Federal Reserve Bank of Chicago

Prenatal Sex Selection and Girls' Well-Being: Evidence from India

Luojia Hu and Analía Schlosser

REVISED
October 2014
WP 2010-11

Prenatal Sex Selection and Girls' Well-Being: Evidence from India

Luojia Hu
Economic Research
Federal Reserve Bank of Chicago
230 S. LaSalle Street
Chicago, IL 60604
Phone: (312) 322 8122
e-mail: lhu@frbchi.org

Analia Schlosser
The Eitan Berglas School of Economics
Tel Aviv University
P.O.B. 39040
Ramat Aviv, Tel Aviv, 69978
ISRAEL
Phone:+972-54-4902414
e-mail: analias@post.tau.ac.il

Prenatal Sex Selection and Girls' Well-Being: Evidence from India[1]

Luojia Hu
Federal Reserve Bank of Chicago

Analía Schlosser
Tel Aviv University

October , 2014

ABSTRACT

We study the impacts of prenatal sex selection on girls' well-being in India. We show that high sex ratios at birth reflect the practice of prenatal sex selection and apply a triple difference strategy to examine whether changes in health outcomes of girls relative to boys within states and over time are systematically associated with changes in sex-ratios at birth. We find that an increase in prenatal sex selection leads to a reduction in girls' malnutrition, in particular, underweight and wasting. We further explore various underlying channels linking between prenatal sex selection and girls' outcomes.

Keywords: Son preference, prenatal sex selection, ultrasound, sex ratio at birth, gender discrimination, child health.

JEL codes: J13, J16, I1, O12

[1] We thank Angus Deaton, Prashant Bharadwaj, Seema Jayachandran, Elaine Liu, Nancy Qian, Steven Pischke, two anonymous referees, and seminar participants at the Federal Reserve Bank of Chicago, Princeton labor lunch, University of Lausanne, University of Zurich, Ben Gurion University, Bar Ilan University, Wellesley College, UIUC, 2011 SOLE meetings, CESifo Venice Workshop, CEMFI, Norwegian School of Economics, and NBER Summer Institute for helpful comments. We also thank Yotam Shem Tov for excellent research assistance. The opinions expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System. Schlosser thanks the Pinhas Sapir Center for research support.

I. INTRODUCTION

Son preference in India and other Asian countries has been documented extensively. A large number of studies have shown that parental preference for boys is manifested in gender differences in the intrahousehold allocation of resources and medical care, which result in gender differentials in child outcomes such as nutrition, morbidity and mortality and, in extreme cases, female infanticide (see, for example, Arnold et al., 1998; Deaton, 1997; George, 1997; Miller, 1981; Kishor, 1993; Rose, 1999; Rosenzweig and Schultz, 1982; Pande, 2003; and Barcellos et al., 2010). The diffusion of prenatal sex determination technologies since the 1980s has provided parents with an alternative way to achieve the preferred sex composition of children. Indeed, starting from the late 1980s, there has been a steep increase in male to female ratios at birth, which has been attributed to the increasing practice of sex selective abortion. A growing number of studies have examined the factors affecting the prevalence of prenatal sex selection and have assessed its effect on sex ratios at birth.[1] Little is known, however, about the impact of this phenomenon on the well-being of the girls who are born.

Prenatal sex selection may affect girls' outcomes through different channels. First, girls might be more likely to be born into households that have weaker son preference so that they would be more likely to be wanted.[2] Second, with access to prenatal sex selection, parents may be less likely to rely on fertility stopping rules to achieve their desired sex composition of children so that girls might be born into smaller families and receive a larger share of household inputs. Third, girls' outcomes might change if the characteristics and living environment of households that do not practice prenatal sex selection are different from those that practice it. Finally, prenatal sex selection may affect girls' outcomes if an increase in the scarcity of girls changes women's value in labor and marriage markets subsequently affecting parental incentives to invest in girls.

In this paper, we study the impact of prenatal sex selection on the well-being of girls in India and investigate some of the different channels through which prenatal sex selection may affect girls' outcomes. India is an interesting environment for studying the impact of prenatal sex selection due to its large regional variations in parental preference for sons and the striking differences in the use of prenatal sex selection technologies over time and across regions. Using a comprehensive data set on child outcomes and household characteristics, we apply a *triple-*

[1] We review this literature in the next section.

[2] This is a similar rationale to that proposed by Donohue and Levitt (2001) and Ananat et al. (2009) in the context of abortion legalization in the US.

differences approach to examine whether changes in the nutritional status and mortality rates of girls relative to boys within states and over time are systematically associated with changes in the prevalence of prenatal sex selection. In the absence of a direct measure of parental access to prenatal sex selection, we use the ratio of males to females at birth (MFR) in the child's state and birth cohort and provide evidence showing that increases in MFR at birth reflect the practice of prenatal sex selection starting from the late 1980s. We then analyze whether changes in the outcomes of girls within states and over time in comparison with boys can be attributed to changes in MFR at birth. We then turn to explore the various channels that might link between prenatal sex selection and girls' outcomes by analyzing whether changes in MFR within states and over time appear to be associated with changes in household characteristics, fertility, breastfeeding practices, and son preference.

Our paper is related to a limited number of recent studies, most of them conducted concurrently with our own, that examine the effects of prenatal sex selection on girls' outcomes. Shepherd (2008) compares between families with high versus low likelihood of using prenatal sex selection in India and finds inconclusive evidence for a link between prenatal sex selection and a decline in female child's mortality. Lin et al. (2009) exploit variation in the use of prenatal sex selection in Taiwan over time and birth order generated by the legalization of abortion and find a positive association between prenatal sex selection and female survival. Almond et al. (2010) use variation in the diffusion of ultrasound technology across regions in China and find that prenatal sex determination is associated with an increase in girls' neonatal mortality but do not find any changes in postnatal investments in girls as manifested in breastfeeding duration and vaccination rates.[3]

Our paper adds to this limited number of studies in several aspects. First, compared to some other countries, India's high degree of heterogeneity in female discrimination and the differences in the practice of prenatal sex selection across regions makes it an optimal case for studying the effects of prenatal sex selection. This is also due to India's large variation in family size across households, which provides an opportunity to examine the extent and impact of prenatal sex selection across different parities and study the effects of prenatal sex selection on fertility decisions. Second, we conduct a more comprehensive analysis on the impacts of prenatal sex

[3] Other recent related studies are Portner (2010), which looks at the determinants of sex-selective abortion in India though not their effect on outcomes, and Bhardadwaj and Nelson (2010) which examine gender differences in prenatal investment in countries with strong son preference, including India, although they do not estimate the impact of sex-selective abortion.

selection by examining both mortality and nutritional outcomes. Our analysis of nutritional status provides important findings on intermediate outcomes that are prevalent in a high proportion of Indian children and have vital life-long consequences for human capital development and well-being (see, e.g. Currie, 2009 and Case and Paxson, 2010). Third, we go a step further and explore different channels through which prenatal sex selection affects child outcomes such as selection, family size, preferences, and breastfeeding duration. Finally, our empirical approach, which is based on triple-differences models, allows us to control for several confounding factors such as regional variation in son preference and differential trends across regions, providing a powerful way to reject alternative explanations for the observed results.

Our results show that an increase in the practice of prenatal sex selection leads to a reduction in the prevalence of malnutrition (in particular, underweight and wasting) among girls. Our results are robust to the inclusion of state-year fixed effects suggesting that our findings are unlikely to be driven by state specific time varying unobserved factors associated with changes in the incidence of prenatal sex-selection. Our results are also robust to several checks that assess the likelihood of possible biases due to state-specific differential trends by gender associated with prenatal sex selection.

Additional results show no evidence that prenatal sex selection leads to a selection of girls into households with more resources since we do not observe a differential improvement in household socio-economic characteristics of girls. However, we find that girls are more likely to be born in families with weaker son preferences. We also find evidence of a larger reduction in family size and an increase in breastfeeding duration for girls. On the other hand, prenatal sex selection does not appear to be associated with a reduction in girls' mortality or a regional decline in reported preferences for sons.

The rest of the paper is organized as follows: In the next section, we review the literature and describe the institutional background of unbalanced sex ratios and prenatal sex selection in India. Section III discusses the conceptual framework for analyzing the effects of prenatal sex selection on child outcomes. Section IV describes the data. Section V lays out the empirical strategy and Section VI presents the results. Finally, Section VII concludes.

II. BACKGROUND AND INSTITUTIONAL FRAMEWORK

Imbalanced sex ratios have been documented in India as early as the 19th century and throughout the 20th century (see Visaria, 1971; Miller, 1981, 1984; Dyson and Moore, 1983; and Sen, 1992, 2003 who introduced the concept of "Missing Women"). Until the late 1970s, gender imbalances were mostly manifested at older ages and not particularly at birth and were attributed to excess female mortality due to maltreatment and neglect and in extreme cases to female infanticide (see, for example, Dreze and Sen, 1997 and Das Gupta, 1987).[4]

Several studies have documented sharp increases in male to female ratios at birth since the late 1980s, especially in northern and western states, which are regions historically known for strong son preference and gender discrimination (see, for example, Das Gupta and Bhat, 1997; Arnold et al., 2002; Bhat, 2002; Bhaskar and Gupta, 2007; and Retherford and Roy, 2003). These studies have shown that the increase in sex ratios at birth is concurrent with the spread of prenatal sex determination technologies, thus suggesting that sex selective abortion is likely to be a major contributing factor.[5] Bhalotra and Cochrane (2010) estimate that about 0.48 million girls per year were selectively aborted in India during 1995-2005, which represents 6.2 percent of all potential female births. Estimates for northern and western regions are considerably higher. For example, Kulkarni (2007) estimates that out of 168,997 expected female births in Punjab in 2001, 19 percent (31,648) went missing.

Sex selective abortion requires two steps: sex determination of the fetus and an abortion. Abortion was legalized in India under the Medical Termination of Pregnancy Act (MTP) in 1972. Despite its legalization, most abortions are still practiced in unofficial and non-regulated facilities.[6] Access to abortion services is not difficult in India, even in the remotest areas of the country (Duggal, 2004). Its costs vary by region, type of facility, method, and gestation period ranging from US$4.5 to US$16.5 (Ravindran, 2002; Sundar, 2003).

Sex determination during pregnancy was first made possible in the late 1970s by the use of amniocentesis (Jefferey et al., 1984) and became more accessible with the introduction of

[4] Other possible explanations attributed to the lower number of females include census underenumeration of girls or different patterns of age misreporting by sex. While these factors might explain part of the gender imbalances, Visaria (1971), Miller (1981, 1984), and other researchers show that the main cause for imbalanced sex ratios until the late 1970s was sex differentials in mortality.

[5] Changes in the enumeration of girls or misreporting of age are unlikely to explain this upward trend since similar increases in sex ratios at birth are also observed among Indian populations living in the US, Canada, and the UK where birth registration is nearly complete and accurate (see Abrevaya, 2009; Almond and Edlund, 2008; Almond et al., 2009; and Dubuc and Coleman, 2007). We provide in the next section further evidence which suggests that prenatal sex selection is the main cause for the upward trend in sex ratios at birth.

[6] It is difficult to obtain precise figures on abortion rates. Using indirect estimation techniques based on 1991 birth rates and population counts, Chhabra and Nuna (1993) assess that about 6.7 million abortions are carried out every year, with the majority of them taking place in informal facilities. A similar estimate (6.4 million) is provided by Duggal and Ramachadran (2004).

ultrasound technologies in the 1980s. In subsequent years, a continued decline in desired fertility coupled with a slower decline in the total number of desired sons increased the pressure to have sons at lower parities, thus raising the demand for prenatal sex selection (Das Gupta and Bhat, 1997). At the same time, economic development and trade liberalization accelerated the supply of prenatal sex determination technologies. Following the policy reforms of the 1990s, multinationals such as General Electric started manufacturing ultrasound machines in India improving access to prenatal sex determination and increasing the practice of prenatal sex selection throughout the country (Geroge, 2006). The diffusion process took place from urban to rural areas and from households of high socioeconomic status to those of low socioeconomic status (Khanna, 1997).[7]

In an attempt to eliminate the practice of sex selective abortion and in response to increasing public pressure from several NGOs and women's organizations, the government of India passed the Prenatal Diagnostic Techniques Regulations and Misuse Act (PNDT Act) in 1994, thus making it illegal to use ultrasound or amniocentesis in order to determine the sex of a fetus. However, this legislation proved to be ineffective and the practice of sex selective abortion continued to spread (see e.g., George, 2002 and Kishwar, 1995). In subsequent years, several private ultrasound clinics opened, mobile clinics and portable ultrasound machines made prenatal sex selection accessible in small towns and rural areas (Krugman, 1998), and most of the sex selective abortions moved to underground clinics (Everett, 1998).

The publication of sex ratio figures from the 2001 Census revealed a further increase in MFR at ages 0-6, which led to increased public pressure on the government to enforce and expand the legal power of the PNDT Act.[8] Recent reports indicate some improvement in the enforcement of the Act. However, sex selective abortion is still being practiced extensively and enforcement of the law appears to be difficult, if not impossible, to achieve (see e.g., Subramanian and Selvaraj, 2009; Portner, 2010).

[7] Ultrasound costs vary widely across regions and between formal and informal facilities. Arnold et al. (2002) provide an average estimate of US$10-20.

[8] In 2001, the Centre for Enquiry into Health (CEHAT) and the Mahila Sarvangeen Utkarsh Mandal (MASUM) filed a public interest litigation against the Union of India and all the state governments for the non-implementation of the PNDT Act and for the inclusion of all emerging technologies that could be used for sex selective abortion. As a result, the Indian Government decreed in 2001 that manufacturers of ultrasound equipment could sell the machines only to registered clinics, which must maintain strict records of their use. For a description of the PNDT Act and its enforcement, see Mallik (2003) and George (2002).

III. CONCEPTUAL FRAMEWORK

There are various channels through which parental preference for sons can affect girls' outcomes in the absence of a technology that allows them to realize their preferences for sex composition. First, parents can allocate more resources to children that provide them with higher utility (i.e. boys) as long as boys and girls are substitutes in parents' utility function (Rosenzweig and Schultz, 1982) or when capital markets are imperfect (Becker and Tomes, 1979 and 1986).[9]

Parents can also attempt to adjust excess fertility and undesired sex composition of children through selective allocation of resources that results in the mortality of the less-desired children (Simmons et al., 1982). Alternatively, if parents continue to bear children until they achieve their desired number of sons, girls will have, on average, a larger number of siblings than boys (Yamaguchi, 1989). Therefore, even if parents treat boys and girls equally within a household, girls can be expected to grow up in larger families with fewer resources per child (Jensen, 2003). In addition, parental preference for boys is likely to affect birth spacing (Ben-Porath and Welch, 1976) and duration of breastfeeding (Jayachandran and Kuziemko, 2009) since parents may wish to have a child sooner following the birth of a daughter, especially if they are approaching the end of the mother's fecund years. Both factors are known to affect children's health.

Increased access to prenatal sex selection may improve the outcomes of girls who are born through various channels. First, girls may be more likely to be born into families that have a lower disutility from girls and therefore may suffer less from discrimination in resource allocation within the household. As suggested by Goodkind (1996), prenatal sex selection may substitute for postnatal sex discrimination. Second, prenatal sex selection allows parents to achieve their preferred number of boys and girls without the need of having additional children.[10]

[9] Differences in parents' utility from boys and girls might be rooted in taste differences or in differences in the costs and benefits of raising boys versus girls (Ben-Porath and Welch, 1976). Arnold et al. (1998) provide a review of the literature and highlight three main channels that affect parental preferences for sons. The first is economic utility due to differences in labor market productivity between boys and girls, security in the case of illness, support in old age and inheritance rules. A second channel is related to social utility, which provides families with more sons a higher status and additional sources of income in the form of dowry payments. A third channel is religious utility, which is derived from the performance by sons of important religious functions, especially upon the death of his parents. Girls, in contrast, are seen as an economic drain on family resources because of the dowry system and the high cost of weddings.

[10] On the other hand, the ability to choose a child's gender increases the expected utility of children and might therefore increase the likelihood of having additional children. Family size might also increase with access to prenatal sex selection if the net cost of raising boys is lower than the cost of raising girls (Ben-Porath and Welch, 1976) or if boys yield higher utility than girls and there is substitution between utility from children and ordinary consumption or between utility from boys and utility from girls. Therefore, the

Consequently, family size will not necessarily be larger for girls. Prenatal sex selection might also improve girls' outcomes if parents of girls internalize the potential increase in the value of women as they become scarcer in marriage and labor markets.[11] Although it can also be the case that as girls become scarcer, parents are more likely to want girls but do not need to invest so much in their quality. In addition, prenatal sex determination might have a negative impact on girls' outcomes if parents reduce prenatal investment in female fetuses in response to prenatal sex determination rather than having an abortion (see, for example, Bharadwaj and Nelson, 2010 and Almond et al., 2010). Finally, girls' outcomes might also change (improve or worsen) through selection if parents who practice prenatal sex selection have different characteristics than parents who do not. Taken together, these various channels imply that the sign of the overall impact of prenatal sex determination on the outcomes of girls who are born cannot be determined a priori and remains an empirical question.

IV. DATA

The data for our empirical analysis are taken from two major data sources. We first use population counts from the Indian census to examine long term changes in MFR. We then use microdata from the National Family Health Survey (NFHS) to link between changes in MFR and prenatal sex selection and study how changes in MFR affect girls' outcomes.

