NBER WORKING PAPER SERIES

CHILD GENDER AND PARENTAL INVESTMENTS IN INDIA:
ARE BOYS AND GIRLS TREATED DIFFERENTLY?

Silvia H. Barcellos
Leandro Carvalho
Adriana Lleras-Muney

Child Gender And Parental Investments In India: Are Boys And Girls Treated Differently?
Silvia H. Barcellos, Leandro Carvalho, and Adriana Lleras-Muney
NBER Working Paper No. 17781
January 2012
JEL No. I15,J16

<u>ABSTRACT</u>

Although previous research has not always found that boys and girls are treated differently in rural India, son-biased stopping rules imply that estimates of the effect of gender on parental investments are likely to be biased because girls systematically end up in larger families. We propose a novel identification strategy for overcoming this bias. We document that boys receive significantly more childcare time than girls. In addition boys are more likely to be breastfed longer, and to be given vaccinations and vitamin supplementation. We then present suggestive evidence that the differential treatment of boys is neither due to their greater needs nor to the effect of anticipated family size.

Silvia H. Barcellos
Rand Corporation
1776 Main Street
Santa Monica, CA 90401
silvia@rand.org

Leandro Carvalho
Rand Corporation
1776 Main Street
Santa Monica, CA 90401
carvalho@rand.org

Adriana Lleras-Muney
Department of Economics
9373 Bunche Hall
UCLA
Los Angeles, CA 90095
and NBER
alleras@ECON.UCLA.EDU

I. Introduction

Women in developing countries fare worse than men in many dimensions: they obtain less schooling, have lower labor force participation, earn lower salaries, are more likely to be poor and often lack fundamental rights, such as voting rights or the right to earn property (Duflo 2005). One often cited extreme manifestation of this phenomenon is that mortality rates are substantially higher for girls than for boys in many developing countries (Chen, Huq and D'Souza 1981; Arnold, Choe and Roy 1998, Sen 1990), although this is not true in the Western World (United Nations Secretariat 1988). These patterns are particularly marked in countries with "son preferences" such as India, where families have explicit preferences for having sons over daughters (Pande and Astone 2007).

Surprisingly, though, the previous literature does not always support the hypothesis that these differences in outcomes are due to explicitly different treatment of boys and girls. Although many papers find that boys receive more healthcare (Basu 1989, Ganatra and Hirve, 1994), are breastfed for longer (Kuziemko and Jayachandra 2010), and are more likely to be vaccinated (Borooah 2004) than girls, others find no evidence of differential investments. For example, Hariss (1995) finds that girls in India receive just as much nutrition as boys, and Deaton (2003) reports that vaccination rates are identical for boys and girls in India. Most notably, Deaton (1997) reviews studies that use the "adult goods method" and finds that there is no evidence of parents spending more on boys than girls.[1] Duflo (2005) concludes that"[e]*ven in the countries where the preference for boys is strongest, it is hard to find evidence that girls receive less care than boys under normal circumstances.*"[2]

However, previous work has assumed that boys and girls live in families with similar characteristics, in terms of both observables and unobservables. Yet this assumption is incorrect if families have a preference for sons and follow male-biased stopping rules of childbearing, (Yamaguchi 1989, Jensen 2005) which appears to be the case in India.[3] As a consequence, these empirical estimates of differential treatment are biased. In particular, if couples' fertility is driven by their desire to have a certain number of boys, then girls will end up in larger families on average. If in turn children in larger families have fewer per capita resources, as hypothesized by

---

[1] In more recent work, Kingdon (2005) finds parents spend more on boys than girls.
[2] Households do favor boys in bad times (Bherman 1988, Rose 2000, Miguel 2005, Maccini and Yang 2009).
[3] For example, families with fewer boys have shorter birth intervals, are more likely to want more children and to continue having children, and are less likely to use contraception (see Clark 2000 for a review).

Jensen (2005), then estimates of differential treatment will be biased upwards: it will appear as if girls get less on average but in fact this is due to the larger family size (and thus lower per capita resources), not to parental behavior. On the other hand if there are returns to scale, then estimates of differential treatment will be biased downwards.

We propose a novel empirical strategy that addresses this issue. It relies on the observation that—in the absence of sex-selective abortion—a child's sex at birth is randomly determined. If that is the case, then families who just had a boy are identical to families who just had a girl. So any differences we observe in terms of parental inputs can be attributed to the sex of the newborn. However, a correlation will develop over time between the youngest child's gender and the family characteristics, because families with a newborn daughter are less likely to stop having children. To overcome this problem, we restrict our sample to families with children who are still "young enough" whose mothers haven't had the opportunity to have other children. Our data suggest that families with boys and girls between zero to 15 months of age (and possibly a bit older) look identical in terms of observables—we use them to study whether boys receive more inputs than girls.

Our analysis allows us to rule out that observed differences in investments are driven by family size; this is important because it affects how one would design policies to improve the lot of girls. If girls get less because they live in poorer larger families, then transfers to those families would help girls; so would policies to reduce family size. If on the other hand parents would like to devote more resources to boys no matter what, then transfers to the same families might not help girls. In that case female-focused interventions might be needed.

The second contribution of this paper is to use our identification strategy to investigate whether boys and girls are treated differently in terms of an important but not frequently studied type of investment in children: *childcare time*. Beginning with Becker (1965), economists have recognized that, in addition to money, time is a key input into the "child production function." Time is particularly important to the extent that it is complementary to many other inputs. For example, feeding children requires both food and the time to cook it and feed the children. However no estimates of gender differences in parental time allocation based on dedicated time-

use surveys exist for developing countries.[4] Using data from the Indian Time Use Survey, we are able to investigate whether families spend more time on childcare after the birth of a son rather than a daughter. We also study gender differences in other frequently used measures of parental investments such as vaccinations, using data from the Indian Demographic and Health Survey (hereafter DHS). Following the literature suggesting that gender differences in India are present mostly in rural areas, our main results focus on rural households. However we also present results for urban areas.

Our results indicate that families treat boys and girls differently. Rural households with an infant boy under the age of one spend roughly 30 minutes more per day—or 14% more time—on childcare than households with an infant girl. The quality of the childcare also appears to be higher for baby boys. This gender difference exists for different types of childcare, including both supervision and physical care. The effect is larger for households with only one child under the age of six: they spend more than 60 minutes more per day (about 30% more) on childcare when their youngest is a boy. Our results also show that boys are more likely to be vaccinated, to be breastfed longer, and to be given vitamin supplements. In general we find these inputs to be at least 10% higher for boys than for girls in rural areas. We do not find evidence that boys fare better than girls in terms of anthropometric measures, but anthropometrics measure outcomes rather than inputs. We discuss various explanations for this one anomalous result.

