sex ratio is unclear.

the distribution of desired TFR in that year. This value of θ is 0.188, or 18.8% of individuals.[24] I assume this fraction remains fixed over time, as does every other factor such as son preference, access to ultrasound, and costs of sex-selective abortions. I calculate how the sex ratio evolves after 1981 simply due to falling desired TFR. The goal is to determine how much of the *trend* in the sex ratio between 1981 and 2011 is explained by fertility decline.

The results of this exercise are shown in Figure 5. The blue line is the projected sex ratio; it is rising over time, though not as much as the actual sex ratio (plotted in gray). The projected sex ratio matches the trend more closely for the second half of the period. This pattern is consistent with the diffusion of ultrasound in the 1980s and early 1990s being the main driver of the rising sex ratio during that period, and then fertility decline playing a major role in the last decade, once access to ultrasound was widespread; families could more fully optimize with respect to their fertility and gender preferences once they were unconstrained by technology. Overall, declining fertility leads to an increase in the sex ratio that is 54% of the actual increase in the sex ratio over the 1981 to 2011 period.[25]

To summarize, the survey questions allow one to quantify the effect of declining fertility preferences on the desired sex ratio, and this magnitude suggests that one half of the recent trend in the sex ratio in India could be explained by declining fertility. Undoubtedly, another key factor is the increasing availability of prenatal sex-diagnostics, so it is reassuring that the projection does not explain all (or more than all) of the actual increase in the sex ratio.[26]

3.4 Net effect of female education on the sex ratio

Thus far, the thought experiment of this paper has been to hold son preference fixed and examine the effects of declining fertility on the sex ratio. In the next and final

[24] One could also allow θ to vary with desired family size. For example, if θ were higher for those who want a smaller family size, then the population-average θ would be increasing over time. However, the data do not allow one to pin down how θ varies with the level of desired fertility, so the additional structure on θ would have to be imposed by assumption.

[25] Alternatively, one can identify θ by matching the 2011 level and run the projection backwards in time. In this case, $\theta = 0.282$ and the change in desired TFR explains 82% of the time trend in the sex ratio. One can also vary the assumed natural child sex ratio. If one matches the 1981 level assuming the natural sex ratio is 1.03, then $\theta = .100$ and fertility decline explains 28% of the sex ratio trend.

[26] Several other factors affecting the sex ratio were also likely changing over time including average son preference in the population, the composition of births across women of different son preference, the costs and risks of abortion, and child mortality.

analysis, I relax this assumption to highlight the dual effects that progressive forces have on the sex ratio. Specifically, I examine the effect of female education. Female education is hypothesized to promote progressive attitudes and influence both son preference and desired fertility.

Previous work has found that maternal education reduces stated son preference (Pande and Astone, 2007). However, even if female education leads to a desire for fewer sons at any given family size, there is an offsetting effect. Maternal education is also associated with a decline in fertility (Dreze and Murthi, 2001; Osili and Long, 2008; Duflo, Dupas, and Kremer, 2012). As this paper has argued, a smaller family size leads to a higher desired sex ratio, all else equal. Thus, the net effect of female education on the desired sex ratio is theoretically ambiguous.

Table 6 examines these component and combined effects of education on the desired sex ratio, focusing on the subsample of mothers. The key independent variables are dummy variables for the respondent's level of education. Interpreting the coefficients as causal effects of education requires that education be uncorrelated with unobserved factors that affect fertility outcomes. To help reduce the likelihood of omitted variable bias, the regressions control for the husband's level of education and household wealth. For household wealth, I use principal component analysis to construct the first principal component of household asset ownership and dwelling characteristic variables, and control for a cubic polynomial in this variable.[27]

Column 1 of Table 6 examines the effect of female education on the percent sons desired at the randomly-specified family size, the unbiased measure of son preference. The categories for educational attainment are completing grade 10 or higher, completing grade 8, and completing primary school with less than primary as the omitted category. The negative coefficients indicate that education reduces women's son preference. Column 5 is similar to column 1 but, for parsimony, uses only a dummy for having finished grade 8 or higher (which 40% of respondents have done). Completing grade 8 reduces the desired proportion sons by 4 percentage points.

Columns 2 and 6 examine the effect of education on desired family size. (In contrast to most previous studies on education and fertility, the outcome here is desired family size for the respondent's child rather than herself.) The outcome is based on the standard

[27]The variables included in the principal component analysis are listed in the notes to Table 6. The results are nearly identical when instead separate control variables for each of the 36 wealth variables are included.

DHS-type fertility preference question which elicits the ideal family size. More education is associated with a smaller desired family size.