The NFHS is a large-scale, multi-round survey conducted in a representative sample of women and their families throughout India.[12] Three rounds of the survey were conducted in 1992-3, 1998-9 and 2005-6. Each round covered approximately 90,000 households which contained more than 500,000 individuals and was designed to provide state-level and national-level estimates. The survey includes detailed information on the demographic and socioeconomic background of the household members, as well as additional modules designed to investigate health, fertility and mortality. One major advantage of these data is that they record complete birth histories

impacts of prenatal sex selection on family size are ambiguous. Nevertheless, the gender gap in family size is expected to narrow or even reverse.

[11] This and the first channel imply that girls are becoming "more wanted". Although the first channel works through selection across households with different preferences, whereas this channel works through a general equilibrium effect changing preferences over time for all households, even those who do not practice prenatal sex selection.

[12]More precisely, it is a representative sample of women of reproductive age (e.g. ever married women at ages 13-49 in the first round, ever married women at ages 15-49 in the second and all women at ages 15-49 in the third round). The NFHS was designed along the lines of the Demographic and Health Survey (DHS) that has been conducted in many developing countries since the 1980s.

(including deaths and children living outside the household) of a large number of women over a long period, making it feasible to compute MFR at birth by state and cohort.[13]

We pooled the three survey rounds and selected only households with ever married mothers aged between 15 and 49 in order to obtain a consistent sample across the three rounds. **Table 1** reports summary statistics for boys and girls included in our main samples (columns 1, 2, 4, and 5) and differences by gender (columns 3 and 6). The sample reported in columns 1-3 (*nutrition sample*) includes the youngest two children aged less than three of ever-married women sampled in one of the three NFHS rounds who have valid anthropometric data.[14] There are 76,500 children (36,940 girls and 39,560 boys) who satisfy these criteria. The sample reported in columns 4-6 (*mortality sample*) includes all children of ever-married women born within the 10 years preceding each survey round. This adds up to 357,809 children (172,472 girls and 185,337 boys). Household characteristics reported in the table are used as control variables in the empirical analysis.

Most children (about 75 percent) in both samples live in rural areas. About half of the children have mothers with no formal education and about 30 percent have fathers with no formal education. Mothers' age at first birth is relatively low at 19, on average. At the time of the survey, mothers in the *nutrition sample* were 25 years old, on average, and they had an average of 3 children. Mothers in the *mortality sample* are 29 years old on average at the time of the survey and have an average of about 4 children.[15]

Columns 3 and 6 show that girls tend to be born into more disadvantaged families than boys. Their families have lower wealth levels, lower parental education, and a lower degree of exposure to mass media.[16] Another fact worth noting is that girls appear to have more siblings than boys in

[13] Note that MFR estimates reported in many studies are based on census data which record only children who are alive and live in the household at the time of the census. Therefore, they are affected by differential mortality and living arrangements.

[14] The first round of the NFHS collected anthropometric data for the youngest two children in the household who were under four years of age. The second round restricted the anthropometric data collection to the youngest two children in the household under the age of three, while the third round extended the data collection to all children in the household under five years of age. To be consistent across survey rounds, we restrict the sample to the youngest two children under three years of age. In practice, our results are insensitive to these restrictions.

[15] Note that the average number of children in our samples does not represent completed fertility as most women are still in their fertile years.

[16] The index for media exposure is defined by the sum of indicators for exposure to TV, radio and newspapers or magazines. Each indicator receives a value of one if the mother reported exposure of at least once a week or almost every day. The wealth index is a constructed index provided in the NFHS data. The

the mortality sample (columns 4-6), which is consistent with parental stopping rules in fertility behavior and son preference. There is no similar difference in the *nutrition sample* (columns 1-3), probably due to the fact that a large proportion of children in the sample comes from households with incomplete fertility. Nevertheless, we see in both samples that mothers of girls are more likely to report that they want to have another child. Differences in this variable are larger in the *nutrition sample* since mothers are more likely to have uncompleted fertility. We will return to the differences in household characteristics and family size in section VI.2 when we discuss possible mechanisms underlying our main results.

V. EMPIRICAL STRATEGY

The research question we seek to address in this paper is: what is the causal effect of prenatal sex selection on the outcomes of girls who are born? The ideal experiment we would like to study would be based on random assignment of access to ultrasound or any technology that facilitates prenatal sex selection. But this is not feasible. Moreover, we do not observe whether individual households practice sex selective abortion. Our empirical strategy is therefore based on using aggregate male-female ratios at birth (MFR), which are a consequence of households behavior, to proxy for the prevalence of prenatal sex selection and exploit the variation in MFR across states and over time to estimate the impacts on girls' outcomes while trying best to isolate other confounding factors.

In subsection V.1 we show how male-female ratios at birth (MFR) vary across time and states and demonstrate that increases in MFR at birth are a consequence of prenatal sex selection. In subsection V.2 we explain how we use this variation to study the impacts of prenatal sex selection on girls' outcomes.

V.1. Incidence of prenatal sex selection and male-female ratios at birth (MFR)

V.1.a. Long run evidence of state-level changes in MFR

As mentioned above, one potential limitation of this analysis is that we do not directly observe the practice of sex selective abortion. However, we do observe its consequences, primarily the abnormal sex ratio at birth. **Figure 1** shows state variation in MFR at age 0 as reported in census

index is based on household assets and housing characteristics and denotes the wealth quintile of the household relative to all households sampled in the same survey round.

records from 1961 to 2001.[17] Until the 1980s, MFR at birth did not exceed the normal ranges of 1.03-1.07 found in various large-scale studies (for example, Visaria, 1971 and Jacobsen et al., 1999). Increases in MFR at birth become evident after 1981 and, to a greater extent, after 1991, which overlaps with the introduction of ultrasound technology in India.

Interestingly, there are large variations in MFR (both in their levels and in their growth rates) even across those states that appear to have a strong preference for boys. To illustrate this point, **Appendix Table A1** reports indicators of son preference and selected characteristics by state based on tabulations from the first round of the NFHS and per capita net state domestic product in 1990.[18] States are grouped by region.

The largest increases in sex ratios at birth are found in the northern and western states, which are characterized by a strong degree of son preference. In Punjab, for example, while MFR was within the normal range between 1961 and 1981, it increased dramatically between 1981 and 2001 from 1.06 to 1.29. Punjab also appears to have strong son preference as manifested by the ideal sex ratio reported by mothers (1.48 or 1.47, depending on the measure) and the proportion of mothers who desire a larger number of sons than daughters (0.48).

While strong son preference is found in states with upward trends in MFR, we also observe that in several states in the northeast, central, and eastern regions with similar strong preferences for sons there was either no increase in MFR or only a mild one. In Madhya Pradesh, for example, mothers reported an ideal sex ratio of 1.55 but sex ratios at birth remained close to natural levels (MFR of 1.06 in 2001). Southern states are usually characterized by a low degree of son preference and stable sex ratios at birth.

This suggests that strong son preference cannot alone explain state variation in MFR since there are several states with strong son preference that have not shown any significant increase in MFR. Indeed, a factor that distinguishes between states with similar son preference but different incidence in prenatal sex-selection is economic development. Northern and western states are generally more economically developed than states in the northeast, central and

[17] MFR from census records number of children aged 0 to 11 months who are alive on census date. So, they also reflect gender differentials in infant mortality rates and living arrangements. In the next paragraphs we refer to MFR at age 0 as MFR at birth since, as we will show later, changes in mortality and living arrangements by gender had only a negligible contribution to the overall trends in this variable.

[18] For details about the construction of son preference indicators see Data Appendix.

eastern regions.[19] On the other hand, development and income levels in many of the northern and western states are comparable to those in southern states where sex ratios have remained balanced.

To more systematically examine the relationship between changes in MFR, income and son preferences, we estimated a set of state-level regression models of MFR on state income per capita and son preference. We use MFR data from the 1991 and 2001 census to capture long-term changes, and state income per capita from 1990 and son preferences from the first round of the NFHS survey (1992) as baseline levels of income and son preference. First, we regressed the state-level MFR on a post dummy for year 2001, baseline state income per capita (in thousand Rs.) and average state son preferences and their interactions with the post dummy.[20] In columns 1 and 2 of **Table 2** we report estimates from models that include income or son preferences separately. In column 3 we report estimates from a model that includes both variables. The results show that that the MFR was higher in states that were richer or had stronger son preferences and that MFR changes were also steeper in those states. Moreover, column 3 shows that associations between changes in MFR and income per capita and son preferences are larger while controlling for both factors at the same time. In column 4, we report estimates from a model that includes also interactions between son preference and income. The results show that the states that experience largest increase of MFR are those with both high income and strong son preference.

Overall, evidence above suggests that two factors which characterize states with an increasing MFR are a strong preference for boys and a higher degree of development and modernization (in combination). The fact that there are states with a strong preferences for boys (eastern region) and states with high levels of development (southern region) that have not exhibited significant increases in MFR provide us with a heterogeneous group of states that are comparable to states with increasing MFR across different dimensions.

[19] Appendix **Table A1** also presents some economic and demographic indicators by state for the years 1991-2. As can be seen, northern and western states have a higher wealth index, income per capita, share of households with electricity, and a higher degree of exposure to mass media (TV).

[20] The son preference measure used here is the ratio of ideal number of boys to ideal number of girls reported by mothers (averaged across households in the state). See data appendix for more information about the definition of the variable.

V.1.b. Individual Level Evidence

Figure 1 and Table 2 show a high variation in MFR across states and over time and its relationship with state-level income and son preferences. But is the increase in MFR directly related to the practice of prenatal sex selection? To examine this question, we use the individual level data from the NFHS to test whether the propensity of giving birth to a boy is higher among families who might feel a stronger pressure to have a son and whether this propensity increased when prenatal sex determination became feasible.

In **Table 3**, we report the likelihood of a male birth at parity *N* (two or three) as a function of the sex composition of the older siblings who were alive at the time of conception using a linear probability model. We examine two samples: children born between 1975 and 1989 and children born from 1990 onwards. This split is meant to proxy for the availability of ultrasound technology.[21] As seen in column 1, the probability of a male birth during the pre-ultrasound period did not vary significantly across households according to the sex composition of their previous children. In contrast, column 3 shows that during the post-ultrasound period this probability was significantly higher for households that had only girls. Estimates of the differential probabilities are virtually unchanged after adding controls for household characteristics (columns 2 and 4). Moreover, at parity 3, the differential probabilities (relative to the reference group *two sons*) are larger and statistically significant for families with *no sons* than for families with *1 son* after the 1990s (columns 3 and 4 in row 5 show that we can reject equality of the coefficients on *no sons* and *1 son* at the conventional level). These results suggest that families have strong preferences for at least one boy, and not necessarily an aversion of girls.

In **Figure 2** we plot the differential probability of a male birth by year comparing between households with no boys among previous children and the rest, by parity. In each case, we also stratify the sample by household income and son preferences, looking separately at those who live in states with stronger son preferences (state son preference in first round of NFHS above median) and are relatively wealthy (household wealth index in the fourth and fifth quintile) and at other households. We smooth estimates using a 7-year moving average.

[21] The sample split at 1990 is used for illustration purposes. Similar but smaller estimates for the post-ultrasound period are obtained when we use 1985 as a cutoff year and higher estimates are obtained if we split the sample at 1995.

Panel A of the figure shows that until the mid-1980s there were no differences in the probability of a male birth by sex composition of previous children at parity two. Moreover, there were no differences between richer households from strong son preference states and the rest. There is an increase over time, especially after the 1990s, in the differential probability of a male birth at parity 2 for households with no boys versus the rest. This increase is steeper among wealthier households who live in states with stronger son preferences.

Panel B shows that the differential probability of a male birth at parity three did not change much for households who live in states with weaker son preferences or have lower income. In contrast, we see a sharp increase in the differential probability of a male birth among those who live in states with stronger son preferences and are relatively wealthier that starts in the early 1990s.[22]

Evidence presented in Table 3 and Figure 2 shows that the likelihood of a male birth increased significantly in the 1990s among households who presumably had a stronger desire for a son and had more resources. This suggests that increases in MFR at birth are likely to be a result of access to prenatal sex selection. In order to further examine the link between son preference and sex ratios at birth, additional information on ultrasound use and abortions is needed. Information on the incidence of abortions is difficult to obtain since a large share of abortions take place in unofficial and non-regulated facilities. In addition, abortion is usually misreported, especially if it is carried out for sex selection purposes. Nevertheless we can use information available in the survey about ultrasound use.[23]

Columns 1 and 2 of **Table 4** report the differential likelihood that a mother performed an ultrasound test during a pregnancy of parity N (two or three) as a function of the sex composition of her N-1 previous children who were alive at time of conception. As seen in the table, mothers with no living sons are more likely to perform an ultrasound test in pregnancies at parities two and three. We further examine the link between ultrasound use and prenatal sex selection by

[22] In **Appendix Figure A.1** we plot the F-statistics for Chow tests for a structural break for each of the years in each series by parity. The results of these tests show no evidence for a break in the 1980s for any of the subsamples. For high income households in son preference states (solid lines), we see evidence for a structural break in 1990 (parity 2) and in 1995 (parity 3). In contrast, there is no evidence for such break for all other households (dotted lines).

[23] The second and third round of the NFHS survey asked mothers whether they performed an ultrasound test for each of their births during the three (in round 2) or five (in round 3) years prior to the survey. The second round asks about ultrasound use among women who received at least one prenatal checkup during pregnancy while the third round asks about ultrasound use for each pregnancy. To avoid selection issues we use information on ultrasound use from the third round only and limit the analysis to that sample.

looking at the association between a male birth and the sex composition of older children among mothers who reported conducting an ultrasound test during pregnancy versus mothers who reported no ultrasound use. It is worth noting that most women who do an ultrasound do not practice sex selective abortion. In addition, some women who practice sex selective abortion, especially those with more years of schooling, do not report doing an ultrasound as they are aware of the banning of ultrasound for sex selection. Nevertheless, differences in the differential probability of a male birth between the "ultrasound" and "no ultrasound" sample add supporting evidence for the practice of sex selective abortion.

Estimates in columns 3 and 4 show that for the sample of women who report ultrasound use, the probability that a birth is a boy is substantially higher (8 to 10 percentage points) for those who previously had no sons than the rest. In contrast, there is no clear relationship between sex composition of previous children and probability of a male birth for households who did not perform an ultrasound test. The differential probability of a male birth (reported in columns 5 and 6) is small, not statistically significant and has inconsistent signs across parity 2 and 3. Estimates of the "ultrasound" and "no ultrasound" samples are significantly different at the 10 percent level.

In summary, the evidence presented in this subsection suggests that increases in MFR at birth are likely to be induced by the increasing practice of prenatal sex selection rather than by biological or environmental factors. We therefore use MFR at the state level as a proxy for the practice of prenatal sex selection.

V.2. Effect of Prenatal Sex Selection

We exploit variation in the timing and extent of prenatal sex-selection across states (proxied by changes in MFR) and examine whether the observed variation is systematically associated with changes in girls' outcomes relative to boys. To take advantage of the year-to-year variation in sex ratios at birth we use MFR constructed from the NFHS based on the complete birth histories reported by mothers. Another advantage of using the NFHS data as opposed to the census data is that we observed all children ever born (and not only those who survive until age one).[24] On the other hand, in the survey data, the state-level MFR for a single year can be quite noisy (especially for smaller states). [25] We therefore use a smoothed version of MFR which is computed as a 7-year

[24] In practice, as we show later, the mortality adjustment matters little for our results.

[25] While the result of such measurement error is formally ambiguous in a multivariate regression (see e.g. Wooldridge 2002), it tends to bias regression coefficients in the direction of 0, which would work against us finding significant effects of the MFR.

moving average of the ratio of the number of male births to female births by year and state based on the pooled data of the three rounds of the NFHS survey. We also use an alternative measure of MFR based sex ratios from the census where we interpolate sex ratios for inter-census years using a fourth order polynomial function that fits the NFHS data. Finally, we use a third measure where we re-estimate the function adjusting the census MFR for infant mortality.[26] The data appendix describes how we construct these three measures of MFR in detail.

Figure 3 plots the smoothed MFR from the NFHS, the census MFR with the interpolated function for inter-census years and the function that adjusts for infant mortality. The population size at age 0 in 1991 is reported in parenthesis. Overall, we see a large variation in the trends of sex ratios across states. For example, while MFR increased steeply in Punjab, Haryana and Delhi, there was almost no change in Kerala or Nagaland. In other states, like Tamil Nadu, MFR was relatively stable until 2000 and started to increase only then. The figure also shows that trends in MFR series computed from NFHS or Census are similar, though there are some disparities between both series, especially in smaller states, probably due to sampling error in the NFHS. In addition, adjustments for infant mortality do not seem to have an important effect. In any case, as we show later, our results are very similar irrespective of the MFR series we use.