Overall, our results suggest that the bias associated with family size depends on the timing of the parental investments. The evidence is consistent with a large bias for investments that occur throughout life such as time inputs: we observe that the gender difference in parental time inputs disappears as children age. On the other hand, the bias is small for investments that only occur early in life, such as vaccinations: for these inputs our results are very similar for young and older children. This is what we would expect because these investments are done once early in life and do not change much as children age.

The main limitation of our identification strategy is that it does not explicitly deal with the possibility that *anticipated* family size is responsible for the effects we observe: children in families who just had a girl may receive fewer parental investments if parents anticipate that they will have another child. We address this issue in various ways. Although no approach is perfect,

---

[4] Yeung et al (2001), Lundberg et al (2007), and Mammen (2009) report that fathers spend more time with boys in the US. Rose (2000) reports that in rural India women work fewer days after the birth of a boy than a girl.

the data suggest that breastfeeding is the only input for which anticipated family size is a plausible explanation. This is consistent with Jayachandran and Kuziemko (2010).

Our approach has other limitations. As in previous work, we cannot fully address the issues of sex-selective abortion and differential mortality. These behaviors most likely will bias our estimates of boy-girl differences towards zero, so our effects can be taken to be lower bounds. To limit the potential bias from sex-selective abortion we limit our study to the 1992 DHS, the most recent DHS survey conducted before ultrasound technology became widely available (Bhalotra and Cochrane 2010). We show that for the 1992 survey, child gender is uncorrelated with prenatal characteristics (e.g., prenatal care). This provides further evidence that families did not know the sex of the child and were not able to discriminate before birth, neither in terms of prenatal investments nor through sex-selective abortion.[5] Our results also suggest that the bias associated with sex-selective abortion in our sample is small. This is consistent with Anderson and Ray (2010), who report that prenatal factors account for only 10% of missing women in India.

To assess the bias caused by postnatal mortality, we compute bounds for our estimates and find that mortality potentially generates large biases: differences between boys and girls could be as much as 50% larger than our baseline estimates. Another limitation of our results is that we can only study children who are under the age of two. This is an important subset of the population, because at this age investments have large returns in the short and long run: lower investments in childhood are associated with worse health and economic outcomes in adulthood, and for future generations (Almond and Currie, forthcoming). But we cannot study older children.

Finally we investigate some possibilities for why parents give girls fewer resources. Parents might prefer boys to girls. Investments in boys might yield larger returns (e.g., men have higher wage rates than women). And boys might be seen as needing more resources, though it is not clear why parents would perceive that boys need greater inputs than girls, given that girls have higher mortality than boys from age 6 month onwards. We provide suggestive evidence that boys do not in fact "need" more than girls: if we look at South Africa, a developing country with

---

[5] Bharadwaj and Nelson (2011) use later waves of the Indian DHS (1998-9 and 2005-6) and show that mothers visit antenatal clinics and receive tetanus shots more frequently when pregnant with a boy. They also present evidence that there are no gender differences in prenatal care in the 1992 DHS data, which suggests that ultrasound use was not prevalent at this time period.

data on investments and no evidence of a son preference, we find that there are no systematic gender differences in most inputs. We also find that boy-girl differences in the north of India are larger than in the south. Unless boys need more in the north than in the south, this again does not support differential needs as the single explanation for our findings. Although the evidence is not conclusive, it does suggest that higher returns, or preference for boys, drive the differential investments.

II-Identification issues in the presence of son-biased stopping rules

In this section, we present a simple model of a son-biased fertility stopping rule and extend the previous results by Yamaguchi (1989) and Jensen (2005) to show that, if families follow this rule, then estimates in the boy-girl discrimination literature that assume that boys and girls live in families with similar characteristics are biased. We then propose a method for overcoming the problems that arise in this context.

We begin by presenting suggestive evidence that families in rural India do follow son-biased stopping rules. This is the case if *ceteris paribus* the probability of a family having no more children after a boy is born is higher than after a girl is born. One implication of this behavior is that the probability of a family's youngest child being a boy is an increasing function of the age of the youngest child. At birth, the sex ratio is determined by biological odds. But as the youngest child ages, the sex ratio is increasingly skewed towards boys since families are more likely to stop having children after a boy is born. This prediction is in fact consistent with the data.

In Figure 1, we plot the fraction of boys by age using data from the 1992 DHS (described in greater detail below) . The figure shows that the fraction of boys among all living children is somewhat constant across ages. But among the youngest child in the family, the fraction of boys increases from 51% for children 0-5 months old to 58% for the children 54-to-59 months old. In other words, if a child is still the youngest at age 4, then the probability that the child is a boy is 58%. This is a large deviation from the natural sex ratio at birth. This evidence suggests that in rural India families indeed follow son-biased stopping rules.[6]

---

[6] This pattern also could be driven by excess girl mortality. To gauge its importance, we compute the fraction of boys among all youngest children (including those who died according to the mother) and compare it to the fraction of boys among those alive. The graph suggests that there is excess girl mortality, since the fraction of boys is higher

We use the framework from Yamaguchi (1989) to illustrate how these stopping rules can bias estimates of differential treatment. There is a continuum of families and they all follow the same stopping rule: families continue to have children until they reach their desired number of sons. Families differ in how many sons they wish to have but they have no preference as to family size. The probability of a newborn being a boy is assumed to be constant across families.

In this stylized model, one can show that boys and girls live on average in families with different characteristics. As shown by Yamaguchi (1989)—and by Jensen (2005), who extends the results to a finite fertility set-up—girls will have more siblings than boys on average. A simple example provides the intuition for this result. Consider a family that wants one boy. If the first-born is a boy, then the family stops having children; if the first-born is a girl, the family continues having children. If all families behave this way, then all girls have siblings, but not all boys do. The simulations in Jensen (2005) suggest that the resulting differences in number of siblings can be quite large.

One can also show (see Appendix 1) that if we compare children in families of the same size, girls are on average in families that desire fewer sons than does the family of the average child. The intuition for this result is as follows: suppose that we observe two families who stopped having children after their second child. Family A has a girl and a boy; family B has two boys. Family A stopped having children despite the fact that they have only one boy; family B stopped because they had two boys, but otherwise would have continued. This example illustrates that for families with two children, girls live in families that desire fewer sons than does the average family. As a consequence, within families of the same size, those with a larger number (share) of boys have a larger desired number of sons than those with many girls.