Columns 3 and 7 recast the negative effect of education on desired family size in terms of the corresponding increase in the desired percent sons. The dependent variable is the sample average percent sons desired corresponding to the respondent's desired family size. For example, if the respondent's desired family size is 1, then the dependent variable is 0.849, which is the average percent sons desired among those randomly assigned family size of 1, as reported in Table 4. (One cannot simply use the respondent's answer on the desired percent sons because each respondent was only asked about one randomly assigned fertility level, which might not correspond to her desired family size.) The negative effect of having a grade 8 education on the desired family size is equivalent to an increase in the desired proportion of boys of 2 percentage points (column 7).

Columns 1 to 3 and columns 5 to 7 show the two offsetting effects: Education reduces son preference at any given family size, which should decrease the desired sex ratio, but it also decreases desired family size, which increases the desired sex ratio. The net effect is shown in columns 4 and 8, where the outcome is the percent sons desired at the desired family size using the standard DHS-type question. The point estimate in column 8 suggests a small positive net effect on the desired proportion sons, i.e., a more male-skewed desired sex ratio, but the coefficient is small in magnitude and statistically indistinguishable from zero. Here, the two opposing effects almost exactly offset each other. This null result is noteworthy: A progressive force like female education need not improve the desired sex ratio.

The contrast between the negative effect of mother's education on son preference using the new measure introduced in this paper and the null effect of mother's education on the desired sex ratio also highlights an important methodological point: Using the sex ratio as a measure of son preference can lead to the wrong conclusions. For example, Filmer, Friedman, and Schady (2009) argue that female education does not reduce son preference in India. Their measure of son preference is son-biased fertility stopping behavior, but like the desired sex ratio, this measure conflates son preference and family size preference. If every family wants (say) one son, those that want fewer children are more likely to need to exceed their desired family size in order to have a son and thus are more likely to stop after a son is born. Son *preference* is distinct from its manifestation in the sex ratio or in differential stopping behavior.

4 Conclusion

This paper made both a methodological and substantive contribution related to son preference and skewed sex ratios. The methodological innovation of the paper was a survey question that elicits desired sex composition at a randomly determined level of total fertility specified to the respondent. This question could be a valuable addition to surveys that aim to measure son preference because it generates a measure of son preference that, unlike those based on the standard fertility preference questions, is not biased by the fact that individuals who desire a larger family size tend to have stronger son preference.

A first substantive contribution was to quantify how much declining fertility contributes to the worsening of the male-skewed sex ratio in India. The smaller the family size, the less likely a family is to have a son by chance. Thus, declining fertility is one force that is driving up the rate of sex-selective abortions and other behaviors that lead to missing women, and this paper shows that this channel is quantitatively important. The estimates suggest that falling fertility could explain half of the increase in the sex ratio that has occurred in India over the past thirty years.

Tracing out how the desired sex composition varies with family size is also revealing about the specific nature of son preference. Families appear to strongly want one son rather than always preferring having a son over a daughter. At a family size of one, the vast majority of respondents want that one child to be a son, but at a family size of two, having one daughter and one son is much more preferred to having two sons. Moreover, at a family size larger than three, respondents prefer to have more daughters than sons. The non-homothetic nature of son preference suggests that favoritism toward boys is not driven primarily by economic considerations such as low earning capacity of girls; the cultural importance of eldest sons looms large.

Another contribution was to show that factors that lead to more progressive attitudes, such as female education, need not improve the sex ratio. While female education reduces the desired sex ratio at any given family size, because it also reduce desired family size, it is not guaranteed to ameliorate the skewed sex ratio. The substantive finding here was that mother's education does reduce son preference but it does not reduce the desired sex ratio. Meanwhile, the methodological lesson was that using the sex ratio as the measure of son preference would have led to the incorrect conclusion that female education has no effect on son preference.

More broadly, the conceptual point emphasized in this paper is that son preference should be thought of as a vector of desired number of sons at different possible fertility levels. The manifestation of son preference in the sex ratio depends on the son preference vector, but also on the desired fertility level. Modernization might shift the son preference curve down, but, as importantly, it will cause movement along the curve by reducing desired fertility. The joint evolution of son preference and desired total fertility determines how the sex ratio will change over time and in response to different interventions.

One direction for future work is to elicit preferences not only for the "bliss point" of sex composition and total fertility, but also the relative disutility of different deviations. For example, at a family size of two, most individuals want one son and one daughter. However, a reasonable guess is that these families would be more inclined to resort to a sex-selective abortion to achieve this preferred sex composition if they naturally had two daughters than if they had two sons; having zero sons is much more undesirable to them than having zero daughters. Similarly, some families will prefer to try again for a son and use stopping rules rather than sex-selective abortions. Quantifying the asymmetry in how much individuals dislike having more boys versus more girls than their ideal sex composition and how much they dislike going beyond their desired family size would help further characterize son preference and family size preference and how the skewed sex ratio and total fertility will evolve in the future.

References

Anukriti, S. (2013): "The Fertility-Sex Ratio Trade-off: Unintended Consequences of Financial Incentives," working paper, Boston College.

Arnold, F., et al. (1998): "Son Preference, the Family-Building Process and Child Mortality in India," Population Studies, 52, 301–315.