Using these alternative MFR measures, we estimate the following equation:

$$(1)\, y_{ist} = \alpha_{s0} + \alpha_{s1}\, female_i + \delta_{t0} + \delta_{t1}\, female_i + x_i\,'\beta + \pi_0 MFR_{st} + \pi_1\left(MFR_{st} * female_i\right) + \varepsilon_{ist}$$

where y_{ist} is the outcome of child *i* in state *s* born in year *t*, α_{s0} and α_{s1} are vectors of gender-specific state fixed effects, δ_{t0} and δ_{t1} are vectors of gender-specific year-of-birth fixed effects and x_i is a vector of individual characteristics that include indicators for twin birth, residence in an urban area, religion, mother's and father's level of education, mother's age (grouped), wealth quintiles, mass media exposure and mother's age at first birth. MFR_{st} is the Male-Female Ratio at birth for the cohort born in year *t* in state *s*. ε_{ist} is the error term, which is composed of a state-

[26] We also experimented with lower and higher order polynomials and the results (not reported here but available upon request) are largely similar.

specific random element that allows for any type of correlation within observations of the same state across time and an individual random element.[27]

The parameter of interest is π_1, which captures the effect of increases in prenatal sex selection (proxied by increases in MFR) on changes in girls' outcomes. This approach is essentially a continuous version of a *triple-difference* estimation strategy. This strategy has the advantage that it allows us to control for state-level fixed factors that differentially affect boys and girls (for example, the degree of discrimination against girls in a state). We can also control for state-time-varying factors that affect boys and girls similarly within each state and cohort and which might be correlated with changes in MFR over time (for example, improvement in access to health care facilities). In addition, we control for differential trends in boys' and girls' outcomes at the national level in a very flexible way by including gender-specific year of birth fixed effects.

Identification relies on the assumption that changes in MFR within a state over time are unrelated to other unobserved factors that could *differentially* affect male and female outcomes. In this regard, it is important to note that we always include in our model the main effect of MFR, which gives an indicator of the effect of unobserved time varying factors at the state level correlated with MFR, on boys' outcomes. As we show later, estimates of MFR are never significant and have inconsistent signs across outcomes, reducing one of the major concerns regarding omitted variable bias. Our results are also robust to replacing the main MFR effect with a set of full interactions between state- and year-of-birth fixed effects, suggesting that state-time varying factors are unlikely to confound our findings. We further assess the plausibility of the main identifying assumption in Appendix B, where we also consider more general specifications that allow for gender-specific coefficients in all covariates and control for time-varying state characteristics interacted with gender, thus also allowing them to have a differential effect by gender.

A possible concern regarding identification could arise if increases in MFR within a state over time were related to increasing discriminatory preferences against girls rather than increasing access to prenatal sex selection technology. While we cannot rule out this alternative explanation, we note that in this case, our estimates would be biased against finding any improvement in girls' outcomes and would therefore provide a lower bound of the effects of prenatal sex selection. On

[27] In all the specifications, we use the national sampling weights provided in the surveys and cluster standard errors by state. As we show in the next section our results are not sensitive to the use of this specific weighting scheme.

the other hand, note that our estimates capture the overall effect of an increase in prenatal sex selection on girls' outcomes. This includes also any possible general equilibrium effect resulting from parents internalizing future changes in women's value due to an increase in the scarcity of girls. We will turn back to this issue in Section VII below when we examine the possible channels linking between prenatal sex selection and girls' outcomes.

We do not include household fixed effects in the estimating equations. This is due to several reasons: First, identification in a household fixed-effect model would rely on a comparison between siblings of the opposite sex. However, households with children of opposite sex cannot be viewed as a randomly selected sample since the sex composition of children is affected by parental stopping rules in fertility behavior and prenatal sex selection. In addition, given that our main sample includes children born within three years prior to the survey date, a household fixed-effects model would generate estimates for a selected sample of households (i.e. those which had two births within the last three years). Note also that girls' outcomes might change due to selection into different households or a differential reduction in family size. We are therefore interested in the well-being of girls across all families and not necessarily within the same household.

VI. EMPIRICAL RESULTS

1. MAIN RESULTS ON NUTRITIONAL OUTCOMES

Girls' well-being can be analyzed across two dimensions: the allocation of household inputs such as material resources, time, health care and feeding practices or children's outcomes such as nutrition, morbidity and mortality. We focus our analysis on children's nutrition and mortality since they embed information on various types of parental input. These outcomes are measured more easily than most others and are therefore less likely to suffer from measurement error or recall bias. We do not examine morbidity since it is more likely to suffer from bias due to gender differences in diagnosis and recall.[28]

We measure children's nutritional status by anthropometric indicators based on height, weight and age. In particular, we consider three indicators of malnutrition: stunting, underweight and wasting. All three are defined based on z-scores, which are computed by subtracting the median and dividing by the standard deviation of a reference population of the same age and

[28] For example, Timaeous et al. (1998) report sex bias in reporting and recalling episodes of disease and sex differentials in childhood risks of illness and illnesses severity perception.

gender. Specifically, a child is considered stunted if his/her height-for-age is at least 2 standard deviations below the median of the reference population (or the associated z-score is smaller than -2). An underweight child has a weight-for-age at least 2 standard deviations below the median, and a wasted child has a weight-for-height at least 2 standard deviations below the median.

The three indicators capture malnutrition from different perspectives. Stunting reflects long-term malnutrition or cumulative nutrition from conception and is also affected by recurrent or chronic illnesses. Wasting measures acute malnutrition and represents the failure to receive adequate nutrition in the period immediately preceding the survey and may be the result of inadequate food intake or a recent episode of illness leading to weight loss. An important feature of the wasting indicator is that it does not depend on the accuracy of age reporting. On the other hand, it is more sensitive to seasonal shocks. Underweight is a composite index of chronic or acute malnutrition. Note that z-scores are normalized by gender and age so that they take into account that boys and girls may follow different growth trajectories.[29] Our analysis uses z-scores based on the US National Center for Health Statistics (NCHS) standard, which was the most commonly used measure until 2006.[30] About 18 percent of the children aged 0 to 35 months included in our sample have missing values in at least one of the anthropometric indicators. Nevertheless, we do not find any significant gender differences in the likelihood of having a missing value in these indicators. Moreover, we do not find any associations between state variation in MFR and the likelihood of missing anthropometric data for girls or boys (see Appendix **Table A2**). [31]

A methodological challenge in the analysis of nutritional outcomes on the basis of anthropometric data is the availability of these data only for surviving children. Therefore, if prenatal sex selection has any impact on mortality, our analysis of nutrition will be based on a selective sample. For example, if increases in MFR are associated with a reduction in female child mortality, they might lead to an increase in the proportion of girls who are close to a survival

---

[29] In an alternative specification, we included age-in-months and its interaction with *female* (instead of child's year-of-birth) in order to control for any differences in child outcomes that could result from differences in the age distribution of boys and girls. The results are virtually identical to those presented here.

[30] A new international reference population was published by the World Health Organization in 2006. As we show in Table A2, our results are not sensitive to the specific reference chart used to define z-scores.

[31] Height was not measured in the first round of the NFHS in Andhra Pradesh, Himachal Pradesh, Madhya Pradesh, Tamil Nadu and West Bengal because height measuring boards were not available at that time (IIPS, 1995). Nevertheless, our results for weight are highly consistent with the results for height and our main results are unchanged when we limit the sample to states that have anthropometric data in the three survey rounds. We therefore believe that the lack of height data in round 1 for some states is unlikely to bias the main results.

threshold, thus attenuating the estimated effect of MFR on nutritional status. As shown below, the effects on mortality are negligible or not significant, which mitigates the concern about selectivity issues in the nutritional status results.

Table 5 reports estimates for *MFR* and *female*MFR* from a linear probability model for the likelihood of being underweight, wasted, or stunted.[32] Estimates from a basic model that includes no covariates, except for a female dummy, and gender-specific state and year-of-birth fixed effects are reported in columns 3 and 4. Columns 5 and 6 report estimates for the full model specified in equation (1) which controls also for the household characteristics reported in Table 1. The key parameter of interest is the coefficient of the interaction term *female*MFR* (column 6), which is negative for all three outcomes and, except for stunting, is statistically significant. Note that the coefficients for the main effect of *MFR* (column 5) have inconsistent signs across outcomes and are never significant, suggesting that changes in MFR are unrelated to changes in the nutritional outcomes of boys. This last finding is important since it suggests that increases in MFR at the state level are unlikely to be associated with other changes in unobservables that affect children's nutritional status. It is also worth noting that our estimates for the interaction between MFR and female are largely unchanged after the addition of household covariates (see columns 4 and 6) minimizing some concerns of omitted variable bias.[33]

Note that the key *female*MFR* coefficient reflects the differential change in girls' outcomes that results from a one-unit increase in MFR. For example, the estimate for underweight from column 6 indicates that a one standard deviation increase in MFR (0.07 points) is associated with a 4 percentage point reduction in the proportion of girls who are underweight, which is a reduction of 8 percent relative to the mean outcome for girls. Taking Punjab as an example, this would mean that the increase in MFR of 0.2 observed between the first and third round of the NFHS is associated with a 10 percentage point reduction in the proportion of girls who are underweight.

Columns 7 and 8 of Table 5 report estimates for the models that use MFR constructed from census data and inter-census interpolation and columns 9 and 10 report estimates from models that use the same data but adjust census numbers for infant mortality. Estimates are highly similar

[32] Marginal effects from logit models provide similar results.

[33] However, if there is selection in which families have girls (as a result of prenatal sex selection), these covariates would be endogenous. This is why we report results with and without covariates. Nevertheless, as we show later, we find no evidence on selection based on these observable characteristics.

to the results based on NFHS data and suggest that girls' chances to be underweight or wasted are lower in states where prenatal sex selection increased. In all models estimates of the interaction term between MFR and female for stunted are negative and similar in magnitude to the estimates for wasted but they are less precise.[34] However, in contrast to underweight and wasted, the main MFR effect is positive (although not significant). Therefore, our conclusions for the effects of prenatal sex selection on children nutritional status apply to a larger extent to weight and weight for height.[35]

We also conducted additional analyses to test the sensitivity of our results by using alternative definitions for nutritional outcomes, alternative model specifications, and alternative weighting schemes. The analysis is described in **Appendix B** and the results are reported in **Appendix Tables A3** and **A4**. Overall, we find our main results are robust under those different choices.

In addition, we estimated Difference-in-Differences models for boys and girls separately. The results (reported in **Appendix Table A5**) are consistent with our triple difference in difference models in that both sets of results show relative improvement of girls' outcomes compared to boys. Moreover, the DID results show that the relative improvements of girls outcomes are actually absolute improvements, rather than merely reflecting less deterioration of girls' outcomes compared to boys.

So far our findings are consistent with the hypothesis that girls outcomes improved due to the increased use of prenatal sex selection. While an accurate measure of the latter is hard to come by, we explore some limited information on ultrasound use in the data to test whether girls in families that are more likely to perform ultrasound test experienced a greater improvement in health. The third round of the NFHS asked women whether they have undergone an ultrasound test in each of the pregnancies that took place within the five years preceding the survey. We use this information to predict women's likelihood of doing an ultrasound test for the whole sample

[34] One possible explanation is that there is more measurement error in height than in weight for young children given that height is measured lying on a board for children under the age of 2. For example, in our sample, the standard deviation of the z-score for height is 1.68, compared to 1.33 and 1.22 of the z-scores for weight and for weight-for-height, respectively. The conventional ranges are 1.10 to 1.30 for height, 1.00 to 1.20 for weight and 0.85 to 1.10 for weight for height (see WHO 1997).

[35] Note that the three nutritional outcomes are not independent of each other. For example, weight-for-height is mechanically related to weight-for-age and height-for-age. So in our case, if we see a large effect on "underweight" (based on weight-for-age) and a moderate effect on "wasting" (based on weight-for-height), then it should not be surprising to also see a somewhat limited effect on "stunting" (based on height-for-age).

based on women's predetermined characteristics.[36] Next, we examine changes in the frequency of female births and health outcomes within states and over time by mothers' predicted likelihood of ultrasound use. The idea is to examine how outcomes change within households who have a similar likelihood of doing an ultrasound test as the technology spreads over time.

We performed two exercises and report the results in **Table 6**.[37] Column 1 of Panel A reports estimates of a model where we regress an indicator for a female birth on the probability of ultrasound use, year of birth and interactions between both variables, controlling for state fixed effects. We see that the interaction term between the probability of ultrasound and year of birth is negative, suggesting that the proportion of female births declines more over time among households that have a higher probability of performing an ultrasound test.

In columns 2 through 4 we regress the nutritional outcomes on the probability of ultrasound use, year of birth and gender and the interactions between these variables including also a gender-specific state fixed effect. The triple interaction term between female, probability of ultrasound and year of birth is negative for all three outcomes and statistically significant for underweight and stunted, suggesting that there is a larger reduction in the proportion of girls who are malnourished over time among households who are more likely to perform an ultrasound test.

Panel B of the table report results based on a subsample of children whose mothers are unlikely to have undergone an ultrasound test (those with a predicted probability below the 25 percentile – 0.078). For this subsample we checked whether the probability of a female birth changes over time by regressing a female dummy on year of birth and state fixed effects. As seen in column 1 of Panel B, there is no change over time in the probability of female births in this subsample. In columns 2 through 4 of Panel B we checked whether there was a change in the nutritional outcomes for these girls by regressing the nutritional outcomes on year of birth and its interaction with female while controlling for gender specific state fixed effects. The results show no reduction in the proportion of girls who are wasted. Results for underweight and stunted show a declined over time for both boys and girls, but the improvement seen among girls is less

[36] We predict this probability using the following covariates: residence in an urban area, religion, mother's and father's level of education, mother's age, wealth quintiles, index of mass media exposure, mother's age at first birth and state fixed effects. We then use the estimated model to predict the probability of ultrasound test for all women in the three survey rounds.

[37] We computed bootstrapped standard errors to account for the fact that the predicted probability of ultrasound use is estimated. The adjustment is done using the procedure described by Petrin and Train (2003).

pronounced as evidenced by the positive estimate of the interaction term between year of birth and female. Overall we see no evidence for a differential improvement in girls' outcomes in households who are unlikely to perform an ultrasound test.

Taken together, results presented in this table are consistent with our previous findings and suggest that girls' outcomes improved as a consequence of the diffusion of prenatal sex selection. In addition, it seems that an increase in girls' scarcity does not affect health investments in girls for households who are unlikely to practice sex selective abortion; a point that we discuss further below.

2. MECHANISMS

Broadly speaking, there are two ways in which access to prenatal sex selection can lead to changes in outcomes: selection and treatment. The selection effect arises from girls being born into families with different characteristics and son preferences that result in changes in the average outcomes of the girl population. The treatment effect, on the other hand, comes from changes in the behavior of a given household, in response to access to prenatal sex selection (or changes in MFR).

Within this framework, we examine four specific channels through which access to ultrasound technology can affect health outcomes -- family characteristics, family size, wantedness of a child's gender, and breastfeeding practices. The first channel is a selection effect while the latter three can be either selection, treatment, or both.

The improvement in girls' nutritional status could be the result of selection, such that girls are born into families with better endowments. It could also be explained by changes in family size. Girls could be increasingly born into smaller families because families with weaker son preferences might also desire smaller size (selection) or because families became less reliant on stopping rules in fertility behavior (treatment). Similarly, girls could become more wanted due to selection into families with weaker son preferences, or due to changes in preferences for *all* parents as they internalize the change in the future value of women as they become a scarce commodity in marriage and labor markets, even in families that do not practice prenatal sex selection. Finally, girls could be breastfed for longer due to selection into different households or due to a change in mothers' behavior as a response to changes in desired number of children or internalization of changes in MFR.

2.a. Are girls born into better-endowed families?

In order to assess the selection effect, we examine whether family characteristics of girls (such as parental education, mother's age at first birth and household wealth status) improved relative to boys in states where MFR has increased. This is done by estimating a triple-differences equation in which we regress each of the family characteristics on *MFR* and *female*MFR* in a model that controls for state fixed effects, year of birth fixed effects and their interactions with a female indicator.

Table 7 reports estimates for the main MFR effect and the interaction between *MFR* and *female*. The table also reports gender differences in household characteristics (column 3). Estimates of the main MFR effect reported in column 4 (rows 1-6) suggest that states with increasing MFR experienced improvement in some family characteristics, in particular, an increase in the level of parental education and mother's age at first birth and a decline in the likelihood of living in a rural area. This is consistent with the fact that the practice of prenatal sex selection is related to economic development. On the other hand, there is no association between MFR and maternal age or the household wealth index.

Estimates for the interaction between *MFR* and *female* reported in column 5 (rows 1-6) are all small, not significant and have inconsistent signs across the different variables. These results suggest that there is no differential improvement in household characteristics among families with girls relative to families with boys and therefore the improvement in girls' nutritional status is unlikely to be explained by the fact that girls are born into *"better"* families.

The positive association between MFR and some household characteristics raises some concerns regarding our main results for girls' nutritional outcomes if an improvement in household characteristics has a differential effect on the nutritional status of boys and girls. Nevertheless, as reported in Table A4 and discussed in the previous section, our main results were virtually identical to those obtained from an enriched model in which we allow for a differential effect of household characteristics by gender. This suggests that the improvement in parental education or other household characteristics did not have a differential effect on girls' nutritional status.

2.b. Are girls born into smaller families?

As noted above, a direct consequence of parental stopping rules in fertility behavior and son preference is that girls are more likely to be born in larger families. With access to prenatal sex selection, parents of girls would not necessarily need to have more children in order to attain the desired sex composition. Family size could also decrease through selection, if parents who do not practice sex selective abortion have, in general, fewer children. As a result of a decrease in family size, girls might fare better due to an increase in family resources per child.

We examine whether family size has differentially changed for girls relative to boys in regions with increasing MFR. To this end, we regress family size on *MFR* and *female*MFR*. The model also controls for state fixed effects, year of birth effects and their interactions with a female indicator.