In Yamaguchi's stylized model, families always obtain their desired number of sons, so we would learn their preferences simply by observing the number of sons they have. In reality, because families have finite fertility and imperfect fertility control, the desired number of sons is unobservable. So even if we control for observed differences in family size and gender composition, there will be unobserved differences in the families into which the average girl and average boy are born. If families that desire a larger number of sons invest less in girls (or more

among the survivors. However, the extent of the bias is small relative to the effect of stopping rules. This is confirmed by the pattern that we observe among all children (rather than the youngest): the fraction of boys rises for this group but the increase is small, much smaller than what is observed among the youngest child.

on boys) than other families, then these stopping rules imply that the previous estimates of discrimination are biased, as we now discuss.

Suppose that we estimate boy-girl differences by regressing some measure of child investment on a constant and a boy dummy (as in Sen and Sengupta 1983; Das Gupta 1987; Sommerfelt and Piani 1997):

$$Z_{ih} = \alpha_0 + \alpha_1 * B_{ih} + u_{ih},$$

where Z_{ih} is investment in child i in household h, B_{ih} is a dummy that is equal to one if child i in household h is a boy, and u_{ih} is an error term. Son-biased stopping rules imply that B_{ih} is correlated with family size. Therefore, B_{ih} will also be correlated with the error term, and α_1 will be biased if child investment depends on the number of children in the family. The sign of the bias may be different for different measures of child investment. On the one hand, children in large families may have to share resources with more siblings (e.g., food)—this is the issue that Jensen (2005) investigates. On the other hand, children in large families may *ceteris paribus* receive more investments if there are large returns to scale for the child's investment (e.g., vaccination in public campaigns, or supervision and teaching).

Given that girls tend to be part of larger families than boys, it may seem reasonable to control for family size. Suppose then that we estimate the following model (Oster 2009):

$$Z_{ih} = \alpha_0 + \alpha_1 * B_{ih} + X_{ih}\rho + u_{ih},$$

where we are now controlling for X_{ih}, a vector that includes the number of siblings or dummies for the sex-composition of siblings. This strategy essentially compares outcomes of boys and girls in families of the same size. However son-biased stopping rules imply that, conditional on family size, girls tend to be in families that want girls more than other families do. In other words, the child's sex is not exogenous; it is correlated with parental preferences for the gender composition of children. A similar argument applies to studies that use the adult goods method championed by Deaton (1989). In general, the sign of the bias is unknown and depends on the relationship between preferences for the gender composition of children and the treatment of boys and girls (see Appendix 2). For example, if all families invest the same amount in boys but families who want boys invest less in girls, then OLS estimates of α_1 are biased downwards because the average girl is in a family that wants fewer boys; thus she receives more child investments than she would have had she been "assigned" to a random family. Controlling for

family size also poses other challenges, because it is related to many unobserved determinants of parental inputs (such as income) and therefore is potentially correlated with the error term.

This discussion raises the question of why we are interested in knowing how girls would be treated if they were randomly assigned to families: the fact is, they are not. Controlling for family size, they are disproportionately in families that want them relative to the average family. But knowing the extent to which families want to treat girls differently is important because it informs policy. Suppose for instance that, as suggested by Jensen, all of the differences we observe are driven by family size. Then policies that transfer income to large and/or poor families will benefits girls. However, if this is not the case and parents want to devote more resource to boys, then transfers to these families will not necessarily benefit girls. There are similar implications for family planning policies. For instance, policies that ban sex-selective abortion might in fact inadvertently hurt girls—this is precisely what Lin, Liu and Qiang (2010) find for Taiwan where female infant mortality increased after abortion was banned.

III. Empirical Strategy

Our empirical strategy relies on the observation that in the absence of sex-selective abortion the child's sex is randomly determined at birth. If that is true, then families who just had a boy are identical to families who just had a girl. Therefore, any differences we observe in terms of parental inputs can be attributed to the sex of the newborn. However over time this is no longer true: families that follow a son-biased stopping rule are more likely to stop having children after a boy. In time a correlation will develop between the youngest child's sex and preferences: families with N children that stop after the birth of a girl tend to like girls more than families with N children that stop after the birth of a boy. To overcome this problem, we restrict our sample to families in which the youngest child is "young enough" as determined by our data: we select our sample such that baby-boy and baby-girl families look identical in terms of their observable characteristics. Formally, we estimate whether boys and girls are treated differently using the following equation:

$$Z_{ih} = \alpha_0 + \alpha_1 * B_{ih} + X_{ih}\rho + u_{ih}.$$

The OLS estimate of α_1 is an unbiased estimator of the parameter of interest if the child's sex is exogenous (conditional on X) —i.e., $Cov(B_{ih}, u_{ih}/X)=0$. Our identifying assumption

is that the child's sex is exogenous at birth *for children who are young enough*. In the next section, we provide evidence that predetermined characteristics (in particular number and gender of siblings) are not correlated with gender for very young children. We also show that, as the model above predicts, this no longer holds true as the family's youngest child gets older. Notice that, if gender is indeed random, then we do not need to condition on any variables.[7] Conditioning on predetermined variables should have no impact on our point estimates and should reduce the standard errors (if these variables predict parental investments).

Our assumption may fail if there is sex-selective abortion against girls or excess girl mortality. We test this directly in the data by comparing the characteristics of families with a baby girl and a baby boy. Still it is possible that families differ in terms of unobservables. Sex-selective abortion and excess female mortality most likely bias our estimator against finding boy-girl differences: because the surviving girls are expected to be in families that like girls more than the average family, they should receive more care than they otherwise would have.[8] Thus, our estimates can be taken as lower bounds of the effect of gender in child investments.

IV- Testing Random Assignment and Selecting the Estimation Sample

To test whether the gender of the youngest child is uncorrelated with predetermined family characteristics, we restrict the sample to children who are the youngest in their families and estimate the following linear equation

$$I(boy_{ia} = 1) = X_i \beta_a + \varepsilon_{ia},$$

where the dependent variable is an indicator of whether child *i* in age category *a* is a boy, and *X* is a set of predetermined characteristics. Independence implies that $\beta_a = 0$, namely that the *Xs* do not jointly predict the gender of the child. The prediction is that we will not reject the null for very young children, but that we will always reject it for children that are "old enough."