Basu, A. M. (1999): "Fertility Decline and Increasing Gender Imbalance in India, Including a Possible South Indian Turnaround," Development and Change, 30(2), 237–263.

Basu, D., and R. De Jong (2010): "Son Targeting Fertility Behavior: Some Consequences and Determinants," Demography, 47(2), 521–536.

Bhat, P. M., and A. Zavier (2003): "Fertility Decline and Gender Bias in Northern India," Demography, 40(4), 637–657.

Campbell, E. K., and P. G. Campbell (1997): "Family Size and Sex Preferences and Eventual Fertility in Botswana," Journal of Biosocial Science, 29(2), 191–204.

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

Das, N. (1987): "Sex Preference and Fertility Behavior: A Study of Recent Indian Data," Demography, 24(4), 517–530.

Das Gupta, M. (1987): "Selective Discrimination Against Female Children in Rural Punjab, India," Population and Development Review, 13(1), 77–100.

Das Gupta, M., and P. M. Bhat (1997): "Fertility Decline and Increased Manifestation of Sex Bias in India," Population Studies, 51, 307–315.

Dreze, J., and M. Murthi (2001): "Fertility, Education, and Development: Evidence from India," Population and Development Review, 27(1), 33–63.

Duflo, E., P. Dupas, and M. Kremer (2012): "Education, HIV, and Early Fertility: Experimental Evidence from Kenya," MIT working paper.

Dyson, T., and M. Moore (1983): "On Kinship Structure, Female Autonomy, and Demographic Behavior in India," Population and Development Review, 9(1), 35–60.

Ebenstein, A. (2010): "The "Missing Girls" of China and the Unintended Consequences of the One Child Policy," Journal of Human Resources, 45(1), 87–115.

Filmer, D., J. Friedman, and N. Schady (2009): "Development, Modernization, and Childbearing: The Role of Family Sex Composition," The World Bank Economic Review, 23(3), 371–398.

Garenne, M. (2011): "Sex Ratios at Birth in African Populations: A Review of Survey Data," Human Biology, 76(4).

Goverment of Rwanda (2009): "National Population Projections," .

Guilmoto, C. Z. (2009): "The Sex Ratio Transition in Asia," Population and Development Review, 35(3), 519–549.

Jensen, R. (2003): "Equal Treatment, Unequal Outcomes? Generating Sex Inequality Through Fertility Behavior," Mimeo, Harvard University.

Kishor, S., and K. Gupta (2009): "Gender Equality and Womens Empowerment in India. National Family Health Survey (NFHS-3), India, 2005-06," Report, International Institute for Population Sciences, Mumbai, India.

Osili, U. O., and B. T. Long (2008): "Does Female Schooling Reduce Fertility? Evidence from Nigeria," Journal of Development Economics, 87(1), 57–75.

Pande, R. P., and N. M. Astone (2007): "Explaining Son Preference in Rural India: The Independent Role of Structural Versus Individual Factors," Population Research and Policy Review, 26(1), 1–29.

Park, C. B., and N.-H. Cho (1995): "Consequences of Son Preference in a Low-Fertility Society: Imbalance of the Sex Ratio at Birth in Korea," Population and Development Review, pp. 59–84.

Portner, C. C. (2014): "Sex Selective Abortions, Fertility, and Birth Spacing," mimeo, Seattle University.

Republic of Kenya (2009): "2009 Population and Housing Census," .

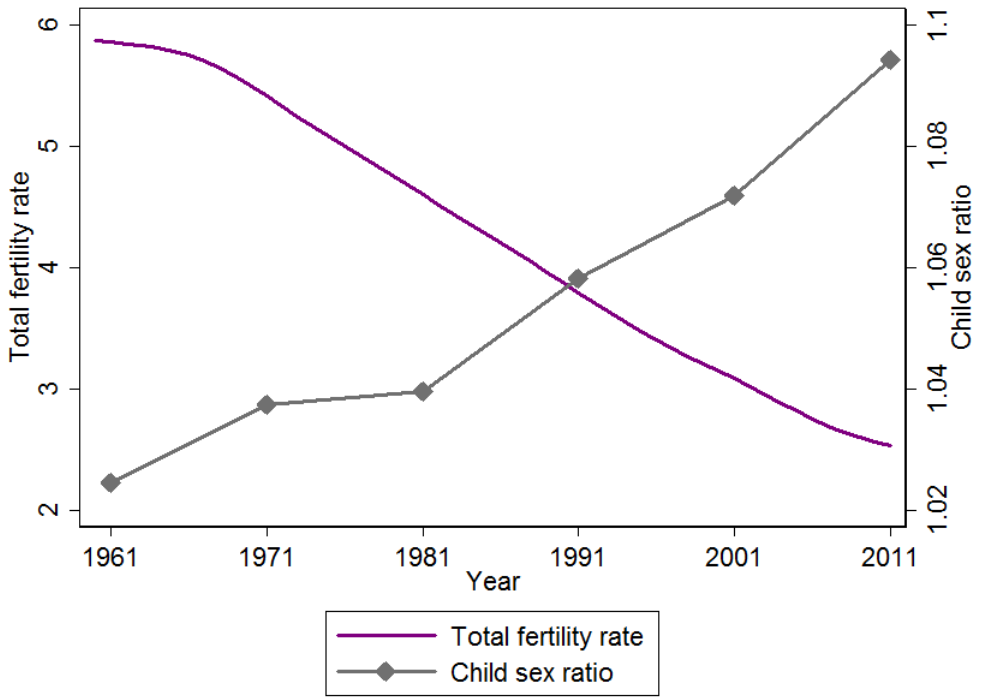
Rosenzweig, M. R., and K. I. Wolpin (1993): "Maternal Expectations and Ex Post Rationalizations: The Usefulness of Survey Information on the Wantedness of Children," Journal of Human Resources, 28(2), 205–229.