Row 7 of Table 7 reports the coefficients of *MFR* and *MFR* interacted with *female*. The estimate reported in column 4 shows a negative association between MFR and family size which may be due to various factors. One possibility is reverse causality. That is, a desire to have fewer children may increase the demand for sons at lower parities, thus increasing the incidence of prenatal sex selection. The association between MFR and family size could also be due to unobserved factors that affect both MFR and fertility. For example, economic development is generally associated with a reduction in fertility and an increase in access to ultrasound technology. A third explanation is a causal link from access to prenatal sex selection to family size. Parents with access to prenatal sex selection are more likely to attain their preferred sex composition of children without the need of having additional children. Note, however, that reverse causality and omitted factors are expected to produce a similar association between MFR and family size for boys and girls while the last one is expected to have a differential effect by gender with larger reductions for girls. As seen in column 5, the estimate for the interaction between *female* and *MFR* is indeed negative (-0.923) and marginally significant (s.e.=0.534). Moreover, estimates reported in the last row of the table show a significant larger reduction in the average number of older siblings for girls relative to boys (estimate=-0.980; s.e.=0.557) which is explained by the reduction in the probability that girls are born at higher parities. These results confirm two expected consequences of prenatal sex selection: a larger reduction in family size for girls relative to boys and a decline in the proportion of girls born at higher parities. Both factors are likely to increase the amount of household resources available for girls.

To further assess the impacts of prenatal sex selection on family size, we also examined changes in number of children over time by households' probability of ultrasound test using the same two-stage procedure applied for the nutritional outcomes. Results reported in **Appendix Table A6** are in line with our findings using MFR.

2.c. Are girls more likely to be wanted?

An additional expected consequence of prenatal sex selection is that girls would be more likely to be born in families that want them. Girls might also be more likely to be wanted if their parents internalize the potential change in the value of women as they become a scarce commodity even in families that do not practice prenatal sex selection. A unique feature of our study is that we are able to assess these channels by examining the association between MFR and parental preferences for sons.

To investigate whether girls are differentially more likely to be born into families with weaker son preference in places with increasing MFR, we regress the various measures of son preference on MFR and MFR*female along with other controls. Note that we cannot establish a causal relationship between MFR and son preference since the link between the two can go in both directions. In addition, there may be other unobserved factors that affect both simultaneously. Nevertheless, it is still interesting to examine the association between MFR and son preference, in particular, the differential association by child gender as any bias related to reverse causality or omitted variables is expected to affect boys and girls similarly.

We use three indices of son preference: two alternative measures based on the ratio of the ideal number of sons to the ideal number of daughters reported by mothers and an indicator for whether the mother reports a desire to have a larger number of sons than daughters. We also define a birth-specific measure capturing the "wantedness" of the child gender. Specifically, we define an indicator for a birth being of wanted gender if the mother's desired number of children of the child's gender is greater than the actual number of older siblings of his/her gender who are alive at the time of the child's conception.

We report in **Table 8** estimates for MFR and its interaction with female for the sample of children aged 0 to 35 months.[38] Results show that for the three measures of son preference (rows 1-3), the coefficient of MFR is positive and the interaction term between MFR and female is negative. This suggests that in places with increasing MFR, boys are more likely to be born into families with stronger son preferences while girls are more likely born into families with weaker son preferences (relative to boys).[39] The last row of the table reports the coefficient for wanted

[38] The sample size is larger than our nutrition sample since we do not restrict it to children who have valid anthropometric data.

[39] To address the issue of reverse causality, we also analyze a separate sample of younger women aged 15 to 20 and examine the relationship between their reported son preference and the MFR of children born 5-10 years prior to the survey date. Presumably reverse causality would be less relevant in these cases. We

gender. Both the main MFR effect and its interaction with female are positive (although they are not statistically significant), and they are consistent with other estimates reported in this table.

Taken together, our analysis on wantedness, suggests that that children born in places with an increasing use of SSA are more likely to satisfy parental preferences in terms of gender composition.

2.d. Are girls receiving better parental care?

An improvement in girls' nutritional status may be a direct consequence of better care and treatment in the family since girls are more likely to be wanted when prenatal sex selection is available. Parental treatment is difficult to assess since we do not directly observe parents behavior. Nevertheless, we can examine mothers' reports on breastfeeding duration as one possible indicator for child care and treatment.[40]

Medical and public health research has found that breastfeeding has large benefits for children's health, especially in environments with poor sanitary conditions (see, for example, The World Health Organization, 2000). The WHO recommends that breastfeeding in developing countries should continue, in combination with supplementary foods, up to the child's second birthday or beyond (WHO, 1991).

Although breastfeeding is nearly universal in India, girls tend to be breastfed for a shorter period than boys. Parents with son preference may breastfeed their sons for a longer period than their daughters either due to direct discriminatory treatment or passively through stopping rules in fertility behavior as mothers stop breastfeeding their daughters sooner in order to conceive again quickly (Jayachandran and Kuziemko, 2010). With access to prenatal sex selection girls' breastfeeding duration might increase either due to lower parental discrimination in child care

look at two different samples: ever married women aged 15-20 and ever married women aged 15-20 who have no children (we cannot look at single women because son preference was asked only to married women). The results reported in **Appendix Table A7** show that son preference changed little in states with increasing sex ratios.

[40] Another potential measure of parental input is vaccination. However, as discussed in Barcellos et al. (2010), evidence from the NFHS data on gender differential in vaccination rates has been mixed. The results vary dramatically depending on whether vaccination is measured based on information recorded on vaccination cards or based on mothers' reports. In this context, selection bias could be a big concern since very few mothers (only about 30%) had a vaccination card and they are also more likely to have a card for boys than for girls. An additional problem with vaccination data from the NFHS is an apparent decline in the record of immunization information in the third round of the NFHS in some states due to logistic and methodological issues, especially among children with no vaccination card (see Chandran et al. 2011). For these reasons, we do not focus on this outcome.

and treatment or due to a decline in parental pressure to conceive a son after the birth of a daughter.

We examine the effect of prenatal sex selection on girls' breastfeeding duration by estimating equation (1) with the dependent variables being indicators of whether the child was breastfed for at least 12, 18 or 24 months. In order to take into account the possibility of right-censoring in duration, the indicators are defined conditional on children being at least 12, 18 or 24 months old, respectively, at the time of the survey.

We estimate the model using our main sample (the youngest two children born within 3 years prior to each survey round). Results are reported in **Table 9.** The differences in outcome means reported in column 3 show a disadvantage for girls in breastfeeding duration that widens with age. For example, girls are one percentage point less likely than boys to be breastfed for at least 12 months, about 4 percentage points less likely to be breastfed for at least 18 months and almost 6 percentage points less likely to be breastfed for at least 24 months.

Estimates for the interaction between *MFR* and *female* suggest some increase in duration of girls' breastfeeding at long durations (24 months or longer) in states with higher increases in MFR. These results suggest that breastfeeding practices may be another channel through which prenatal sex selection improves girls' nutritional outcomes. The increase in breastfeeding duration is also consistent with our results that show an increase in the proportion of girls who are wanted and a larger reduction of family size for girls. As discussed earlier, these two factors are likely to explain gender difference in breastfeeding duration.

3. MORTALITY

Gender preferences and discrimination may also affect child mortality. Indeed, a large literature has documented excess female mortality which becomes evident after the age of one month. In this section, we examine the effect of prenatal sex selection on children survival probabilities. We focus on the following set of binary outcomes: surviving beyond the age of 1 month, 1 year and 5 years.[41] For each outcome, we estimate a linear probability model with a specification similar to that of equation (1).[42] We consider two samples: the first includes the last two children born within 3 years prior to each survey and parallels the sample selection criteria used for the

[41] To deal with censoring, the indicators are only defined for children who would be "old enough" (i.e., one month, one year and five years old, respectively) at the time of the survey.
[42] Results from logit models (not reported here) provide similar findings.

nutritional outcomes and the second includes all children born within 10 years prior to each survey.

The results are reported in **Table 10**. Overall, we do not find significant changes in female survival probabilities in states with a high incidence of prenatal sex selection. The coefficients on the key interaction term *MFR*female* (reported in columns 6 and 11) are small in all outcomes and samples and none of them is statistically significant.

These results are somewhat at odds with our previous findings on nutritional outcomes, family size and breastfeeding duration. Particularly puzzling is why we find a differential improvement in female nutritional status but do not see any significant increase in female survival probabilities. One possible explanation is that families that substitute between prenatal and postnatal female discrimination are not the same families characterized by excess female child mortality. A second possible explanation is that differential recall (by gender) of deaths and measurement error may be biasing the estimated MFR effects towards zero. [43] By comparison, the analysis of nutrition outcomes of living children (born within 3 years prior to the survey) is based on anthropometric measures observed at the time of the survey and is thus less likely to be affected by these biases. A third possible explanation is that while the nutritional status of surviving girls has improved the magnitude of the improvement is still not large enough to increase the likelihood of survival for marginal girls. Medical research suggests that the leading contributors to child mortality are respiratory ailments and infectious and gastro-intestinal diseases. Malnutrition, while often underlying and exacerbating these diseases, is not by itself a fatal factor, except in severe or extreme cases.[44] It is also possible that other types of parental investment, such as preventative care, have a more important impact on reducing diseases and thus mortality.

VII. CONCLUSION

In this paper, we study the impact of prenatal sex selection on the well-being of girls in India, which may come about through selection or/and treatment effect.

[43] However, as we show in **Appendix C**, we find no evidence for differential changes in underreporting (or recalls) of female deaths in the NFHS or any association between underreporting and sex selection.

[44] Our indicators of malnutrition (i.e. anthropometric measures that are two standard deviations below the median of the reference population) are often thought to reflect mild cases. Indeed, when examining comparable samples (children aged 0-35 months), we find only a moderate correlation between the malnutrition indicators and mortality at the state level in each of the survey rounds. Our results are consistent with those of Hill and Upchurch (1995) who report a lack of association between female mortality disadvantage and nutritional disadvantage across several developing countries.

We explore these issues using data from the National Family Household Survey. We proxy parental access to prenatal sex selection using the ratio of male to female births in the year and state in which the child was born and provide evidence showing that starting in the 1990s high sex ratios at birth reflect the practice of prenatal sex selection. We then analyze whether changes in girls' outcomes within states and over time in comparison to boys are associated with changes in sex ratios at birth.

We find that an increase in the practice of prenatal sex selection leads to a differential reduction in the incidence of malnutrition (underweight and wasting) among surviving girls. We find *no association* between increases in prenatal sex selection and boys' outcomes suggesting that our findings for girls are unlikely to be driven by unobserved state-time-varying factors associated with the increasing practice of prenatal sex selection. Our results are highly robust to several checks that assess the likelihood of possible biases due to state-specific differential trends by gender associated with prenatal sex selection.

An exploration of the channels linking between prenatal sex selection and girls' outcomes shows no evidence for a selection of girls into families with different observable socio-demographic characteristics. However, consistent with some of the expected consequences of prenatal sex selection, we find an increase in the proportion of girls born in families with weaker son preferences. In addition, we find evidence of a larger reduction in family size for girls and a larger decline in the average number of older siblings for girls, which is explained by the reduction in the probability that girls are born at higher parities. We also find some suggestive evidence of better treatment of girls as reflected in breastfeeding duration. On the other hand, we do not find an increase in girls' survival probabilities or a reduction in son preference as reported by mothers in states with an increasing incidence of prenatal sex selection.

REFERENCES

Abrevaya, J. 2009. "Are There Missing Girls in the United States? Evidence from Birth Data." American Economic Journal: Applied Economics, Vol. 1, pp. 1-34.

Almond, Douglas and Lena Edlund. 2008. "Son-Biased Sex Ratios in the 2000 United States Census." Proceedings of the National Academy of Sciences, of the United States of America, Vol. 105 No. 15, pp. 5681-5682.

Almond, Douglas, Lena Edlund, and Kevin Milligan. 2009. "Son Preference and the Persistence of Culture: Evidence from Asian Immigrants to Canada." National Bureau of Economic Research, Working Paper No. 15391.

Almond, Douglas, Hongbin Li, and Lingsheng Meng. 2010. "Son Preference and Early Childhood Investments in China", unpublished manuscript.

Ananat, Elizabeth O., Jonathan Gruber, B. Levine, and Douglas Staiger. 2009. "Abortion and selection." The Review of Economics and Statistics, Vol. 91. No. 1. pp. 124–136.

Arnold, Fred, Sunita Kishor, and T.K. Roy. 2002. "Sex-Selective Abortions in India." Population and Development Review, Vol. 28 no. 4, pp. 759-785.

Arnold, Fred, Minja Kim Choe, and T.K. Roy. 1998. "Son Preference, the Family- Building Process and Child Mortality in India." Population Studies, Vol. 52, No. 3, pp. 301-315.

Anderson Siwan and Debraj Ray. 2012. "The Age Distribution of Missing Women in India." Economic and Political Weekly, Vol. 47, No. 47-48, pp. 87

Barcellos, Silvia, Leandro Carvalho and Adriana Lleras-Muney. 2010. "Child Gender and Parental Investment in India: Are Boys and Girls Treated Differently?." Working paper.

Becker, Gary and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." Journal of Political Economy, Vol. 87, pp. 1153-89.

Becker, Gary and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." Journal of Labor Economics, Vol. 4, pp. S1-S39.

Ben-Porath, Yoram and Finis Welch. 1976. "Do Sex Preferences Matter?" Quarterly Journal of Economics, Vol. 90 No. 2, pp. 285-307.

Bhalotra, Sonia and Tom Cochrane. 2010. "Where Have All the Young Girls Gone? Identification of Sex Selection in India." IZA Discussion Paper No. 5381.

Bharadwaj, Prashant and Leah K. Nelson. 2010. "Discrimination Begins In the Womb: Evidence Of Sex-Selective Prenatal Investments." unpublished manuscript, UCSD.

Bhaskar, Venkataraman, and Gupta Bishnupriya. 2007. "India's development in the era of growth." Oxford Review of Economic Policy, Vol 23, No. 2, pp. 135-142.

Bhat, Mari P. N. 2002. "On the Trail of 'Missing' Indian Females – II: Illusion and Reality", Economic and Political Weekly, Vol. 37. No. 52, pp. 5244-5263.

Case Anne and Christina Paxson. 2010. "Causes and Consequences of Early Life Health", Demography, Vol. 47S, pp. S65-S85.

Chhabra, Rami, and Sheel C. Nuna. 1993. "Abortion in India: an overview." New Delhi: Virendra Printers.

Chandran S. Anil, Ramachandra Srikrishna and Sekhar P. Satya. 2011. "National Family Health Survey-3 Reported Low Full-Immunization Coverage Rates in Andhra Pradesh, India: Who is to be Blamed?" Journal of Public Health, pp. 1-7.

Currie, J. 2009. “Healthy, wealthy and wise: socioeconomic status, poor health in childhood, and human capital development”, Journal of Economic Literature, Vol. 47 No. 1, pp. 87–122.

Das, Gupta, Monica. 1987. “Selective Discrimination Against Female Children in Rural Punjab.” Population and Development Review, Vol. 13 No. 1, pp. 77-100.

Das, Gupta M., and P. N. Mari Bhat. 1997. “Fertility Decline and the Increased Manifestation of Sex Bias in India.” Population Studies Vol. 51 No. 3, pp. 307-315.

Deaton, Angus. 1997. “The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.” Baltimore and London: John Hopkins University Press.

Donohue, John J., and Levitt Steven. 2001. “The Impact of Legalized Abortion on Crime.” Quarterly Journal of Economics, Vol. 116, no. 2, pp. 379-420.

Dreze, Jean, Sen Amartya. 1997. “Indian Development: Selected Regional Perspectives.” Oxford: Oxford University Press.

Dubuc, Sylvie, and Coleman David. 2007. “An Increase in the Sex Ratio of Births to India-Born Mothers in England and Wales: Evidence for Sex-Selective Abortion.” Population and Development Review, Vol. 33, No. 2, pp. 383-400.

Duggal, Ravi 2004. “The Political Economy of Abortion in India: Cost and Expenditure Patterns”, Reproductive Health Matters Vol. 12 No. 24, pp. 130-137.

Duggal, Ravi, and Ramachandran Vimala. 2004. “The Abortion Assessment Project – India: Key Findings and Recommendations.” Reproductive Health Matters, Vol. 12, pp. 122-129.

Dyson, Tim, and Mick Moore. 1983. “On Kinship Structure, Female Autonomy, and Demographic Behavior in India.” Population and Development Review 9 (1): 35-60.

Everett, Jana 1998. “Indian Feminists Debate the Efficacy of Policy Reform: The Maharashtra Ban on Sex-Determination Tests”, Social Politics, Col. 5 No. 3, pp. 314-337.

George, Sabu. M. 1997. “Female Infanticide in Tamil Nadu, India: From Recognition Back to Denial?.” Reproductive Health Matters, Vol. 10, pp.124-31.

George, Sabu M. 2002. “Sex Selection/Determination in India: Contemporary Developments.” Reproductive Health Matters, Vol. 10 No. 19, pp. 190-192.

George, Sabu M. 2006. “Sex Ratio in India.“ The Lancet*,* Vol. 367, p. 1725.

Goodkind, Daniel. 1996. "On Substituting Sex Preference Strategies in East Asia: Does Prenatal Sex Selection Reduces Postnatal Discrimination?” Population and Development Review, Vol. 22, No. 1, pp. 111-125.