We use India's 1992 DHS (also known as the National Family Health Survey), a large representative survey that contains several variables determined before birth. The DHS surveyed

---

[7] There is evidence suggesting that the sex ratio at birth may be correlated with birth order, parental age, mother's education and marital status (Almond and Edlund 2007 and Chahnazarian 1988). But these effects are very small and can only be detected using very large samples of births (Yamaguchi 1989, Almond and Edlund 2007).
[8] The estimator also could be biased upwards. Girls that survive might be healthier than boys and thus need less care than boys. But this seems unlikely, because the mortality rates for girls remain higher than the mortality rates for boys for the entire postnatal period.

ever-married women of reproductive ages, each of whom was interviewed separately and asked questions about her characteristics and reproductive history. The DHS files contain full birth histories: there is a record for every child born, including date of birth and gender, whether the child has died, and whether s/he continues to live at home. For every child born we know the characteristics of the mother, and we can compute the number of siblings by gender and age (including the number of those who have died). We use only the 1992 survey to minimize the bias attributable to sex-selective abortion: the previous literature suggests that ultrasound technology became widespread in India only in the mid 1990s, particularly after 1995 (Bhalotra and Cochrane 2010). We also focus on rural households, as the previous literature has done.[9] The final data set contains one observation per family and includes children (excluding twins[10]) born to women ages 15 to 49 living in rural areas in 25 states.

Next we pool children into 12-month age-groups and run a joint test for each age-group.[11] We use these results to determine at which age the test starts to systematically reject the null. Table 1 shows all of the predetermined characteristics of the child and the mother that we can include, 20 in all. We selected variables that were most likely to be *predetermined before birth* for *all* children, regardless of their age. Therefore, we do not include location, household composition, marital status, spouse characteristics, or number of living siblings, because some of these characteristics may be affected by the gender of one's children (in fact we later document that household composition is affected by the gender of the youngest). However we do look at family size and composition at birth, maternal characteristics, and use of prenatal care. Prenatal care use allows us to investigate whether ultrasound technology was important in 1992: access to and use of prenatal care should predict the gender of the baby only if individuals know that before birth and if they differentially treat boys and girls in terms of prenatal variables. If this is not the case, then sex-selective abortion—a more radical type of differential treatment—should be less of a concern. However prenatal care variables are not available for older children (only for those under age four).

---

[9] Most papers looking at gender discrimination also concentrate on rural populations, e.g. Sen and Sengupta (1981), Rosenzweig and Schultz (1982), Behrman and Deolikar (1989), Subramaniam and Deaton (1991), Ganatra and Hirve (1994), Subramaniam (1996), Rose (2000), Pande (2003), Pande and Astone (2007), Oster (2009)

[10] We exclude twins so that we can define the family's youngest child's sex.

[11] We pool children into age groups in order to minimize the likelihood that we do not reject the null because of small sample sizes.

Figures 2a and 2b display the results of our test graphically. They plot the p-value of the joint test that the *X*s do not predict the gender of the youngest child. Figure 2a includes all children up to age five. For this full sample, we include all predetermined characteristics except prenatal care, which is not available for the older children. Figure 2b repeats the exercise including prenatal care use, and considers children up to age four only. The first point in either figure corresponds to children 0-to-11 months old. Starting with Figure 2a, we observe that for the youngest group (the first p-value reported), we cannot reject the null that families whose youngest is female look the same in terms of observable characteristics as families whose youngest is a boy. For living children, we can reject the null at the 5% level for the first time for the age group 19-30 months. Thereafter, we reject the null often. When we include prenatal care (Figure 2b) we reject the null at the 10% level for the first time for children ages 17-28 months. Based on these results, and to be conservative, we keep children all ages 0-15 months for our analysis.

Table 1 shows the results of our tests in more detail, for both our final estimation sample and older children. For each predetermined characteristic, we test whether the means are the same for families whose youngest is a boy versus those whose youngest is a girl. At the bottom of the table we report the p-value from the joint test that all characteristics predict gender. For the sample of the youngest children, no coefficient is significant at the 5% or 10% level, and the joint test cannot reject the null that all characteristics do not predict gender. This is true whether we perform the test for the subset of prenatal care variables or include the other characteristics as well. For comparison, we report in the last two columns the result of the tests for the youngest children who are ages 16-47 and 48-59 months. For the 16-47 group we have all the same covariates as your our sample, for the 48-59 group we have fewer predetermined characteristics. For the children 16-47 months, two of the variables we examine are statistically different at the 5% level, even though the joint p-test does not reject the null. Interestingly, we now observe that if the youngest is male, he is more likely to have more sisters. This is consistent with son-biased stopping rules. He is also more likely to have a mother who speaks Hindi, a characteristic that predicts son preferences (Pande and Astone 2007). Finally, for the oldest children (for whom we do not observe prenatal care), we see that four out of 14 of the predetermined family and mother characteristics are statistically significant predictors of gender. The joint test rejects the null, despite the fact that this sample is substantially smaller.

12

Note that if the bias from sex-selective abortion were large enough, even our youngest sample would not appear to be balanced between boys and girls.[12] Also given that prenatal care use does not predict gender, we conclude that there is not enough use of ultrasound technology and sex-selective abortion in our data to matter. This is consistent with Bharadwaj and Nelson (2011) who find no gender differences in prenatal care in the 1992 DHS but do find that boys receive more prenatal investments than girls in the 1998-9 and 2005-6 waves of the Indian DHS.

To assess the effect of excess mortality of girls on our results, we also plot in Figure 2 the p-value of the test for the sample of ever-born children (including children whose mothers reported them to have died by the time of the survey). Our results are basically the same, implying that the observed differences are mainly the result of stopping rules. However, as expected we would reject the null at the 10% level for living children for the first time at ages 17 months, but if we looked at *all* children we would reject the null starting at 19 months. Therefore excess mortality matters, but only among the oldest group.

Two caveats remain. First, as in other tests of random assignment, our test is imperfect because we can only observe that the samples are identical based on observables—it is possible that they are different based on unobservables. Second, although our samples are large, they are not large enough to precisely identify the age at which the covariates become unbalanced. In summary, the data support the assumption that gender is as good as "randomly assigned" among the youngest children, 15 months and younger. We use this sample to estimate whether girls receive fewer resources than boys, beginning with parental time.