Sen, A. (1992): "Missing Women," British Medical Journal, 304, 586–87.

Westoff, C., and N. Ryder (1977): "The Predictive Validity of Reproductive Intentions," Demography, 14(4), 431–453.

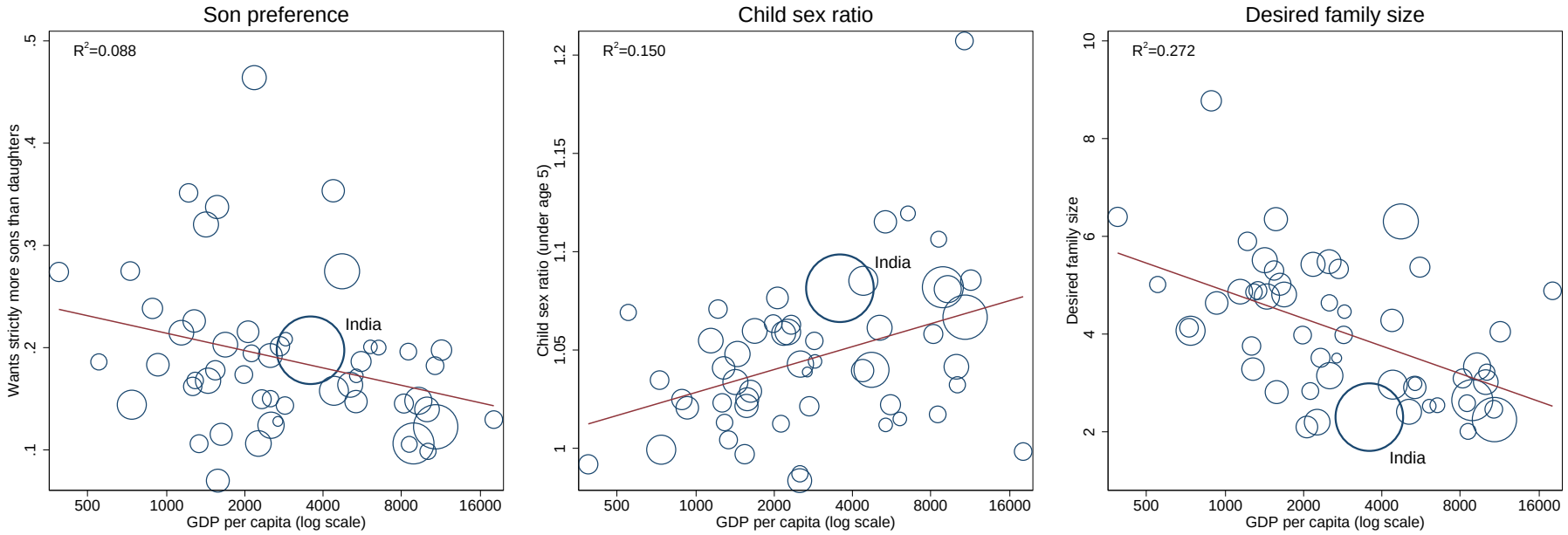
Yamaguchi, K. (1989): "A Formal Theory for Male-Preferring Stopping Rules of Childbearing: Sex Differences in Birth Order and in the Number of Siblings," Demography, 26(3), 451–465.