Griliches, Zvi and Hausman, Jerry A.. 1986. “Errors in Variables in Panel Data”, Journal of Econometrics, Vol. 31 Issue 1.

Hill, Kenneth, and Dawn M. Upchurch. 1995. “Gender Differences in Child Health: Evidence from the Demographic and Health Surveys.” Population and Development Review, Vol. 21 No. 1, pp. 127-151.

Hu, Luojia and Analia Schlosser. 2012. “Trends in Prenatal Sex Selection and Girls’ Nutritional Status in India,” CESifo Economic Studies, Vol. 58, No. 2, pp.348-372.

IIPS (International Institute for Population Sciences). 1995. National Family Health Survey (MCH and Family Planning): India 1992–93. Bombay: International Institute for Population Sciences.

Jacobsen, R., Møller, H., and Mouritsen A. 1999. “Natural Variation in the Human Sex Ratio.” Human Reproduction, Vol. 14 No. 12, pp. 3120-3125.

Jayachandran, Seema, and Ilyana Kuziemko. 2009. “Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India.” NBER Working Paper No. 15041.

Jefferey, Roger, Jefferey Patricia, and Andrew Lyon. 1984. “Female Infanticide and Amniocentesis.” Social Science and Medicine, Vol. 19 No. 11, pp. 1207-1212.

Jensen, Robert. 2003. “Equal Treatment Unequal Outcomes? Generating Sex Inequality Through Fertility Behavior” Mimeo, Harvard University.

Khanna, Sunil K. 1997. “Traditions and reproductive technology in an urbanizing north Indian village.” Social Science & Medicine, Vol 44, No. 2, pp. 171-180.

Kishor, Sunita. 1993. “May God Give Sons to All: Gender and Child Mortality in India.” American Sociological Review, Vol. 58, No. 2, pp. 247-265.

Kishor, S. 1993. “May God Give Sons to Us All: Gender and Child Mortality in India”, American Sociological Review, Vol. 58 No. 2, pp. 247-265.

Kishwar, M. 1995. “When Daughters are Unwanted: Sex Determination Tests in India.” Manushi 86: 15-22.

Krugman, Andrea. 1998. “Being Female Can Be Fatal: an Examination of India’s Ban on Pre-Natal Gender Testing.” Cardozo Journal of International and Comparative Law, Issue 1, pp. 215-238.

Kulkarni, P. M. 2007. “Estimation of Missing Girls at Birth and Juvenile Ages in India”. Paper commissioned by the United Nations Population Fund, UNFP.

Lin Ming-Jen, Jin-Tan Liu, and Nancy Qian. 2009. “More Missing Women, Fewer Girls Dying: The Impact of Abortion on Sex Ratios at Birth and Excess Female Mortality in Taiwan”, NBER Working Paper w14541

Mallik, Rupsa. 2003. “‘Negative Choice’ Sex Determination and Sex Selective Abortion in India.” Abortion Assessment Project – India. Center for Enquiry into Health and Allied Themes (CEHAT), Survey No. 2804 & 2805. Maharashtra, India.

Mahapatra, Prasanta (2010) “An Overview of the Sample Registration System in India”, Prince Mahidol Award Conference and Global Health Information Forum.

Miller, Barbara. 1981. “The endangered sex: neglect of female children in rural north India.” Ithaca: Cornell University Press. [re-published with a new chapter, "The Girl Child in India in the Year 2000," pp. 197-217, Oxford University Press, 1997].

Miller, Barbara. 1984. “Daughter Neglect, Women's Work and Kinship: Pakistan and Bangladesh Compared.” Medical Anthropology Vol. 8, No. 2, pp. 109-126.

Pande, Rohini P. August 2003. “Selective Gender Differentials in Childhood Nutrition and Immunization in Rural India: The Role of Siblings.” Demography Vol. 40, No. 3, pp 395-418.

Petrin, Amil and Train Kenneth 2003. “Omitted Product Attributes in Discrete Choice Models”, NBER Working Paper 9452.

Portner, Claus C. 2010. “Sex Selective Abortion, Fertility and Birth Spacing.” mimeo, University of Washington.

Ravindran, T.K.S. 2002. “Gender Gaps in Research on Abortion in India: A Critical Review of Selected Studies”. The Gender and Reproductive Health Research Initiative. CREA. New Delhi.

Retherford, Robert D., and T. K. Roy. 2003. “Factors affecting sex-selective abortion in India and 17 major states.” National Family Health Survey Subject Report, No. 21.

Rose, Elaina. 1999. “Consumption Smoothing and Excess of Female Mortality in Rural India.” The review of Economics and Statistics, Vol. 81 No. 1, pp. 41-49.

Rosenzweig, Mark R., and Paul Schultz. 1982. “Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India.” The American Economic Review, Vol. 72, No. 4, pp. 803-815.

Saikia, Nadita, Domantas Jasilionis, Faujdar Ram, and Vladimir M. Shkolnikov (2011), “Trends and Geographic Differentials in Mortality Under 60 in India”, Population Studies, Vol. 65, No 1, pp. 73-89.

Shepherd, Mary Elizabeth. 2008. “The Selective Abortion in India: The Impact on Child Mortality.” Cambria Press.

Sen, Amartya. 1992. “Missing Women.” BMJ, vol. 304, pp. 586-587.

Sen, Amartya. 2003. “Missing Women – Revisited.”, BMJ, vol. 327, pp. 1297-1299.

Simmons, George B., Celeste Smucker, Stan Bernstein, and Eric Jensen. 1982. “Post-Neonatal Mortality in Rural India: Implications of an Economic Model.” Demography, Vol. 19 No. 3, pp. 371-389.

Subramanian, S. V., and S. Selvaraj. 2009. “Social Analysis of Sex Imbalance in India: Before and After the Pre-Natal Diagnostic Techniques (PNDT) Act.” J Epidemiol Community Health Vol. 63, pp. 245-252.

Sundar, Ramamani, 2003. “Abortion Costs and Financing a Review”. Abortion Assessment Project – India. Centre for Enquiry into Health and Allied Themes, Mumbai.

Timaeus, Ian, Harries Katie, Fairbairn Francesco. 1998. “Can use of health care explain sex differentials in child mortality in the developing world.” in UNITED NATIONS (Ed.), Too young to die: genes or gender? United Nations Population Division, New York.

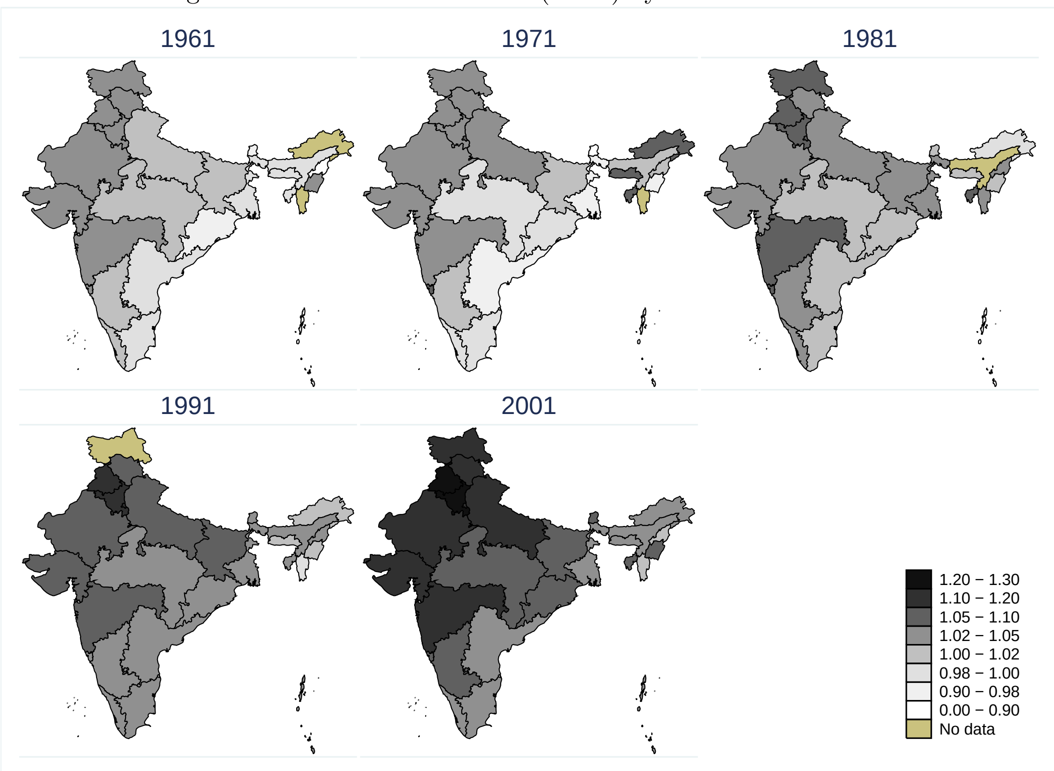
Visaria, Pravin M. 1971. “The Sex-ratio of the Population of India.” Census of India 1961, Volume 1: Monograph 10. New Delhi: Ministry of Home Affairs.

Wooldridge, Jeffrey 2002. “Economtric Analysis of Cross Section and Panel Data.” Cambridge, MA: MIT Press.

World Health Organization. 2000. “Effect of Breastfeeding on Infant and Child Mortality Due to Infectious Diseases in Less Developed Countries: A Pooled Analysis.” Lancet, Vol. 355, No. 9202, pp. 451-455.

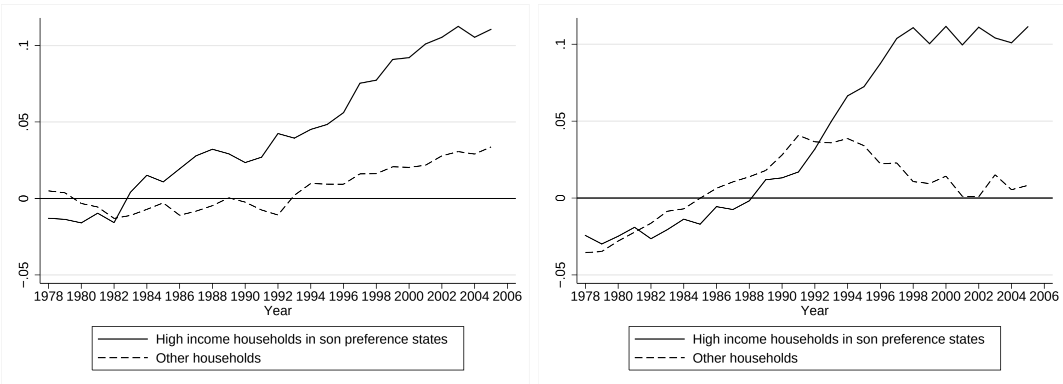
Yamaguchi, Kazuo. 1989. “A Formal Theory for Male-Preferring Stopping Rules of Childbearing : Sex Differences in Birth Order and in the Number of Siblings.” Demography, Vol. 26, No. 3, pp. 451-465.

Figure 1: Male to Female Ratio (MFR) by State - Census data



Notes: The figure plots estimates of the Male-to-Female Ratio by state based on census counts of children alive at age 0 (0 to 11 months) for the different years.

Figure 2: Differential Probability of a Male Birth by Year of Birth



(a) Parity 2 (b) Parity 3

Notes: The figure plots estimates for the differential probability of a male birth by year - (7-year moving average) comparing between households with no boys among previous children and the rest (panel a for parity two and panel b for parity three). The full line plots the differential probability for households in the 4th or 5th wealth quintile who live in states with son preferences above median (according to first round of NFHS). See data appendix for the definition of son preference measure. The dotted line plots the differential probability for all other households.

Figure 3: MFR By State

Notes: The figure plots the Male-to- Female Ratio by state based on alternative data sources. The solid line (NFHS) is based on a 7-year moving average of the ratio of the number of male births to female births from the NFHS data, the dashed line (Interpolat.) interpolates the Census MFR for inter-census years, where the interpolation is based on a fourth order polynomial function estimated from the NFHS data. The dotted line (interpolate. adj. IM) is based on census MFR adjusted for infant mortality and inter-census years interpolated using NFHS data. The dots (Census) denote the MFR computed from census data (number of children aged 0-11 months at the census date).

Table 1. Summary Statistics

	Nutrition sample			Mortality sample		
	Girls	Boys	Difference	Girls	Boys	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Urban	0.236	0.237	-0.002	0.232	0.235	-0.003
			(0.003)			(0.002)
Index of mass media exposure	0.799	0.824	-0.025	0.738	0.755	-0.017
			(0.010)			(0.005)
Wealth index	2.82	2.86	-0.041	2.74	2.77	-0.030
			(0.009)			(0.009)
Mother's Age	25.2	25.3	-0.098	28.7	28.7	-0.047
			(0.024)			(0.019)
Mother's age at 1st birth	19.0	19.0	0.000	18.7	18.7	-0.023
			(0.025)			(0.013)
Mother's education						
No education	0.541	0.538	0.004	0.613	0.607	0.006
			(0.004)			(0.003)
Primary school	0.158	0.148	0.009	0.146	0.147	-0.001
			(0.003)			(0.002)
Secondary school	0.247	0.256	-0.010	0.200	0.203	-0.004
			(0.005)			(0.002)
Higher	0.053	0.057	-0.004	0.040	0.042	-0.002
			(0.002)			(0.001)
Missing	0.001	0.001	0.000	0.001	0.001	0.000
			(0.000)			(0.000)
Father's education						
No education	0.300	0.291	0.009	0.341	0.339	0.002
			(0.003)			(0.002)
Primary school	0.193	0.189	0.004	0.205	0.201	0.003
			(0.005)			(0.003)
Secondary school	0.381	0.389	-0.008	0.346	0.350	-0.004
			(0.006)			(0.002)
Higher	0.119	0.124	-0.005	0.101	0.103	-0.002
			(0.003)			(0.002)
Missing	0.006	0.007	-0.001	0.006	0.007	0.000
			(0.001)			(0.000)
Religion						
Hindu	0.792	0.792	0.001	0.791	0.794	-0.003
			(0.002)			(0.002)
Muslim	0.158	0.156	0.002	0.161	0.156	0.005
			(0.003)			(0.001)
Other religion	0.049	0.052	-0.003	0.047	0.049	-0.002
			(0.002)			(0.002)
Missing	0.001	0.001	0.000	0.001	0.001	0.000
			(0.000)			(0.000)
Number of children in the family	2.920	2.927	-0.007	3.861	3.736	0.125
			(0.009)			(0.015)
Mother wants another child	0.486	0.363	0.123	0.292	0.211	0.081
			(0.009)			(0.006)
Sample size	36,940	39,560	76,500	172,472	185,337	357,809

Notes: The table reports summary statistics for boys and girls (cols. 1,2,4, and 5) included in the analysis samples and differences between the characteristics of girls and boys (cols. 3 and 6). Standard errors of the differences clustered at the state level are reported in parenthesis. The samples pool rounds 1, 2, and 3 of the NFHS. The *nutrition sample* reported in columns 1-3 includes the last two children under three years of age of ever married women with valid anthropometric data. The *mortality sample* reported in columns 4-6 includes all children born within the last 10 years preceding the survey date of ever married women. Observations are weighted using national-level weights.

Table 2. Predictors of State Changes in MFR

	(1)	(2)	(3)	(4)
Income per capita	0.999		1.209	-5.519
	(0.456)		(0.369)	(4.141)
Income per capita x 2001	1.080		1.193	-4.073
	(0.340)		(0.331)	(3.026)
2001	-1.890	-3.604	-12.65	10.89
	(1.860)	(4.411)	(4.927)	(13.094)
Ideal MFR		11.84	13.99	-7.887
		(2.976)	(2.889)	(13.493)
Ideal MFR x 2001		5.394	7.520	-9.607
		(3.512)	(3.319)	(9.442)
Income per capita x Ideal MFR				4.943
				(3.136)
Income per capita x Ideal MFR x 2001				3.870
				(2.257)
Number of observations	52	52	52	52
R-sq	0.330	0.254	0.576	0.632

Notes: The Table reports regression estimates where the dependent variable is the state MFR (in 1991 or 2001). The explanatory variables are income per capita in 1990 (in thousand Rs), ideal MFR (averaged at the state level), a dummy for year 2001 and interactions between income or ideal MFR and year 2001 dummy. The model in column (4) includes also interactions between income per capita and ideal MFR. See data appendix for definition of ideal MFR. Standard errors clustered at the state level are reported in parenthesis.

Table 3. Differential Probability of a Male Birth at Parity N as a Function of the Sex Composition of Older Siblings

	Born between 1975-1989		Born after 1989	
	No controls	Full controls	No controls	Full controls
	(1)	(2)	(3)	(4)
A. Parity 2 (omitted category=Boy)				
No sons	-0.005	-0.005	0.019	0.020
	(0.003)	(0.003)	(0.007)	(0.007)
Sample Size	*45,198*	*45,198*	*74,422*	*74,422*
B. Parity 3 (omitted category=Boy-Boy)				
No sons	-0.005	-0.006	0.039	0.036
	(0.011)	(0.011)	(0.012)	(0.012)
1 son	0.005	0.005	0.012	0.011
	(0.007)	(0.007)	(0.009)	(0.009)
p-value for test (No sons=1 son)	0.268	0.206	0.003	0.004
Sample Size	*30,804*	*30,804*	*45,200*	*55,289*

Notes: The table reports the differential probability of a male birth at parity 2 (panel A) and parity 3 (panel B) as a function of the sex composition of older siblings. The samples include all children born in the 15 years prior to each survey date. Estimates reported in columns 1-2 are for children born between 1975 and 1989. Estimates reported in columns 3-4 are for children born in 1990 or afterwards. Regression estimates reported in columns 2 and 4 are from models that control also for twin status, mother's age, mother's education, mother's age at first birth, indicators for mother's religion, father's education, mother's mass media exposure, wealth, and rural/urban status. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis. Sample sizes are reported in *italics* .