V-Results from the Time Use Survey

We begin by investigating whether families spend more time taking care of children if their youngest child is a boy. We use data from the Indian Time Use Survey (hereafter TUS) conducted from July 1998 to June 1999 by the Social Statistics Division of the Central Statistical Organization of India. The TUS asked about the time use of all household members over five years of age during the previous 24 hours. The diary section was open-ended in terms of both describing the activities and giving beginning and ending times, with each activity identified as

---

[12] We performed another test of sex-selective abortion: we looked at whether the preceding birth interval was shorter for boys than for girls, but again we found small and statistically insignificant differences.

multiple (simultaneous) or not.[13] The survey collected data in six states chosen to be representative of the different regions of the country (Gujarat, Haryana, Madhya Pradesh, Meghalaya, Orissa and Tamil Nadu). There were 12,750 rural and 5,841 urban households interviewed, totaling roughly 75,000 respondents. Following the existing literature, we focus on rural households. We also analyze time-use data corresponding to "normal" days only (excluding holidays, etc).[14] The main variable of interest is the amount of time spent on childcare by household members over age five. We follow Guryan et al. (2008) as closely as possible, and classify the following activities as childcare: physical care of children (washing, dressing, feeding); teaching, training and instruction of own children; accompanying children to places; travel related to care of children; and supervising children.

These data have some limitations for our analyses. Aside from containing information on only six states, there is very little information about the participants. That effectively prevents us from reproducing our test that the predetermined covariates do not predict gender.[15] Also, families cannot be identified, only households. And we can only identify the youngest child in the *household* (not in the family): for this reason, we restrict the sample to those who are the children or grandchildren of the household head.[16]

The most important limitation of the TUS is that we do not know the identity of the child who was being cared for, we only know that individuals reported being occupied with childcare.[17] However this feature has one advantage: because the questions on childcare do not refer to a particular child, respondents are less likely to systematically bias their responses based on the gender of their youngest child.

Age in months is not available in the TUS. We look at children under age one for our main results; this group is closest to the experimental sample in the DHS. Because the TUS is small, and because there is substantial age-heaping at age one (which appears to differ by gender,

---

[13] The activities were coded into 176 different types. For simultaneous activities, field workers determined the main activity and distributed the total time spent according to the relative importance of activities.
[14] This excludes "abnormal" days when there are guests, someone is sick, or there is a festival, as well as "weekly variants", but most days are included. All households are interviewed for at least one normal day.
[15] We cannot reject the null in this sample either but we only have a few covariates. Results available upon request.
[16] Children who do not live with their biological parents receive less care on average, and it is possible that this differs by gender—for example, families are much more likely to adopt girls than boys. We restrict the sample to avoid these complications. We also exclude households with more than one child at the youngest age so we can define the sex of the youngest (if a boy and a girl are both aged three, we cannot tell who is the youngest).
[17] The survey did not ask the respondent who was present when an activity was performed.

according to Coale and Demeny 1967, and Bhat 1990), we also report results for children under age two. For comparison and to assess the bias among older children, we also report the results for children ages two-to-five years of age.

Table 2 presents summary statistics. Households with children under two on average spent more than three hours on childcare per day, while households with older children spent a little less than two hours.[18] Women provided more than 80% of the total time spent on childcare by the household. About 70% of childcare consisted of the physical care of children. During roughly half of the time devoted to childcare, the caregiver reported no simultaneous activity: we use this as a measure of the quality of childcare.

Preliminary evidence of differential treatment by gender is presented in Figure 3, which shows the cumulative distribution of childcare by gender of the youngest child under age one.[19] The baby-boy distribution appears to first-order stochastically dominate that of baby girls, suggesting that boys receive more childcare than girls. To obtain estimates of the effect of gender on childcare time, we estimate

$$Z_h = \alpha_0 + \alpha_1 * B_h + X_h \rho + u_h,$$

where Z_h is the total amount of time that all members in the household spent on childcare, and B_h is a dummy for whether the household's youngest child is a boy. We present the results with and without controlling for predetermined household-level covariates, X_h. The standard errors are estimated using White's correction for heteroskedascity, and we use the survey weights.[20]

The main results in Table 3 indicate that families spend more time on childcare when the youngest is a boy than when the youngest is a girl. The first column estimates a simple OLS model where the dependent variable is the total number of minutes spent on childcare, including zeroes. It shows that households where the youngest child is a boy spend roughly 32 minutes

---

[18] Although these numbers seem small, they are comparable to those from other countries. For example, Guryan et al. (2008) in Table 4 report that the average *weekly* childcare time for an adult with children ranges from four hours (South Africa) to about nine hours (US). Assuming that there are three adults per household on average this translates into roughly two (South Africa) to four (US) hours per day at the household level. The most likely reason why the numbers are so low is that individuals only report childcare when it is performed as a primary activity (exclusively)—previous research (Fedick et al. 2005) suggests that estimates of total childcare time are about three to four times larger when time spent with children (though not reported as childcare) is included.

[19] About 7% of households report spending no time (collectively) on childcare, even though they have an infant.

[20] We also estimated standard errors taking the survey design into account and found similar results (available upon request). The TUS had a sophisticated sampling scheme with three levels of stratification and clustering, and consequently there were many strata with one sampling unit. To account for all these features many assumptions have to be made, thus we opted for showing the OLS standards errors in the main tables.

more per day taking care of children than households whose youngest child is a girl, or about 14% more relative to the mean. Column 2 shows that this estimate is robust to controlling for religion, ethnicity, and the area of land that the household owns. In Column 3, we estimate a logit of whether the household spends any time on childcare. Although the estimates are positive, they are not statistically significant, and they are somewhat small (about 4%). If we estimate an OLS model instead for those that report *some* care, we find that households whose youngest child is a boy spend roughly 24 minutes more (about 10% more) per day on childcare than households whose youngest child is a girl. Column 5 estimates a Tobit model, which accounts for censoring at zero. Again, we find a statistically significant increase in childcare of about 15%.

Panels B and C investigate how these effects vary with age. The effects are similar, though smaller, for children under two (Panel B). Panel C presents the results for households whose youngest is between two and five years of age. Regardless of the specification, we do not find any statistically significant effect of gender. In fact, all of the coefficients have the "wrong" sign. Thus, we fail to find evidence of differential treatment among this older group. This suggests that either the estimates for the older children are biased because of changes in family size, or that the nature of parental time inputs is different among older children.

Table 4 looks at whether the effects of gender differ based on observable household characteristics and on the type of care. For reference purposes Column 1 reproduces our main estimates from Table 3. In Column 2, we interact gender of the youngest with the number of other children in the household under the age of six, which is also added as a control. The coefficient on gender is larger now, and the interaction with number of children is negative. If the youngest is the only child under six and a boy, then the household spends 44 minutes more on childcare. However there is no difference if there are four or more other children under the age of six. It would appear that when there are many small children, there is simply "no room" to provide differential treatment.

We observe the same pattern for the older sample (Panel B), for which all coefficients are significant. Column 3 restricts attention to families with no other children under age six: for these families, all childcare is directed towards the youngest (for our main results, we do not know who in the household is receiving the childcare). We find that infant boys receive 60 more minutes of care than infant girls (about 30% more).