Figure 1: Trend in the total fertility rate and child sex ratio in India



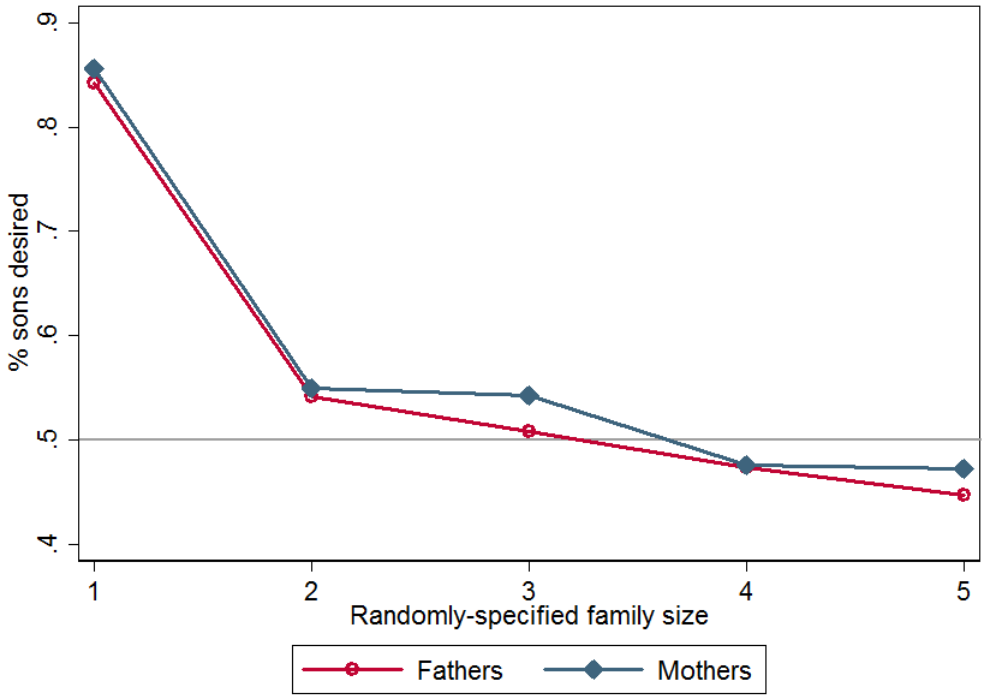
Notes: Data sources are Indian Census of Population (child sex ratio, every 10 years) and World Bank Indicators (total fertility rate, annual). Child sex ratio is the ratio of males to females among children 0 to 6 years old.

Figure 2: Cross-country relationships: GDP per capita versus son preference, child sex ratio, and desired family size



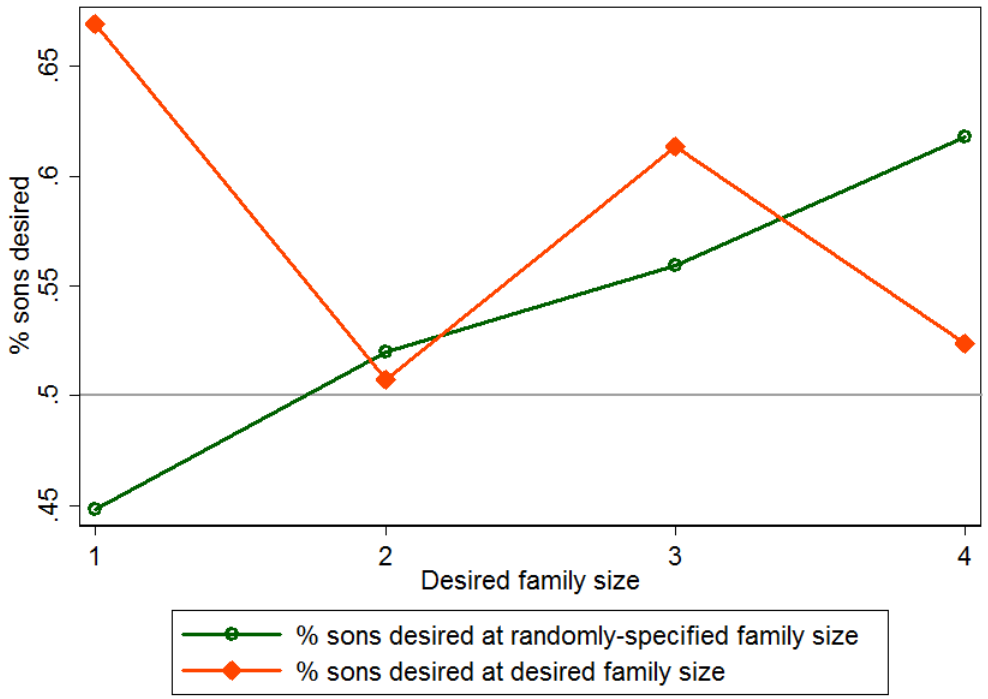
Notes: Data are from Demographic and Health Surveys (DHS) of ever-married women age 15 to 49. The sample comprises countries that conducted a DHS between 2005 and 2014; the most recent DHS per country is used. Each circle represents a country; the size of the circle is proportional to the DHS sample size. The R^2 is based on an unweighted regression where the dependent variable is the outcome on the vertical axis and the independent variable is log GDP per capita. The red line plots the fitted values from the regression. Each of the three univariate relationships is statistically significant at the 5% level. *Wants more sons than daughters* and *desired family size* are based on the DHS question that asks the respondent how many children she would ideally want and what their sex composition would be, if she could go back to the start of her childbearing. *Child sex ratio* is the ratio of boys to girls among respondents' currently-alive children age 5 and under. GDP per capita is the World Bank's PPP-adjusted value in US dollars for the survey year.