Table 4. Sex-Ratios and Ultrasound Use

	Ultrasound Test		Pregnancy outcome = boy			
	Full Sample		Ultrasound Sample		No Ultrasound Sample	
	No controls (1)	Full controls (2)	No controls (3)	Full controls (4)	No controls (5)	Full controls (6)
A. Parity 2 (omitted category=1 son)						
No sons	0.025	0.019	0.077	0.078	0.025	0.025
	(0.011)	(0.008)	(0.016)	(0.017)	(0.017)	(0.016)
Sample Size	*14,048*	*14,048*	*5,098*	*5,098*	*8,950*	*8,950*
B. Parity 3 (omitted category=2 sons)						
No sons	0.111	0.081	0.124	0.103	-0.010	-0.013
	(0.027)	(0.022)	(0.056)	(0.055)	(0.026)	(0.026)
1 son	0.030	0.019	0.034	0.025	0.008	0.007
	(0.015)	(0.015)	(0.068)	(0.067)	(0.033)	(0.033)
p-value for test (No sons=1 son)	0.000	0.000	0.001	0.006	0.218	0.217
Sample Size	*8,069*	*8,069*	*1,922*	*1,922*	*6,147*	*6,147*

Notes: Columns 1 and 2 report the differential likelihood that a mother performs an ultrasound test during pregnancy as a function of the sex composition of her older children. Columns 3 and 4 report the differential likelihood of a male birth as a function of the sex composition of the older children among mothers who reported to have performed an ultrasound test during that pregnancy. Columns 5 and 6 report the differential likelihood of a male birth as a function of the sex composition of the older children among mothers who reported that they did not performed an ultrasound test during that pregnancy. Panels A and B report estimates for parities 2 and 3, respectively. The omitted category in all regressions is one son for parity two and two sons for parity three. Regression estimates reported in columns 2, 4 and 6 are from models that control also for twin status, mother's age, mother's education, mother's age at first birth, indicators for mother's religion, father's education, mother's mass media exposure, wealth, rural/urban status, and year of birth and state fixed effects. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis. Sample sizes are reported in *italics* .

Table 5. Effects on Nutritional Status of Children

			MFR from NFHS							
			Basic specification		Full controls		MFR from Census and inter-census interpolation		MFR from Census adjusted for infant mortality and inter-census interpolation	
Outcome	Sample size (1)	Outcome mean (2)	MFR (3)	Female *x* MFR (4)	MFR (5)	Female *x* MFR (6)	MFR (7)	Female *x* MFR (8)	MFR (9)	Female *x* MFR (10)
Underweight	76,314	0.485	-0.162	-0.535	-0.095	-0.536	-0.025	-0.431	-0.024	-0.426
			(0.336)	(0.160)	(0.318)	(0.132)	(0.249)	(0.135)	(0.253)	(0.135)
Wasted	69,784	0.179	-0.050	-0.247	-0.029	-0.250	-0.122	-0.301	-0.117	-0.292
			(0.284)	(0.109)	(0.275)	(0.104)	(0.220)	(0.124)	(0.218)	(0.126)
Stunted	69,571	0.433	0.316	-0.287	0.386	-0.293	0.385	-0.280	0.350	-0.265
			(0.230)	(0.180)	(0.236)	(0.164)	(0.292)	(0.181)	(0.293)	(0.182)

Notes: Sample sizes are reported in column 1. Means of the dependent variables are reported in Column 2. Columns 3 and 4 report regression estimates for MFR and MFR interacted with a female dummy from a linear probability model that controls for state, survey round, year of birth fixed effects and their interactions with gender. Columns 5 through 10 report regression estimates from models that control also for twin status, mother's age, mother's education, mother's age at first birth, indicators for mother's religion, father's education, mother's mass media exposure, wealth, and rural/urban status. Estimates in columns 3 though 6 come from models that use MFR based on the NFHS data as the main explanatory variable. Estimates in columns 7 and 8 come from models that use MFR constructed from census data and inter-census interpolation and estimates in columns 9 and 10 come from models that use the same data but adjust census numbers for infant mortality. See data Appendix for details about the construction of the MFR series. The sample includes children aged 0 to 35 months born in the last two births of ever married women sampled in rounds 1-3 of the NFHS surveys. Sample sizes are smaller for wasted and stunted because height measures were not taken in round 1 for the following 5 states: Andhra Pradesh, Himachal Pradesh, Madhya Pradesh, Tamil Nadu and West Bengal. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Table 6. Outcomes by Probability of Ultrasound Use

	Female (1)	Underweight (2)	Wasted (3)	Stunted (4)
A. Full Sample				
Prob. of ultrasound use	5.251	-10.184	-1.191	-21.513
	(1.505)	(4.857)	(4.304)	(6.876)
Year of birth	-0.000	-0.010	-0.001	-0.013
	(0.000)	(0.001)	(0.002)	(0.002)
Prob. of ultrasound use x Year of birth	-0.003	0.005	0.001	0.011
	(0.001)	(0.002)	(0.002)	(0.003)
Female x prob. of ultrasound		16.392	4.818	13.363
		(6.024)	(5.836)	(5.789)
Female x Year of birth		0.004	0.003	0.002
		(0.002)	(0.002)	(0.002)
Female x prob. of ultrasound x year of birth		-0.008	-0.002	-0.007
		(0.003)	(0.003)	(0.003)
Number of observations	97,347	76,167	69,654	69,444
B. Low probability of ultrasound sample (25 percentile - prob.< 0.078)				
Year of Birth	-0.00038	-0.007	-0.001	-0.012
	(0.00041)	(0.001)	(0.001)	(0.002)
Female x Year of birth		0.004	0.003	0.002
		(0.002)	(0.002)	(0.002)
Number of observations	24,630	18,306	16,981	16,941

Notes: The table reports the probability of a female birth (column 1) and nutritional outcomes (columns 2 through 4) by the estimated probability of ultrasound use and additional covariates. Panel A uses the full sample of children aged 0 to 35 months born in the last two births of ever married women sampled in rounds 1-3 of the NFHS surveys. Panel B includes only children born in households with a predicted probability of ultrasound use below the 25 percentile (prob.<0.078). Estimates reported in column 1 come from models that control also for state fixed effects. Estimates reported in columns 2-4 come from models that control for state fixed effects interacted with gender. Observations are weighted using national-level weights. Bootstrapped standard errors clustered at the state level are reported in parenthesis.

Table 7. Family Characteristics of Girls vs. Boys

Outcome	Sample size (1)	Outcome mean (2)	Females -Males (3)	MFR (4)	Female *x* MFR (5)
Mother's age	76,500	25.3	-0.098	2.952	-1.817
			(0.024)	(2.268)	(1.703)
Mother's age at 1st birth	76,500	19.0	0.000	4.863	0.785
			(0.025)	(1.704)	(0.873)
Mother's education	76,394	3.62	-0.118	6.489	0.243
			(0.042)	(1.769)	(1.419)
Father's education	76,032	6.15	-0.137	4.538	0.682
			(0.051)	(1.215)	(2.275)
Wealth index	76,500	2.84	-0.041	-0.268	-0.039
			(0.009)	(0.623)	(0.522)
Rural	76,500	0.763	0.002	-0.300	0.158
			(0.003)	(0.121)	(0.159)
Number of children	76,500	2.92	-0.007	-1.421	-0.923
			(0.009)	(0.672)	(0.534)
Number of older siblings	76,500	1.82	-0.026	-1.172	-0.980
			(0.009)	(0.593)	(0.557)

Notes: Column 2 reports means of the dependent variables and column 3 reports the female-male differential. Columns 4 and 5 report regression estimates for MFR and MFR interacted with a female dummy from a model that includes gender specific state and year of birth fixed effects. The sample includes children aged 0 to 35 months born in the last two births of ever married women sampled in rounds 1-3 of the NFHS surveys. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Table 8. MFR and Son Preference

Outcome	Sample size (1)	Outcome mean (2)	Females -Males (3)	MFR (4)	Female *x* MFR (5)
Ideal MFR I	82,040	1.37	-0.090	0.413	-0.284
			(0.009)	(0.208)	(0.155)
Ideal MFR II	82,040	1.35	-0.076	0.387	-0.343
			(0.009)	(0.201)	(0.155)
Wants more sons than daughters	85,504	0.36	-0.063	0.115	-0.216
			(0.005)	(0.144)	(0.121)
Child of wanted gender	85,686	0.71	-0.149	0.242	0.143
			(0.009)	(0.136)	(0.091)

Notes: The dependent variables in rows 1 and 2 are based on the ratio of the ideal number of sons to the ideal number of daughters reported by mothers. The dependent variable in row 3 is an indicator for whether the mother reports a desire to have a larger number of sons than daughters. The dependent variable in row 4 is an indicator for whether the child is of wanted gender which is constructed by comparing between mother's responses to ideal number of sons and daughters and sex composition of older siblings. See data appendix for more details about the construction of these variables. Means of the dependent variables are reported in Column 2. Column 3 reports differences in means of the dependent variables between females and males. Columns 4 and 5 report regression estimates from a model that includes state and year of birth fixed effects and their interactions with female and controls also for twin status, mother's age, mother's education, mother's age at first birth, indicators for mother's religion, father's education, mother's mass media exposure, wealth, and rural/urban status. The sample includes children aged 0 to 35 months born in the last two births of ever married women sampled in rounds 1-3 of the NFHS surveys. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Table 9. Breastfeeding Duration

Outcome Breastfed for:	Sample size (1)	Outcome mean (2)	Females -Males (3)	MFR (4)	Female *x* MFR (5)
At least 12 months	59,670	0.881	-0.011	0.093	-0.058
(age>=12 months)			(0.004)	(0.132)	(0.133)
At least 18 months	43,295	0.707	-0.038	0.128	-0.206
(age>=18 months)			(0.007)	(0.190)	(0.230)
At least 24 months	29,537	0.544	-0.056	0.037	0.391
(age>=24 months)			(0.007)	(0.186)	(0.155)

Notes: Means of the dependent variables are reported in Column 2. Column 3 reports differences in means of the dependent variables between females and males. Columns 4 and 5 report regression estimates for MFR and MFR interacted with female from a linear probability model that controls for state and year of birth fixed effects and their interactions with female and the set of covariates specified in Table 5. In addition, the models control for age in months and age in months interacted with a female dummy. The minimum age restriction used in each of the regressions is specified in the first column of the table. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Table 10. Effects on Survival Probability

Outcome	Sample (1)	Children Born within 36 months preceding survey: Sample size (2)	Outcome mean (3)	Females -Males (4)	MFR (5)	Female x MFR (6)	Children Born within 120 months preceding survey: Sample size (7)	Outcome mean (8)	Females -Males (9)	MFR (10)	Female x MFR (11)
Survived beyond first month	Age ≥	98,922	0.958	0.007	-0.010	0.024	356,361	0.952	0.007	-0.001	0.019
	1 month			(0.002)	(0.031)	(0.047)			(0.001)	(0.031)	(0.034)
Survived beyond 12 months	Age ≥	63,961	0.936	0.004	0.028	0.029	323,916	0.925	0.003	0.010	0.003
	12 months			(0.002)	(0.063)	(0.098)			(0.002)	(0.058)	(0.039)
Survived beyond age 5	Age ≥	---	---	---	---	---	188,924	0.892	-0.004	-0.062	-0.031
	60 months								(0.005)	(0.080)	(0.067)

Notes: The sample for columns 2-6 includes all children aged 0 through 35 months born in the last two births to ever married women sampled in rounds 1-3 of the NFHS surveys. The sample for columns 7-11 includes all children born in the last ten years prior to survey date to ever married women sampled in rounds 1-3 of the NFHS surveys. The minimum age restriction used for each outcome is specified in column 1. Columns 3 and 8 report means of the dependent variables and columns 4 and 9 report the female-male differential in the dependent variables. Columns 5 and 6 and columns 10 and 11 report regression estimates for MFR and MFR interacted with a female dummy from a linear probability model that includes gender specific state and year of birth fixed effects and control also for the covariates specified in table 5. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Appendix A – Data description

Definitions of MFR

MFR from NFHS data

We use MFR constructed from the NFHS based on the complete birth histories reported by mothers of children born in the last 15 years. We focus on birth events that took place in the last 15 years because a large proportion of children aged 15 and above do not live in the household so misreporting of births is more likely for older children. In addition, Anderson and Ray (2012) find that under reporting of deaths, and female deaths in particular, is more severe for ages above 15. We pool the three survey rounds and summarize the number of female and male births by state and year of birth. We define MFR as the total number of male births divided by the total number of female births in a given state and a year. To reduce volatility due to sampling error, we use a smoothed version of MFR which is computed as a 7-year moving average of the ratio of the number of male births to female births by year and state. Note that pooling data from the three survey rounds dilutes a bit the problem of recall bias since we count children born in the same year who are at different ages at the survey date. Smoothing the data reduces also possible errors of age misreporting and age heaping. We assessed the sensitivity of our results to the specific smooth window in rows 10 and 11 of Appendix Table A4. Estimates are highly consistent with our main results. As an additional robustness check, we also computed MFR using only children born in the last 10 years (instead of 15). Results (not reported here but available upon request) are virtually unchanged.

MFR from Census

We also construct an alternative measure of MFR that incorporates the Census data for the years 1981, 1991, 2001, and interpolates inter-census years using NFHS data. For each state, we specify a fourth order polynomial function to interpolate the MFR for inter-census years. We estimate the parameters of the function by running a constrained nonlinear regression of the MFR from the NFHS with the constraint that the fitted curve goes through the Census MFR in the 3 census years. We weight each cell by the number of observations in the NFHS data. We also experimented using a higher or lower order of polynomial to interpolate the inter-census years. Results (not shown here but available upon request) are highly consistent with our main results.

MFR from Census adjusted for infant mortality

Because census data reports only population counts of children who are alive at the enumeration day, we also re-estimate the polynomial function adjusting the census MFR for infant mortality. For this purpose, we compiled infant mortality state data from publications of the Sample Registration System (SRS) which is the main data source for Indian vital statistics. Although the SRS data is a much larger sample than the NFHS, it still suffers from sampling error and misreporting (see Saikia et al, 2011 and Mahapatra, 2010). Publications for infant mortality are not available for the exact census years and data are missing for some states. We therefore use 1985 data to adjust census counts of 1981 and 1990 data to adjust census counts of 1991. Infant mortality data is missing for the following states and years: Arunachal Pradesh (1985, 1990), Delhi (1985, 1990), Goa (1985, 1990), Himachal Pradesh (1990), Jammu & Kashmir (1990), Manipur (1985, 1990), Meghalaya (1985, 1990), Mizoram (1985, 1990), Nagaland (1985, 1990), Sikkim (1985, 1990), and Tripura (1985, 1990). For these cases, we use data from the closest year available.

Son Preference Measures

Mothers were asked in NFHS survey about desired number of children (n) , desired number of boys (r^B), desired number of girls (r^G), and desired number of children for which sex would not matter ($r^O = n - r^B - r^G$). We take numeric responses to these questions (about 95 percent of the cases) and define two alternative measures of son preference as follows:

The first measure is computed as

$$\boldsymbol{ideal\ MFR1} = \frac{r^B + 0.5 * r^O}{r^G + 0.5 * r^O}$$

The second measure is computed as

$$\boldsymbol{ideal\ MFR2} = \frac{\frac{n}{2} + \sum_{k=0}^{n}(r^B - k)Pr(B=k)1(k<r^B) - \sum_{k=0}^{n}(r^G - k)Pr(G=k)1(k<r^G)}{\frac{n}{2} - \sum_{k=0}^{n}(r^B - k)Pr(B=k)1(k<r^B) + \sum_{k=0}^{n}(r^G - k)Pr(G=k)1(k<r^G)}$$

where

$$Pr(B = K) = Pr(G = K) = \frac{n!}{k!\,(n-k)!}\left(\frac{1}{2}\right)^n$$

The intuition behind the second measure is the following. Consider a woman who reports wanting *n* children. If she has neutral preference for boys and girls (r^B=0 and r^G=0), her ideal MFR is given by (n/2)/(n/2)=1. Now suppose she wants r^B boys (r^B>0). To find the corresponding ideal MFR, we first find all possible sequences that would violate her preference $(k < r^B)$ where k is the actual number of boys. We then calculate the associated probabilities for these sequences, *Pr(B=k)*, for all relevant *k*, where *Pr(B=k)*, is the pdf for a binomial distribution $Bi(n,p)$, given by the equation above. Next we hypothetically re-assign sex to the undesired sequences to satisfy her preference (assuming she can selectively abort girls) and compute the resulting ideal MFR. Specifically, we adjust the neutral ideal MFR by adding to the numerator and subtracting from the denominator the total number of desired boys from all the undesirable sequences ($\sum_{k=0}^{n}(r^B - k)Pr(B = k)1(k < r^B)$).