16

In Columns 4-7, we show that households spend more time in both private and public childcare if the baby is a boy than if it is a girl (Columns 4 and 6). The amount of childcare time per child is increasing with the number of children for public care, but not for private (Columns 5 and 7). Most interestingly, gender differences in physical care do not decrease with the number of other siblings, whereas the effect of gender disappears for supervising if there are two additional children under age six. Again, we observe similar results for children under age two. These patterns can be explained by the private-versus-public nature of childcare activities. Because supervising is a public type of care, it makes sense that as the number of young children in the household increases, members will spend disproportionally more time in this type of care, and this time will not be closely related to the sex of the youngest child. In contrast, physical care is private, so there is room for differential treatment even when other young children are present.

Column 8 shows estimates of the effect of gender on "exclusive childcare time"—our proxy for quality care—defined as the number of minutes that adults spent caring for children and not doing anything else. Households whose youngest child is a boy provide more exclusive childcare than households whose youngest child is a girl: roughly 60 minutes more per day if there are no other children under age six. In Column 9, we repeat the estimation for urban households. The effect of gender is actually negative, but the sample is small, and the standard errors are large. We also investigate who in the household provides the care. All members report spending substantially more time on childcare if the youngest is a male, even though in general the estimates are significant only for adult women (Appendix Table 1).

Overall, we find that more time is spent on childcare in households whose youngest is a boy, and the quality of this time is higher.

VI-Gender differences in other inputs: additional results from the DHS.

VI.a. Effect of gender on inputs: main results

We now investigate whether there are boy-girl differences in other child investments, using the DHS data on breastfeeding, vitamin A supplementation and vaccinations. The results are reported in Table 5. All estimations use survey weights and correct the standard errors for survey design.

17

First we look at breastfeeding, which is deemed to be the ideal source of nutrition for infants, particularly in developing countries where food is in limited supply, and water and sanitation are poor.[21] We do not find that boys are more likely to have ever been breastfed (defined as ever breastfed, or breastfed for less than one month). This is true in both the linear and non-linear specifications, and regardless of whether we add controls. The effect sizes are precisely estimated zeroes. Most likely this is because more than 95% of children are ever breastfed.

In the next set of columns, we look at the duration of breastfeeding. We estimate censored linear regressions, because many children are still being breastfed at the time of the interview. Alternatively, we estimate a censored log-linear model and an accelerated-failure time model. We find a positive and statistically significant effect of gender on the duration of breastfeeding. If we estimate a proportional hazard model, we find that the odds of stopping breastfeeding are lower for males. The magnitudes suggest that breastfeeding duration increases as much as 40% when the child is a boy, which is consistent with Jayachandran and Kuziemko (2009).

Next we look at whether children are given Vitamin A, which protects against night blindness, measles and diarrhea.[22] Using either a linear or non-linear model, we find that boys are about 13% more likely to receive vitamin A.

Finally, we look at whether mothers have a vaccination card on hand at the time of the interview. Only about 28% of mothers have a vaccination card, but they are 4% more likely to have the vaccination cards of boys. For all outcomes, the results are not sensitive to the inclusion of covariates—the point estimates are almost identical, as one would expect if these characteristics are orthogonal to gender.

We also investigate whether boys are more likely to be vaccinated for specific diseases.[23] At the interview, mothers were asked first for the vaccination cards. If the mother had it, then all of the vaccination history was taken directly from the card. If the card was not available, then

---

[21] See Jayachandran and Kuziemko (2009) for a more detailed discussion of the benefits of breastfeeding in the context of developing countries.

[22] Children between six months and five years of age are supposed to take Vitamin A supplements every six months. The first two doses can be given at the same time that required vaccinations are given.

[23] The recommended vaccination schedule for children in India is as follows: BCG at birth; polio at birth, six weeks, 10 weeks, and 14 weeks; DPT at six weeks, 10 weeks, and 14 weeks; and measles at nine months. BCG protects against tuberculosis and DPT protects against diphteria, pertussis and tetanus.

mothers were asked to provide information on each type of vaccination. We use the information from either source. These results are shown in Table 6. The magnitudes vary depending on the vaccination but we find that boys are about 8 to 12% more likely to be vaccinated than girls. The results are not sensitive to the inclusion of covariates, the use of a non-linear model (Panel B), or whether we restrict the sample to those who are old enough in principle to have received all vaccinations already (Panel C). Oster (2009) and Jayachandran and Kuziemko (2009) find similar results.

However if we restrict the sample to children with vaccination cards (Panel D), then most estimates are small, some are negative, and all are statistically insignificant. However, these results are suspect for two reasons. First, the sample itself is substantially smaller (less than 30% of the original) and it is selected on the basis of gender—we already documented that mothers are more likely to have a card for boys. Moreover, the only study that we are aware of that looked at the quality of vaccination reports found that parental reports were in fact more accurate than vaccination cards (Babu et al. 2011). These results suggest why some previous research has not found large gender differences in vaccinations: some of it was based on vaccination cards (Borooah 2004) and other studies used surveys based entirely on mothers reports, which could be more (or less) reliable than the DHS data. This might be the case in the NSS, which Deaton (2003) uses to draw his conclusions.

VI.b. Does anticipated family size explain our findings?

The main threat to our identification strategy is the possibility that families change their expectations once the gender of the baby is revealed at birth. In particular, parents who desire boys but have a girl might anticipate the need to have more births in the future: they might start saving, or go back to work earlier (Rose 2000), which could explain why girls receive fewer investments. Unfortunately, there is no data on anticipated family size. Therefore we assess its importance in two ways. First, we use a proxy for anticipated family size to see how it affects our results. Alternatively, we estimate the effect of family size itself on inputs in order to assess how omitting anticipated family size might affect our results.

We start by using the question "would you like to have another child or would you prefer not to have any more children?" as a proxy for anticipated family size. Table 7 Panel A shows that if the youngest child is a boy then mothers are less likely to want more children—this

suggests that anticipated family size indeed responds to the gender of the last born. In fact the effect is in fact quite large: about 41% of women want more children, but this falls by 13 percentage points (a 30% decrease) if the youngest is a boy. The effects are the same if we add controls. In Panel B, we reproduce our results controlling for this variable. We find that including this variable as a control has virtually no effect on our coefficients of interest. The point estimates are identical to those in Panel C (which report the baseline results for the estimation sample in this table, which includes only those with older sibs for comparison with the next exercise) with one exception: the coefficients for gender on breastfeeding duration increase (although given the standard errors we would not reject equality).