Figure 3: Percent sons desired by randomly-specified family size



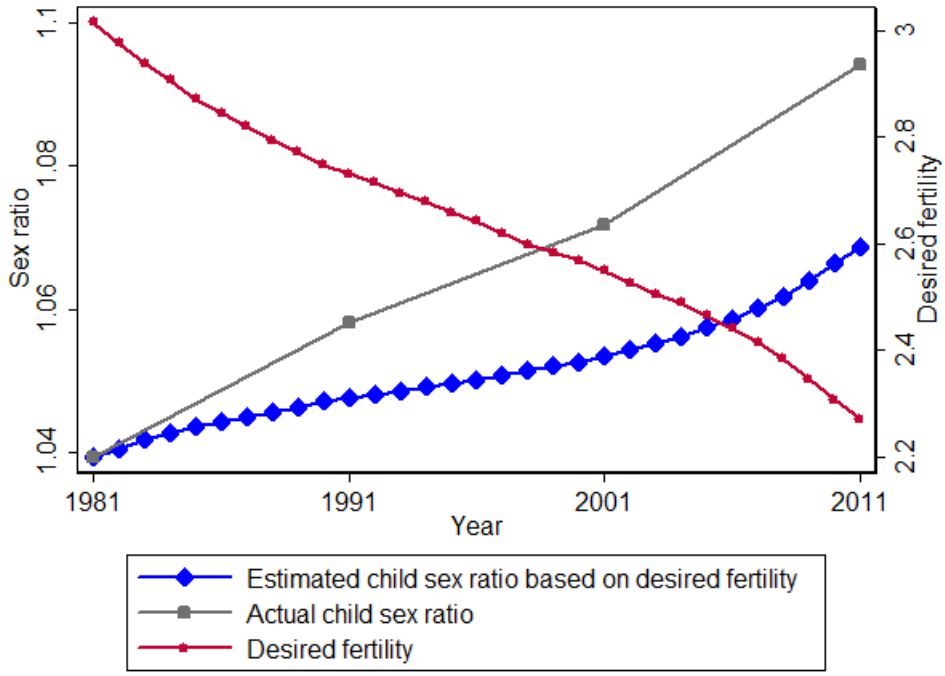
Notes: The height of each data point is the average percent sons desired among respondents asked about the family size indicated on the horizontal axis. Respondents were randomly assigned a hypothetical family size of 1, 2, 3, 4 or 5.

Figure 4: Percent sons desired by randomly-specified versus desired family size



Notes: The green line (circles) shows how son preference, measured using an exogenous fertility level, varies across respondents with different desired fertility levels. The orange line (diamonds) shows how the traditional measure of son preference based on DHS-type fertility questions varies across respondents with different desired fertility levels.

Figure 5: How much of the time trend in the sex ratio can fertility decline explain?



Notes: The gray line (squares) is the child sex ratio from Indian Census of Population data. The blue line (diamonds) projects how the sex ratio changed after 1981 due to declining desired fertility by combining (a) the actual decline in desired fertility over the time period and (b) the survey data described in this paper on how the desired sex ratio varies with the fertility level. The red line (circles) plots the average desired fertility level among women age 25 to 34 in the given year, based on responses in the National Family Health Survey Rounds 1 to 3.

Table 1: Descriptive statistics

	Mothers	Fathers
Age	34.975	40.495
	[5.611]	[6.732]
Illiterate	0.349	0.137
	[0.477]	[0.344]
Finished primary	0.312	0.248
	[0.463]	[0.432]
Finished Class 8	0.180	0.229
	[0.384]	[0.420]
Finished Class 10	0.124	0.266
	[0.329]	[0.442]
Finished Class 12+	0.035	0.120
	[0.184]	[0.325]
Hindu	0.942	0.953
	[0.234]	[0.212]
Muslim	0.055	0.043
	[0.228]	[0.203]
Scheduled caste	0.191	0.168
	[0.393]	[0.374]
Scheduled tribe	0.010	0.010
	[0.099]	[0.100]
Number of children	3.578	3.514
	[1.304]	[1.266]
Percent sons among children	0.542	0.540
	[0.326]	[0.329]
Surveyed student is female	0.563	0.543
	[0.496]	[0.498]
Surveyed student is grade 6	0.537	0.547
	[0.499]	[0.498]

Notes: Sample comprises 3215 mothers and 2460 fathers. The table reports subsample means with standard deviations in brackets.