For example, suppose a family wants 3 children, 1 boy, and the other two do not matter (n=3, r^B=1 and r^G=0). There are eight possible sequences of boys and girls born into this family, each with probability 1/8. There is only one sequence (three girls) that is undesired. The family could replace this with one boy and two girls (e.g. by selectively aborting a third girl until it conceives a boy). Across the 24 births in the eight sequences this would yield 13 boys and 11 girls, or an MFR of 1.18. And this can be verified by the above formula:

$$ideal\ MFR2 = \frac{\frac{3}{2} + \sum_{k=0}^{3}(1-k)Pr(B=k)1(k<1)}{\frac{3}{2} - \sum_{k=0}^{3}(1-k)Pr(B=k)1(k<1)} = \frac{\frac{3}{2} + (1-0)Pr(B=0)}{\frac{3}{2} - (1-0)Pr(B=0)} = \frac{\frac{3}{2} + \left(\frac{1}{2}\right)^3}{\frac{3}{2} - \left(\frac{1}{2}\right)^3}$$
$$= 1.18$$

Similar calculations can be done for the case where she wants r^G girls (r^G>0). The general formula above includes both cases (assuming families would selectively abort boys as well as girls).

One problem with these measures (*ideal MFR1* and *Ideal MFR2*) is that while they work well for most cases, they might become ill-defined in some extreme cases. For example, for women who report that they want all boys ($r^B = n, r^G = r^O = 0$), these measures become ∞ and thus the data is missing.

We therefore consider an additional measure of son preference: an indicator for wanting more sons than girls:

$$\boldsymbol{more\ sons}\ = 1\{r^B > r^G\}$$

Finally, we construct a measure for a child of wanted gender**.** Specifically, we define an indicator for a female birth being wanted to be 1 if the desired number of girls ($r^G + 0.5 * r^O$) is greater than the actual number of older sisters alive at the girl's conception and 0 otherwise. The indicator is defined in a similar way for a male birth.

Appendix B – Robustness Tests

We consider alternative definitions for nutritional outcomes: we use the continuous z-scores and also examine an alternative measure applied sometimes in the medical literature defined as the ratio of a measured value in the individual to the median value of the reference population for the same age and sex, expressed as a percentage: percent of the median. We also tested the sensitivity of our results to the use of specific growth standards. As noted above, in 2006 the World Health Organization published new growth standards for assessing children's nutritional status. We re-defined the anthropometric indicators using the new WHO tables and re-estimated the main specification. Estimates using these alternative outcomes reported in Appendix Table A3 confirm our main findings – girls' nutritional outcomes improved in places with increasing MFR. In this case, the coefficient for the interaction term between MFR and female is also significant for the height outcomes, although, as before, the main MFR effect points to a worse outcome for boys (but not significant). Taken together, the results suggest an increase in the practice of prenatal sex selection leads to an improvement in girls' nutritional outcomes (especially weight and weight for height) among surviving girls.

We also performed additional tests to assess the validity of our identifying assumptions and check for the robustness of the main results. The results of these tests are reported in Appendix **Table A4**. In order to facilitate comparison, we reproduce the estimates obtained in our main specification in the first row of the table. Overall, the additional estimates are qualitatively similar across various models and specifications.

First, we assess the validity of our main identifying assumption that there are no changes in unobserved factors associated with MFR that could affect girls' outcomes. Note that this is partially addressed by the triple-difference strategy which controls for state-level time-varying factors that affect boys and girls similarly by including a main MFR effect. As we note before, the main effect of MFR is never significant and is inconsistent across outcomes suggesting that conditional on other covariates included in our model, changes in MFR are not associated with changes in other factors that affect boys' outcomes. We also consider a more flexible specification with the main MFR effect being replaced by a set of full interactions between state- and year-of-birth fixed effects to allow for unobserved state-specific time-varying factors. Estimates for the interaction term between female and MFR are virtually unchanged for the three outcomes (see Table A4 row 2) suggesting that our results are unlikely to be driven by state-time-varying unobserved factors associated with MFR that affect child outcomes.

Still, there may be state-time-varying factors that differentially affect boys and girls. This could take place if increases in MFR are related to more rapid economic development and modernization, which in turn may affect girls' and boys' health outcomes differentially (for example, by providing better health care access or by means of a reduction in discrimination against girls). A possible way to account for unobserved time-varying factors that differentially affect both genders would be to control for state-gender-specific trends. However, this approach is less suitable to our set-up as we do not have enough observations from cohorts born before prenatal sex selection to extrapolate a long-term trend. Moreover, MFR evolves following a close to linear trend after the diffusion of prenatal sex selection. Therefore, the inclusion of gender-state-specific linear trends absorbs most of the variation in MFR and magnifies any possible measurement error in this variable (Griliches and Hausman, 1986).

We therefore apply an alternative approach where we estimate expanded versions of equation (1) in which we control for household and time-varying state characteristics and their interactions with gender. Specifically, we estimate three models with alternative sets of additional controls. In the first set, we add interactions between household covariates and a female dummy to the basic model. In the second and third specifications, we include, in addition to gender-specific controls for household characteristics, two alternative sets of state-level time-varying covariates interacted with gender. The first set of state variables is compiled from the NFHS data and includes state means of the wealth index and the proportion of households in urban areas, proportion with electricity access, and proportion with access to TV. The alternative set contains state information on net domestic product and per-capita net domestic product (compiled from reports of the India's central bank -- the Reserve Bank of India).

The estimates for the three expanded models, reported in rows 3 through 5 of **Table A4,** are largely similar to those obtained in our main specification suggesting that our results are unlikely to be driven by unobserved changes at the state level correlated with MFR that had a differential impact on boys and girls.

We further assess the likelihood of the main identifying assumption by examining the sensitivity of our results to the inclusion of additional covariates associated with improvement in health outcomes. We first add to our main specification controls for birth weight and its interaction with female. While birth weight could itself be affected by prenatal sex selection, we are interested in the sensitivity of our results to the inclusion of this covariate and its interaction with female as it should absorb any changes at the state level that differentially affect boys and girls nutritional status at birth. Unfortunately, only 30 percent of the sample has data on birth weight. However, almost all children (99 percent) have valid information regarding their size at birth as reported by mothers (very large, larger than average, average, smaller than average or very small). We find that mother's report on size at birth is highly correlated with birth weight for the subsample of children who have valid data on birth weight. We therefore impute birth weight for those with missing data using size at birth and all additional covariates.[1] We then re-estimate our main model after including controls for birth weight, an indicator for imputed birth weight, and their interactions with female. In an alternative specification, we include indicators for size at birth and their interactions with female instead of birth weight.

The results of these two specifications are reported in rows 6 and 7 of **Table A4** and confirm that our estimates of MFR and their interaction with female are largely unaffected by the inclusion of these additional covariates.[2] This result suggests that for our estimates to be biased, there should be unobserved time-varying factors correlated with MFR that differentially affect boys' and girls' nutritional outcomes *after* birth but not *at birth*.

As an additional check, we also estimated a specification that controls for delivery at a health facility and its interaction with female in an attempt to proxy for improvements in the availability of health care that could differentially affect boys and girls outcomes. Again in this case, we find that our results (reported in row 8 of the table) are not sensitive to the addition of these additional covariates.

In rows 9-11 of the table we assessed the robustness of our results to the specific definition of the main explanatory variable. Note that our basic specification includes MFR in a linear fashion although it may be the case that the relationships between MFR and the outcomes of interest are nonlinear. We therefore re-estimated equation (1) after replacing MFR with its natural logarithm. The estimated effects reported in row 9 are of comparable magnitude to our main results.[3] We also experimented with different ways

---

[1] Specifically, we impute birth weight by estimating separate regressions for each gender and survey round of birth weight on size at birth, year of birth, state fixed effects, and all household covariates.

[2] Note that if prenatal sex determination affects birth weight the estimates from these two specifications that control for birth weight and its interaction with female would be attenuated. Interestingly, we find that our estimates are not reduced after controlling for birth weight. Indeed, in an additional analysis we estimated two models where we regressed birth weight or an indicator for small size at birth on MFR and its interaction with female and all additional covariates specified in equation (1) and found that the estimates of MFR and its interaction with female were not significant.

[3] For example, the coefficient on female*ln(MFR) in the underweight regression is -0.609 (s.e.=0.147) which implies that a 0.2 -point increase in MFR from 1.05 to 1.25 (or 17%) is associated with a decrease of 10 percentage points (17*0.609) in the proportion of underweight girls, which is similar to the results reported in Table 5.

of smoothing MFR. For example, we considered both a shorter and a longer window for computing the moving average. The results, which are reported in rows 10 and 11, are similar to those obtained from our main specification.

Finally, we also looked at the sensitivity of our results to the weighting scheme of the data by re-estimating our model using unweighted data (row 12) and using state weights (row 13). All the estimates are similar to the main results.

We also estimate the impacts of prenatal sex selection using a discrete version of the triple-differences approach, where we stratify states by region and compare between changes in outcomes of girls relative to boys in northern and western states (regions with the sharpest increase in sex ratios at birth) versus all the rest. Results were very much in line with the findings presented in this paper. Namely, we observe an improvement in girls' nutritional outcomes (relative to boys) in north-western states. For details, see Hu and Schlosser (2012).

Appendix C – Exploring Differential Recall of Death by Gender and Its Association with Sex-Selection

There is no direct way to measure underreporting of deaths in the NHFS. However, we can partially address this issue by comparing deaths rates of the same cohort over the survey rounds, both for the full sample and by gender. Specifically, we compute the death rate for children during their first year of life (among children who were older than 1 year of age) and compare differences in death rates of the cohorts that are observed in two survey rounds.[4]

We do this by estimating the following model:

$$Died\_0\_1_{istj} = \alpha_s + \delta_t + \beta_0 Post_j + \beta_1 Female_i + \beta_2 (Post_j * Female_i) + \varepsilon_{istj}$$

Where *Died_0_1* is an indicator for whether the child *i* born in state *s* in year *t* and observed in survey round *j* died between ages 0 and 11 months. α_s and δ_t are vectors of state and year of birth fixed effects respectively. *Post* is an indicator for the later survey round (round 2 for cohorts 1989-1991 and round 3 for cohorts 1996-1998) and female is an indicator for a female child. If there are differential changes in the reporting of female deaths over time, we should find that β_2 is statistically significant.

[4] Given the age restrictions, we can compare death rates of the 1989-1991 cohorts between rounds 1 and 2 of the survey and death rates of the 1996-1998 cohorts between rounds 2 and 3 of the survey.

We estimate the model comparing death rates in rounds 1 and 2 and in rounds 2 and 3 and also pooling the three survey rounds. The results are reported in Appendix Table A.8. As seen in columns 1 through 3, the coefficient of the interaction term between Female and Post is very small and never significant suggesting that there are no changes in differential recall of deaths by gender.

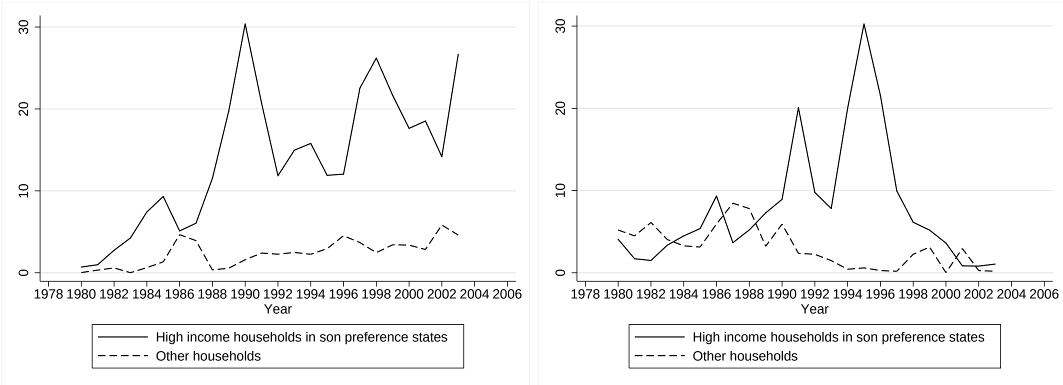
We also examine whether underreporting of deaths is related to sex selection. To do so, we extended the model to include also a proxy for sex-selection (MFR) interacted with a post dummy (indicating the later survey round) and female dummy and the 3-way interactions. To proxy for sex selection, we used the state-level MFR at age 0 from the 2001 Census because we think it is useful to have an MFR measure from a different source than the NFHS and that the year 2001 reflects more clearly where prenatal sex selection takes place.

As can be seen in the results reported in column 4 of the same table (Appendix Table A.8), there is a positive coefficient on MFR*female, suggesting that states that have prenatal sex selection have also relatively higher female mortality rates. However, there is no evidence of any association between recall or differential recall by gender and prenatal sex selection (i.e. the coefficients of the interactions MFR*post and MFR*female*post are not statistically significant from zero, either when looking at each separately or when looking at them jointly).

To summarize, we find no evidence suggesting that there were differential changes in underreporting (or recalls) of female deaths in the NFHS or that underreporting was related to sex selection.

A Appendix

Figure A.1: F-Values for Structural Break Test in Differential Probability of a Male Birth by Year of Birth



(a) Parity 2 (b) Parity 3

Notes: The figure plots F-values for structural break tests in the differential probability of a male birth by year of birth. The full line plots F-values for households in the 4th or 5th wealth quintile who live in states with son preferences above median (according to first round of NFHS). The dotted line plots F-values for the sample based on all other households. See data appendix for the definition of son preference measure. See Figure 2 for details on the probability of male series.

Table A1. State Characteristics

	1990 NDP PC (Rs.)	Urban	Wealth Index	HH with electricity	Son preference: Ideal MFR I	Son preference: Ideal MFR II	Son preference: Wants more sons than daughters	TV exposure	Mother's Illiterate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
North									
Delhi	11,057	0.92	4.79	0.96	1.31	1.28	0.302	0.83	0.37
Haryana	7,508	0.26	3.90	0.86	1.45	1.43	0.451	0.49	0.64
Himachal Pradesh	4,910	0.10	3.62	0.92	1.33	1.31	0.367	0.47	0.50
Jammu & Kashmir	3,816	0.18	3.74	0.88	1.50	1.49	0.492	0.50	0.57
Punjab	8,318	0.28	4.26	0.94	1.48	1.47	0.480	0.57	0.53
Rajasthan	4,191	0.20	2.79	0.54	1.62	1.62	0.577	0.18	0.82
West									
Gujarat	5,891	0.35	3.60	0.78	1.46	1.36	0.424	0.39	0.55
Maharashtra	7,439	0.42	3.54	0.76	1.31	1.28	0.359	0.47	0.50
Northeast									
Arunachal Pradesh	5,398	0.15	3.17	0.62	1.50	1.46	0.430	0.29	0.70
Assam	4,281	0.12	2.44	0.20	1.41	1.41	0.436	0.18	0.59
Manipur	3,976	0.32	3.55	0.64	1.37	1.37	0.435	0.38	0.48
Meghalaya	4,375	0.19	3.10	0.43	1.01	1.00	0.143	0.24	0.51
Mizoram	4,474	0.49	3.82	0.76	1.18	1.18	0.331	0.25	0.08
Nagaland	4,990	0.21	3.64	0.78	1.14	1.14	0.278	0.23	0.43
Tripura	3,370	0.20	2.96	0.47	1.31	1.31	0.327	0.34	0.42
Sikkim	5,302	0.14	3.73	0.80	1.16	1.15	0.218	0.56	0.49
Central									
Madhya Pradesh	4,049	0.22	2.85	0.65	1.55	1.52	0.516	0.27	0.74
Uttar Pradesh	3,590	0.20	2.62	0.34	1.68	1.66	0.565	0.19	0.76
East									
Bihar	2,660	0.15	2.32	0.17	1.65	1.62	0.556	0.13	0.78
Orissa	3,077	0.15	2.42	0.29	1.45	1.43	0.451	0.16	0.67
West Bengal	4,673	0.27	2.67	0.30	1.28	1.28	0.319	0.33	0.51
South									
Andhra Pradesh	4,531	0.26	3.20	0.65	1.33	1.29	0.329	0.39	0.69
Goa	8,797	0.50	4.32	0.92	1.24	1.21	0.278	0.71	0.34
Karnataka	4,598	0.33	3.27	0.66	1.25	1.24	0.271	0.40	0.61
Kerala	4,200	0.28	3.89	0.61	1.15	1.14	0.183	0.42	0.16
Tamil Nadu	4,983	0.35	3.42	0.66	1.07	1.07	0.115	0.50	0.50

Notes: The table reports selected economic and demographic characteristics by state and measures of son preference. Data on net domestic product per capita (NDP PC) in column 1 are reported in Rupees and are based on reports of the Reserve Bank of India. Data for columns 2-9 are based on state averages from the first round of the NFHS. Tabulations for Sikkim are based on the second round of the NFHS as Sikkim was not sampled in the first round. See data appendix for definition of son preference measures.

Table A2. Are Missing Anthropometric Values Correlated with Gender or MFR?