However wanting more children is not random; it is jointly determined with investments. To obtain unbiased estimates, we follow the previous literature (see Schultz 2008) and instrument for wanting more children using the gender of the first born: in families with son-preference, a first born girl should increase the desire for more children. These families are less likely at any point in time to have achieved their optimal number of boys relative to those families whose first born is a boy. The gender of the first born is indeed a good predictor of the desire to have more children: if the first born is a male, then mothers are less likely to want more children: the coefficient (reported in Panel D) is negative and statistically significant—the point estimate is 0.082 (standard error 0.013). When we instrument for wanting more children, we find our results are essentially unchanged for most inputs (panel C). For having a vaccination card, breastfeeding, [24] and for all of the vaccinations, the coefficients on youngest is male remains the same. The coefficients on Vitamin A fall a bit, but the standard errors are large.

An alternative way to gauge the bias created by the (unobserved) anticipated family size is to estimate the effects of family size directly and to assume that the effects of an *anticipated* increase in family size are similar to the effects of an increase in *actual* family size. The results are shown in Appendix Table 2. A simple OLS regression of inputs on family size shows that children in larger families receive fewer inputs. However these results are difficult to interpret because these families also might be poorer. To obtain better estimates, we again instrument for family size using the gender of the first born. We find that the gender of the first born is a good

---

[24] We cannot estimate 2SLS models for duration of breastfeeding, so the Table reports the reduced-form effect of youngest gender and the gender of the first born.

predictor of family size: if the first born is a boy, then family size decreases by about 4% (.13 fewer children).

The 2SLS estimates of the effects of family size on inputs reveal a different picture from the OLS estimates.[25] We find that family size significantly lowers both the incidence and duration of breastfeeding.[26] But all of our other estimates are statistically insignificant. Noticeably, the effect of family size is positive for most vaccinations, probably because there are returns to scale. Therefore we infer that anticipated family size might explain the effects of gender on breastfeeding, but not the effects we observe for other inputs.[27] Thus overall we find very little support for the idea that anticipated family size explains our results.

VI.c. Additional robustness checks

In Appendix Tables 3 and 4, we perform a number of additional robustness checks. The first columns reproduce our main results for reference.

In Appendix Table 3, we report estimates limiting the sample to first-born children— previous literature suggests that sex-selective abortion is less important among first-born (Retherford and Roy 2003). On the other hand, previous literature also suggests that discrimination against girls increases with birth order (Das Gupta 1987), so it is not entirely clear a priori what to expect in this sample. Nevertheless, we find that even among first born, boys appear to receive more inputs, although the magnitudes are smaller and not always significant (however this sample is substantially smaller).

Next we investigate the effect of mortality on our estimates. One advantage of the 1992 DHS data (unlike later waves) is that mothers were asked to report on investments even for children who had died before the interview. Assuming that maternal reports aren't gender biased, we can gauge the effect of mortality on our estimates by simply including these deceased children in our estimation sample. The results do not differ from our main results. Alternatively, we can compute bounds by imputing the missing information under best and worst case

---

[25] We also used twins as an instrument and found similar results, but these were more sensitive to the addition of covariates (results available upon request).

[26] We report reduced forms (instead of 2SLS) for breastfeeding duration because of censoring.

[27] Unfortunately, we cannot assess these effects for childcare time because the survey identifies only households— we have no information on the gender of the first born.

scenarios for children who died before the investment was possible[28] or for whom the maternal report was missing.[29] Our bounds are not very tight as many include zero. However if one assumes that only the upper bounds are likely (deceased girls were treated worse than deceased boys) then our upper bounds imply that our estimates could be substantially underestimated.

Appendix Table 4 shows the results for all children ages 16-47 months (regardless of whether they are the youngest). This sheds light on the extent to which using older children generates bias.[30] For this older sample, the coefficients are similar and still statistically significant, but they are a bit smaller relative to the sample mean than our main estimates. Comparing the magnitudes to those in the first columns (our main results), we can conclude that the bias due to son-biased stopping rules and family size appears to be small for vaccinations, but it is much larger for breastfeeding and for childcare. These results are consistent with our theory: vaccinations are acquired while children are still very young and are still the youngest in the family, whereas other inputs are received at older ages when households will have had a chance to respond with additional children.

Next we report the results for urban areas. Just as in the TUS, these results are smaller and generally insignificant, although most of the times still positive. Overall, our findings support the hypothesis that differential gender treatment is greater in rural areas.

In summary, for all of the measures we looked at, we find that boys are given more inputs than girls. In general, girls receive at least 10% less than boys. To assess the magnitude of these differences, we estimate how much gender differences in investments can explain the higher mortality rates among girls. We use estimates from the literature of the effects of breastfeeding, vitamin A supplementation and vaccinations on mortality.[31] Mortality rates among children 12 to

---

[28] We impute the information for polio/DPT 1st dose if the child died before two months, for polio/DPT 2nd dose before three months, for polio/DPT 3rd dose before four months, for vitamin A before six months and for measles before nine months.

[29] The upper bounds assume that all dead girls would have not received inputs (for dummy variables) or would have been given the 25th percentile of the girls' outcomes distribution. For boys, we assume that had they lived they would all have been given inputs (for dummy variables) or be given the 75th percentile of boys' outcome distribution. For upper bounds, we assume the opposite.

[30] We use all children, not just the youngest to reproduce the exercises that are typically done in the literature.

[31] For each investment, we first calculate the (gender neutral) probability of death conditional on not receiving the investment (p_0) and the (gender neutral) probability of death conditional on receiving it (p_1) using the relative risks estimated in the literature [vitamin A (Rahmathullah et al 2003); breastfeeding (Briend, Wojtyniak, and Rowland, 1988 and WHO 2000); measles (Koenig et al 1990); Polio, BCG and DPT (Moulton et al 2005)], the mortality rate for children 12 to 36 months old (20.3 per 1,000 children) and the fraction of children in this age group receiving the investment (see Jayachandran and Kuziemko 2009 for a more detailed discussion). Let θ_b be the fraction of boys and

36 months of age are 16.7 per 1,000 for boys and 24.2 per 1,000 for girls. A back-of-the envelope calculation suggests that the observed differences in investments (excluding time inputs) explain about 27% of excess girl mortality among children in this age group (or, about two additional girl deaths per 1,000 children).[32]

VII-Investigating reasons for differential investments

We begin by investigating whether boys appear to need more inputs from their parents—this could be the case if boys are more active, or if they get sicker more frequently. The mortality data from India is inconsistent with the view that boys in India need more, because girls have larger mortality rates than boys from about 6 months of age to well into adulthood.