Table 2: Desired number of sons by randomly-specified family size: Sample means

	Randomly-specified family size				
	1	2	3	4	5
Desired # of sons at randomly-specified family size					
0	0.151	0.031	0.020	0.014	0.028
1	0.849	0.846	0.405	0.176	0.153
2		0.124	0.546	0.726	0.350
3			0.029	0.063	0.436
4				0.021	0.018
5					0.015
Average % sons desired	0.849	0.546	0.528	0.475	0.462
Sex ratio desired	5.639	1.204	1.117	0.905	0.858
Observations	1102	1101	1182	1178	1112

Notes: Cells in the first 6 rows report the proportion of respondents who want a particular number of sons at the hypothetical randomly-specified family size. The remaining rows report the aggregated average percent sons, corresponding sex ratio and sample size for each randomly-specified family size.

Table 3: Percent sons desired by randomly-specified family size: Regression results

Dependent variable: % sons desired at randomly-specified family size				
	(1)	(2)	(3)	(4)
Randomly-specified family size	-0.085***		-0.085***	
	[0.003]		[0.003]	
Randomly-specified family size is 2		-0.303***		-0.303***
		[0.014]		[0.013]
Randomly-specified family size is 3		-0.322***		-0.321***
		[0.013]		[0.013]
Randomly-specified family size is 4		-0.374***		-0.374***
		[0.013]		[0.013]
Randomly-specified size is 5		-0.387***		-0.387***
		[0.013]		[0.013]
Controls for baseline characteristics	No	No	Yes	Yes
p-values for equality of coeffs				
Family size 1=Family size 2		0.000		0.000
Family size 2=Family size 3		0.034		0.038
Family size 3=Family size 4		0.000		0.000
Family size 4=Family size 5		0.070		0.083
Observations	5,675	5,675	5,675	5,675

Notes: Each observation is a parent respondent. In columns 3 and 4, all baseline characteristics listed in Table 1 (as well as indicator variables for missing values) are included as control variables. Standard errors, clustered by school, are in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 4: Percent sons desired by randomly-specified versus desired family size: Sample means

	Randomly-specified family size							
	1	2	3	4	5	Avg % sons at randomly-specified family size	Avg % sons at desired family size	Number of observations
Desired family size								
1	0.647	0.527	0.433	0.440	0.393	0.448	0.669	298
2	0.861	0.541	0.530	0.477	0.471	0.520	0.507	4460
3	0.828	0.620	0.609	0.512	0.490	0.560	0.613	284
4 or more	1.000	0.583	0.700	0.500	0.600	0.618	0.524	50

Notes: Each row is a subsample defined by the respondent's desired family size for his or her child. The first 5 columns report the average percent sons desired by respondents randomly assigned to different specified family sizes. The sixth column aggregates the data in the first 5 columns and reports the average percent sons at the randomly-specified family size. The seventh column reports the percent sons desired based on responses to the standard DHS-style fertility question where the respondent chooses the desired family size and sex composition simultaneously.

Table 5: Relationship between desired family size and son preference

	% sons desired at randomly-specified family size		% sons desired at desired family size	
	(1)	(2)	(3)	(4)
Desired family size	0.066***		-0.023***	
	[0.010]		[0.008]	
Desired family size is 2		0.093***		-0.162***
		[0.019]		[0.018]
Desired family size is 3		0.133***		-0.056***
		[0.024]		[0.020]
Desired family size is 4+		0.213***		-0.144***
		[0.040]		[0.021]
Constant	0.439***	0.481***	0.569***	0.669***
	[0.021]	[0.019]	[0.018]	[0.019]
p-values for equality of coeffs				
Family size 1=Family size 2		0.000		0.000
Family size 2=Family size 3		0.011		0.000
Family size 3=Family size 4+		0.041		0.000
Observations	5,092	5,092	5,092	5,092

Notes: Standard errors, clustered by school, are in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 6: Relationship between maternal education and fertility preferences (Sample of mothers)

	% sons desired at randomly-specified family size (1)	Desired family size (2)	Imputed % sons desired at desired family size (3)	Actual % sons desired at desired family size (4)	% sons desired at randomly-specified family size (5)	Desired family size (6)	Imputed % sons desired at desired family size (7)	Actual % sons desired at desired family size (8)
Mother finished primary	-0.015	-0.058***	0.002	-0.007				
	[0.010]	[0.018]	[0.003]	[0.005]				
Mother finished Class 8	-0.039***	-0.117***	0.013***	0.006				
	[0.013]	[0.022]	[0.004]	[0.008]				
Mother finished Class 10+	-0.056***	-0.163***	0.029***	-0.008				
	[0.013]	[0.027]	[0.006]	[0.007]				
Mother finished Class 8+					-0.039***	-0.108***	0.020***	0.003
					[0.010]	[0.017]	[0.004]	[0.006]
Observations	2,883	2,597	2,597	2,597	2,883	2,597	2,597	2,597