	Sample size (1)	Outcome mean (2)	Females -Males (3)	Basic specification: MFR (4)	Basic specification: Female *x* MFR (5)	Full controls: MFR (6)	Full controls: Female *x* MFR (7)
Missing weight	91,374	0.160	-0.002	0.620	-0.048	0.605	-0.042
			(0.002)	(0.368)	(0.091)	(0.352)	(0.093)
Missing weight for height	91,374	0.246	-0.001	0.828	-0.125	0.877	-0.124
			(0.003)	(0.901)	(0.082)	(0.916)	(0.091)
Missing height	91,374	0.248	0.000	0.838	-0.118	0.888	-0.118
			(0.003)	(0.899)	(0.084)	(0.913)	(0.093)

Notes: The dependent variables are indicators for missing values in weight (row 1), weight for height (row 2), or height (row 3). Means of the dependent variables are reported in Column 2. Gender differences in outcomes are reported in column 3. Columns 4 and 5 report regression estimates for MFR and MFR interacted with a female dummy from a linear probability model that controls for state and year of birth fixed effects and their interactions with gender. Columns 6 and 7 report regression estimates from a model that controls also for twin status, mother's age, mother's education, mother's age at first birth, indicators for mother's religion, father's education, mother's mass media exposure, wealth, and rural/urban status. The sample includes children aged 0 to 35 months born in the last two births of ever married women sampled in rounds 1-3 of the NFHS surveys. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Table A3. Effects on Nutritional Status of Children
Using Alternative Anthropometric Indices

Outcome	Sample size (1)	Outcome mean (2)	Basic specification: MFR (3)	Basic specification: Female x MFR (4)	Full controls: MFR (5)	Full controls: Female x MFR (6)
Weight for age	76314	-1.873	0.727	1.038	0.794	1.060
z-score			(0.864)	(0.453)	(0.847)	(0.358)
Weight for height	69784	-0.980	1.506	0.926	1.507	0.963
z-score			(1.225)	(0.394)	(1.235)	(0.408)
Height for age	69571	-1.729	-1.126	0.430	-0.911	0.683
z-score			(0.867)	(0.416)	(0.739)	(0.312)
Weight for height	76314	79.944	9.042	11.665	8.819	11.498
percent of reference median			(8.352)	(5.387)	(8.148)	(3.790)
Weight for height	69784	91.212	13.706	9.893	14.269	9.711
percent of reference median			(11.570)	(4.130)	(11.497)	(4.216)
Height for age	69571	93.390	-4.195	1.451	-3.095	2.312
percent of reference median			(3.333)	(1.537)	(2.782)	(1.181)
Underweight	77,195	0.440	-0.054	-0.519	-0.019	-0.554
new-zscore			(0.273)	(0.154)	(0.258)	(0.139)
Wasted	69,883	0.221	-0.040	-0.246	0.060	-0.211
new-zscore			(0.275)	(0.091)	(0.250)	(0.093)
Stunted	70,221	0.499	0.378	-0.162	0.260	-0.216
new-zscore			(0.201)	(0.132)	(0.203)	(0.119)

Notes: Means of the dependent variables are reported in Column 2. Columns 3 and 4 report regression estimates for MFR and MFR interacted with a female dummy from a linear probability model that controls for state and year of birth fixed effects and their interactions with gender. Columns 5 and 6 report regression estimates from a model that controls also for twin status, mother's age, mother's education, mother's age at first birth, indicators for mother's religion, father's education, mother's mass media exposure, wealth, and rural/urban status. The sample includes children aged 0 to 35 months born in the last two births of ever married women sampled in rounds 1-3 of the NFHS surveys. Sample sizes are smaller for wasted and stunted because height measures were not taken in round 1 for the following 5 states: Andhra Pradesh, Himachal Pradesh, Madhya Pradesh, Tamil Nadu and West Bengal. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Table A4. Robustness Checks

		Underweight		Wasted		Stunted	
		MFR	Female *x* MFR	MFR	Female *x* MFR	MFR	Female *x* MFR
		(1)	(2)	(3)	(4)	(5)	(6)
1	Main Results	-0.095	-0.536	-0.029	-0.250	0.386	-0.293
		(0.318)	(0.132)	(0.275)	(0.104)	(0.236)	(0.164)
2	State x Year of birth FE	---	-0.560	---	-0.267	---	-0.304
			(0.134)		(0.101)		(0.165)
3	Household covariates interacted with	-0.107	-0.514	-0.036	-0.236	0.379	-0.276
	gender	(0.319)	(0.126)	(0.278)	(0.105)	(0.239)	(0.164)
4	State varying controls 1 + HH	-0.091	-0.433	0.140	-0.135	0.415	-0.217
	covariates interacted with gender	(0.254)	(0.115)	(0.181)	(0.114)	(0.175)	(0.112)
5	State varying controls 2 + HH	-0.303	-0.384	-0.009	-0.227	0.168	-0.206
	covariates interacted with gender	(0.289)	(0.092)	(0.240)	(0.110)	(0.170)	(0.127)
6	Birth weight	-0.055	-0.549	-0.028	-0.250	0.415	-0.303
	interacted with gender	(0.327)	(0.149)	(0.270)	(0.124)	(0.269)	(0.175)
7	Size at birth	-0.085	-0.525	-0.031	-0.235	0.383	-0.276
	interacted with gender	(0.325)	(0.129)	(0.282)	(0.100)	(0.237)	(0.150)
8	Delivery at health	-0.092	-0.518	-0.029	-0.243	0.395	-0.288
	facility interacted with gender	(0.318)	(0.132)	(0.274)	(0.105)	(0.235)	(0.166)
9	Log MFR	-0.099	-0.609	-0.029	-0.295	0.463	-0.330
		(0.344)	(0.147)	(0.302)	(0.105)	(0.254)	(0.188)
10	5-year span for MFR	-0.103	-0.426	-0.015	-0.184	0.175	-0.222
		(0.269)	(0.131)	(0.199)	(0.077)	(0.205)	(0.161)
11	9-year span for MFR	0.130	-0.591	0.080	-0.316	0.540	-0.236
		(0.386)	(0.159)	(0.295)	(0.138)	(0.308)	(0.193)
12	Unweighted regression	0.011	-0.474	0.151	-0.290	0.186	-0.245
		(0.235)	(0.112)	(0.203)	(0.103)	(0.137)	(0.101)
13	State weights	0.011	-0.471	0.136	-0.286	0.201	-0.267
		(0.240)	(0.111)	(0.204)	(0.095)	(0.142)	(0.096)

Notes: The table reports estimates from various robustness checks. See Appendix B of the paper for a detailed explanation of each test.

Table A5. Difference in Difference Models for Nutritional Outcomes

	Sample size		Outcome mean		MFR from NFHS		MFR from Census and inter-census interpolation		MFR from Census adjusted for infant mortality and inter-census interpolation	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)	Boys (9)	Girls (10)
Underweight	39475	36839	0.477	0.493	-0.107	-0.621	-0.035	-0.448	-0.033	-0.442
					(0.319)	(0.301)	(0.250)	(0.224)	(0.254)	(0.223)
Wasted	36178	33606	0.186	0.171	-0.036	-0.272	-0.125	-0.418	-0.120	-0.404
					(0.278)	(0.257)	(0.223)	(0.212)	(0.221)	(0.214)
Stunted	36081	33490	0.426	0.440	0.379	0.103	0.380	0.109	0.345	0.089
					(0.239)	(0.231)	(0.291)	(0.222)	(0.292)	(0.221)
Weight for age	39475	36839	-1.864	-1.882	0.836	1.819	0.487	1.131	0.472	1.080
z-score					(0.846)	(0.825)	(0.705)	(0.701)	(0.712)	(0.701)
Weight for height	36178	33606	-0.991	-0.969	1.530	2.456	1.341	2.137	1.243	2.013
z-score					(1.240)	(1.310)	(1.096)	(1.184)	(1.088)	(1.182)
Height for age	36081	33490	-1.722	-1.735	-0.873	-0.268	-0.851	-0.555	-0.778	-0.501
z-score					(0.736)	(0.695)	(0.865)	(0.637)	(0.870)	(0.637)
Weight for age	39475	36839	80.269	79.596	9.304	19.904	5.269	11.613	5.077	10.981
percent of reference median					(8.153)	(8.418)	(7.197)	(7.909)	(7.257)	(7.933)
Weight for height	36178	33606	91.464	90.942	14.490	23.853	12.679	20.311	11.717	19.057
percent of reference median					(11.529)	(12.635)	(10.171)	(11.469)	(10.096)	(11.462)
Height for age	36081	33490	93.476	93.297	-2.944	-0.945	-2.985	-2.086	-2.705	-1.879
percent of reference median					(2.768)	(2.684)	(3.212)	(2.461)	(3.231)	(2.461)

Notes: Sample sizes are reported in columns 1 and 2. Means of the dependent variables are reported in Column 2 and 3. Columns 5-10 report differences-in-difference estimates from models that control for state, survey round, year of birth fixed effects. In addition, the models include the following covariates: twin status, mother's age, mother's education, mother's age at first birth, indicators for mother's religion, father's education, mother's mass media exposure, wealth, and rural/urban status. Estimates in columns 5 and 6 come from models that use MFR based on the NFHS data as the main explanatory variable. Estimates in columns 7 and 8 come from models that use MFR constructed from census data and inter-census interpolation and estimates in columns 9 and 10 come from models that use the same data but adjust census numbers for infant mortality. See data Appendix for details about the construction of the MFR series. The sample includes children aged 0 to 35 months born in the last two births of ever married women sampled in rounds 1-3 of the NFHS surveys. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Table A6. Family Size by Probability of Ultrasound Use

A. Full Sample	
Prob. of ultrasound use	4.560
	(22.483)
Year of birth	-0.015
	(0.007)
Prob. of ultrasound use x Year of birth	-0.004
	(0.011)
Female x prob. of ultrasound	22.707
	(12.019)
Female x Year of birth	0.003
	(0.003)
Female x prob. of ultrasound x year of birth	-0.011
	(0.006)
Number of observations	97347
B. Low prob. of ultrasound sample (25 percentile - Prob.< 0.078)	
Year of Birth	-0.003
	(0.008)
Female x Year of birth	0.003
	(0.006)
Number of observations	24,630

Notes: The table reports changes in family size within states and over time by the probability of ultrasound use. Panel A uses the full sample and Panel B includes only households with a predicted probability of ultrasound use below the 25 percentile (prob.<0.078). Estimates come from models that control for state fixed effects interacted with gender. Observations are weighted using national-level weights. Bootstrapped standard errors clustered at the state level are reported in parenthesis.

Table A7. Son Preference and MFR

	MFR of cohorts born between 5-10 years prior to survey date	
	Ever married age 15-20 (1)	Ever married age 15-20 with no children (2)
Ideal MFR I	-0.105	0.070
	(0.225)	(0.234)
	1.329	*1.300*
Ideal MFR II	0.034	0.135
	(0.197)	(0.211)
	1.307	*1.285*
Wants more sons than daughters	-0.111	0.209
	(0.259)	(0.232)
	0.332	*0.310*
Number of women	26,987	12,452

Notes: The table reports regression estimates of son preference on the state average MFR of cohorts born between 5 and 10 years prior to survey date. Rows one and two report regression estimates where the dependent variable is defined based on the ratio of the ideal number of sons to the ideal number of daughters reported by mothers (ideal MFR). The third row reports regression estimates where the dependent variable is an indicator that equals one if the mother reported a larger ideal number of sons relative to her ideal number of daughters. See Data Appendix for more details about the definition of son preference measures. All estimates come from models that control for state fixed-effects and indicators for mother's year of birth, mother's age, mother's education, mother's age at first birth, mother's religion, father's education, mother's mass media exposure, wealth, and rural/urban status. The sample in column 1 includes all ever married women of ages 15-20 sampled in rounds 1-3 of the NFHS surveys. The sample in column 2 includes all ever married women of ages 15-20 with no children sampled in rounds 1-3 of the NFHS surveys. Means of the dependent variables are reported in italics. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Table A8. Assessing Differential Recall of Female Deaths by Survey Round and Associations with MFR

	Round 1 vs. Round 2 (1)	Round 2 vs. Round 3 (2)	Pooling all rounds (3)	Pooling all rounds + Interaction with MFR (4)
Post	0.005	0.019	0.011	-0.002
	(0.006)	(0.005)	(0.005)	(0.064)
Female	-0.008	-0.003	-0.006	-0.100
	(0.005)	(0.004)	(0.003)	(0.051)
Female x Post	0.000	-0.003	-0.001	-0.015
	(0.007)	(0.008)	(0.006)	(0.073)
MFR x Female	--	--	--	0.085
				(0.045)
MFR x Post	--	--	--	0.011
				(0.057)
MFR x Female x Post	--	--	--	0.013
				(0.063)
Outcome mean	0.076	0.070	0.073	0.073
F-Stat (MFR x Post = MFR x Female x Post = 0)	--	--	--	0.12
Number of observations	68,134	56,285	124,419	124,419

Notes: The table reports estimates from a regression that examines whether the death rate of the same cohort changed over time differentially by gender. The sample in column (1) includes children born between 1989-1991 who appear in rounds 1 or 2 of the survey. The sample in column (2) includes children born between 1996-1998 who appear in rounds 2 or 3 of the survey. The sample in columns (3) and (4) includes the union of columns (1) and (2). In addition to the listed variables, all models control for state and year of birth fixed effects. Observations are weighted using national-level weights. Standard errors clustered at the state level are reported in parenthesis.

Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Using Stock Returns to Identify Government Spending Shocks — **WP-09-03**
Jonas D.M. Fisher and Ryan Peters

Stochastic Volatility — **WP-09-04**
Torben G. Andersen and Luca Benzoni

The Effect of Disability Insurance Receipt on Labor Supply — **WP-09-05**
Eric French and Jae Song

CEO Overconfidence and Dividend Policy — **WP-09-06**
Sanjay Deshmukh, Anand M. Goel, and Keith M. Howe

Do Financial Counseling Mandates Improve Mortgage Choice and Performance? Evidence from a Legislative Experiment — **WP-09-07**
Sumit Agarwal,Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff

Perverse Incentives at the Banks? Evidence from a Natural Experiment — **WP-09-08**
Sumit Agarwal and Faye H. Wang

Pay for Percentile — **WP-09-09**
Gadi Barlevy and Derek Neal

The Life and Times of Nicolas Dutot — **WP-09-10**
François R. Velde

Regulating Two-Sided Markets: An Empirical Investigation — **WP-09-11**
Santiago Carbó Valverde, Sujit Chakravorti, and Francisco Rodriguez Fernandez

The Case of the Undying Debt — **WP-09-12**
François R. Velde

Paying for Performance: The Education Impacts of a Community College Scholarship Program for Low-income Adults — **WP-09-13**
Lisa Barrow, Lashawn Richburg-Hayes, Cecilia Elena Rouse, and Thomas Brock

Establishments Dynamics, Vacancies and Unemployment: A Neoclassical Synthesis — **WP-09-14**
Marcelo Veracierto

The Price of Gasoline and the Demand for Fuel Economy: Evidence from Monthly New Vehicles Sales Data — **WP-09-15**
Thomas Klier and Joshua Linn

Estimation of a Transformation Model with Truncation, Interval Observation and Time-Varying Covariates — **WP-09-16**
Bo E. Honoré and Luojia Hu

Self-Enforcing Trade Agreements: Evidence from Antidumping Policy — **WP-09-17**
Chad P. Bown and Meredith A. Crowley

Too much right can make a wrong: Setting the stage for the financial crisis — **WP-09-18**
Richard J. Rosen

Working Paper Series *(continued)*

Can Structural Small Open Economy Models Account for the Influence of Foreign Disturbances? **WP-09-19**
Alejandro Justiniano and Bruce Preston

Liquidity Constraints of the Middle Class **WP-09-20**
Jeffrey R. Campbell and Zvi Hercowitz

Monetary Policy and Uncertainty in an Empirical Small Open Economy Model **WP-09-21**
Alejandro Justiniano and Bruce Preston

Firm boundaries and buyer-supplier match in market transaction: IT system procurement of U.S. credit unions **WP-09-22**
Yukako Ono and Junichi Suzuki

Health and the Savings of Insured Versus Uninsured, Working-Age Households in the U.S. **WP-09-23**
Maude Toussaint-Comeau and Jonathan Hartley

The Economics of "Radiator Springs:" Industry Dynamics, Sunk Costs, and Spatial Demand Shifts **WP-09-24**
Jeffrey R. Campbell and Thomas N. Hubbard

On the Relationship between Mobility, Population Growth, and Capital Spending in the United States **WP-09-25**
Marco Bassetto and Leslie McGranahan

The Impact of Rosenwald Schools on Black Achievement **WP-09-26**
Daniel Aaronson and Bhashkar Mazumder

Comment on "Letting Different Views about Business Cycles Compete" **WP-10-01**
Jonas D.M. Fisher

Macroeconomic Implications of Agglomeration **WP-10-02**
Morris A. Davis, Jonas D.M. Fisher and Toni M. Whited

Accounting for non-annuitization **WP-10-03**
Svetlana Pashchenko

Robustness and Macroeconomic Policy **WP-10-04**
Gadi Barlevy

Benefits of Relationship Banking: Evidence from Consumer Credit Markets **WP-10-05**
Sumit Agarwal, Souphala Chomsisengphet, Chunlin Liu, and Nicholas S. Souleles

The Effect of Sales Tax Holidays on Household Consumption Patterns **WP-10-06**
Nathan Marwell and Leslie McGranahan

Gathering Insights on the Forest from the Trees: A New Metric for Financial Conditions **WP-10-07**
Scott Brave and R. Andrew Butters

Identification of Models of the Labor Market **WP-10-08**
Eric French and Christopher Taber

Working Paper Series *(continued)*