Notes: Standard errors, clustered by school, are in brackets. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. All regressions control for dummy variables for whether the father has completed primary school, grade 8, or grade 10+, and a cubic polynomial in the first principal component of wealth. The variables used in the principal component analysis include student responses about whether the family owns their house, the house is "pukka" (high quality/made of permanent materials), the house has electricity, the house has a flush toilet, the house has a non-flush toilet, the house has tap water, the house has a separate kitchen, the child had 2 meals each day in the last 7 days, the family owns a cell phone, the family owns a TV, the family owns a radio, the family gets newspapers daily, the family gets magazines, and the family owns a computer. The principal component analysis also includes parent answers about whether the respondent belongs to a scheduled caste, respondent belongs to a scheduled tribe, household owns land, and household owns each of the following items: radio/tape recorder, computer/laptop, television, bicycle, motorcycle/scooter, car/truck/tractor/other 4-wheel vehicle, refrigerator, fan, air conditioner, kerosene or gas stove, kerosene lamp, landline telephone, cell phone, sewing machine, thresher, water pump, bullock cart, livestock, and washing machine. Missing values are replaced with the sample mean and flags for imputed missing values are included in the principal component analysis.

Appendix Table 1: Wealth, total fertility, and son preference in north India

	Haryana	Other Hindi-belt states
Dwelling has piped water	0.40	0.29
	[0.49]	[0.46]
Dwelling has electricity	0.93	0.72
	[0.26]	[0.45]
Wealth index	0.23	-0.04
	[0.88]	[1.08]
Number of births	2.24	2.42
	[1.95]	[2.32]
Desired family size	2.18	2.40
	[0.81]	[0.89]
Wants strictly more sons than daughters	0.22	0.26
	[0.42]	[0.44]
Years of schooling (boys)	4.13	3.40
	[2.54]	[2.56]
Years of schooling (girls)	3.92	3.15
	[2.66]	[2.59]

Notes: Respondents are ever-married women age 15 to 49 from the National Family Health Survey, round 3. The Hindi-belt states, excluding Haryana, are Bihar, Chhattisgarh, Delhi, Himachal Pradesh, Jharkhand, Madhya Pradesh, Rajasthan, Uttar Pradesh, and Uttaranchal. The sample size comprises 2790 respondents for Haryana and 42,608 for the other Hindi-belt states; schooling is based on 2464 children age 7-14 in Haryana and 37,949 children in the other Hindi-belt states. The wealth index is constructed by the NFHS using principal component analysis of several asset ownership and dwelling characteristics variables, and is normalized to be mean 0 and standard deviation 1 for the India-wide sample.

Appendix Table 2: Randomization balance check

	Randomly-specified family size					Equality of means p-value
	1	2	3	4	5	
Age	37.666	37.512	37.445	37.186	37.201	0.392
	[6.601]	[6.813]	[6.946]	[6.457]	[6.733]	
Illiterate	0.264	0.254	0.247	0.262	0.260	0.843
	[0.441]	[0.436]	[0.431]	[0.440]	[0.439]	
Finished primary	0.284	0.294	0.280	0.268	0.295	0.625
	[0.451]	[0.456]	[0.449]	[0.443]	[0.456]	
Finished Class 8	0.187	0.187	0.206	0.213	0.213	0.377
	[0.390]	[0.390]	[0.404]	[0.410]	[0.410]	
Finished Class 10	0.187	0.191	0.194	0.188	0.166	0.637
	[0.390]	[0.393]	[0.396]	[0.390]	[0.372]	
Finished Class 12+	0.077	0.074	0.073	0.070	0.066	0.847
	[0.266]	[0.262]	[0.261]	[0.255]	[0.248]	
Hindu	0.942	0.952	0.950	0.952	0.936	0.365
	[0.234]	[0.214]	[0.218]	[0.213]	[0.245]	
Muslim	0.054	0.044	0.048	0.044	0.060	0.326
	[0.225]	[0.204]	[0.214]	[0.205]	[0.238]	
Scheduled caste	0.175	0.174	0.184	0.176	0.191	0.884
	[0.380]	[0.380]	[0.388]	[0.381]	[0.393]	
Scheduled tribe	0.007	0.007	0.018	0.009	0.008	0.316
	[0.083]	[0.083]	[0.134]	[0.094]	[0.091]	
Number of children	3.531	3.608	3.561	3.535	3.516	0.857
	[1.297]	[1.303]	[1.265]	[1.332]	[1.238]	
Percent sons among children	0.554	0.513	0.541	0.544	0.552	0.083
	[0.320]	[0.325]	[0.329]	[0.332]	[0.327]	
Surveyed student is female	0.535	0.573	0.556	0.560	0.545	0.519
	[0.499]	[0.495]	[0.497]	[0.497]	[0.498]	
Surveyed student is grade 6	0.544	0.542	0.546	0.542	0.531	0.839
	[0.498]	[0.498]	[0.498]	[0.498]	[0.499]	

Notes: Standard deviations in brackets. The statistical test for joint equality of means across subsamples allows for clustering within a school.