We look at the medical literature to assess whether any of the inputs we study are known to provide greater benefits to boys than girls. The meta-analysis of RCTs performed by Beaton et al. (1994) shows that vitamin A supplementation has the identical proportional effect on mortality for boys and girls. Not much is known about other inputs from randomized trials. But the benefits of breastfeeding also appear not to vary by gender, or if anything females appear to benefit more.[33] BCG vaccines appear to benefit girls more (Roth et al. 2006), and so do measles vaccines (Koenig 1990). This evidence, though scant, does not support the idea of greater benefits for boys. Nevertheless, parents could still perceive greater benefits to providing particular inputs to boys. So we empirically investigate this hypothesis.

We perform two data exercises to assess how much of the observed differences in inputs could be driven by boys' greater needs. First we look at South Africa, the only developing country we are aware of with a dedicated time-use survey,[34] a DHS survey, and for which

---

θ_b be the fraction of girls (in the age group) who receive the investment. The difference in the mortality rates of girls versus boys associated to gender differences in the investment is equal to $(\theta_b - \theta_g)*(p_0 - p_1)$. We sum these differences over all investments and divide the total by the difference in mortality rates of girls and boys.

[32] We know of no good estimates of how parental time affects mortality. Assuming that one additional hour of childcare reduces the probability of death (in absolute terms) by 0.0003, the boy-girl difference in time use of roughly 51 minutes translates into a boy-girl difference in mortality of 0.000255. Thus we would further explain 3.4% of excess female mortality.

[33] Klein et al (2011) report that most studies of breastfeeding have not investigated whether the benefits differ by gender, but they find breastfed girls have greater protection against respiratory infections than boys.

[34] We use data from the South African Time Use Survey, conducted in 2000 by Statistics South Africa. Information on time use was collected for persons aged 10 years and above, with two respondents randomly chosen per household (or only one if there was only one household member aged 10 years or more). Data were collected for

fertility patterns suggest no son preference (Gangadharan and Maitra 2003). Second we use the South of India as a baseline for establishing needs: previous research has suggested that in the South preferences for boys are smaller than in the North, but there is no ex-ante reason to believe that needs of children are different in the South.

Beginning with the South African case, Figure 4 plots the cumulative distribution of household childcare time by gender of the youngest among children under the age of one. There is a small difference between the genders: households in which the youngest is a girl are more likely to report no care. Compared to the Indian TUS, South African mean childcare time is lower and a larger fraction of households report spending no time at all on childcare. These differences are easily explained: the Indian TUS collects diaries for all household members aged six and above, whereas the South African TUS only collects time use for one or two eligible members (above age 10).

In Table 8 we report the point estimates for the gender differences: we find that boys are more likely to get any care (the implied marginal effect is about 12%), but that conditional on getting care, girls appear to get more care than boys. This evidence does not strongly support the idea that boys need more childcare time. Table 8 also reports whether boys are given more of all other inputs. Most of the coefficients on the male dummy are statistically insignificant; furthermore, most coefficients are negative and small for most inputs, which suggests that there is no greater need among boys. Of course, this evidence is only suggestive, since it is not clear that South Africa provides a good counterfactual for India.

We look at the needs question in another way by comparing the effects of gender in the North and South of India. If we assume that children in the North and South have the same needs, then needs cannot explain any existing any north-south differences in the data. The South provides a closer counterfactual for needs than does South Africa. However since the time use data only includes six states, this exercise is not very informative for time inputs (there is only one state in that survey, Tamil Nadu, that falls in the traditional definition of South (results available upon request). These results are in Table 8. Interestingly, we do find positive and often significant effects of gender in the South, particularly for breastfeeding. But we also find that the interaction between gender and north is positive for all of our DHS outcomes (except

8,564 households (14,553 respondents). We use data from 521 households whose youngest member is under one year old.

24

breastfeeding), in spite of the fact that sex-selective abortion in the North most likely biases these estimates downward. Only a few of these interactions are statistically significant, but they are all positive, suggesting greater investments in the North for DPT, Polio and Vitamin A. Therefore, with the possible exception of breastfeeding, needs alone cannot account for the larger investments observed for boys.

Our results suggest that needs and anticipated family size cannot explain all of the effects of gender that we observe. We do not have enough information to assess whether parents invest less in girls because of lower returns, the other main reason for differential investments. However, there is growing evidence that this is indeed part of the explanation, as suggested in the seminal paper by Rosenzweig and Schultz (1982). Jensen (2010) and Oster and Millet (2011) document that in India, when the returns to schooling for women increase (as a result of the availability of higher paying jobs in female oriented call-centers), girls stay in school longer. Also Jayachandran and Lleras-Muney (2010) show in a different context that when female adult mortality declines, schooling of girls increases. Relatedly, Qian (2008) shows that in China girls' mortality and education improve when the price of female-intensive crops (and thus female income) rises. Overall, our findings point to either differences in returns or differences in preferences as the main reasons for lower investments in girls. Still is possible that households invest more in girls in dimensions that we cannot observe, such as dowries.

VIII-Other results from the DHS: anthropometric measures and living arrangements

We now look at the effect of gender on height-for-age, weight-for-age, and weight-for-height Z-scores. These are computed by subtracting the median of the reference population and dividing by the standard deviation of the standard population.[35] It is important to normalize outcomes because boys are known to be taller and heavier than girls. We further examine whether gender determines the likelihood of a child being stunted, underweight, or wasted.[36] Importantly, these measures are not ideal for investigating differential treatment. Anthropometric measures are

---

[35] About 15% of children were not measured but this did not differ by gender.
[36] A child is stunted if the height-for-age is two s.d. below the median of reference population (measures chronic under-nutrition); a child is underweight if the weight-for-age is two s.d. below the median (measures both chronic and acute under-nutrition); a child is wasted if the weight-for-height is two s.d. below (measures acute under-nutrition).

outcomes, not inputs, and parental control over them is limited. Height and weight are the result of caloric inputs, but also of other factors such as the incidence of disease and caloric expenditure, which may differ by gender for biological or other reasons.

Table 9 shows that boys fare *worse* than girls for all of the anthropometric measures we use if we rely on the Z-scores provided by the DHS (this is also what Mahajan and Tarozzi 2007 find). These results are hard to reconcile with the previous evidence presented here, that inputs are higher for boys. However, Sommerfelt and Arnold (1998) also find that girls under the age of two have better anthropometric measures than boys in almost all developing countries for which DHS data is available (41 surveys were used).

There are two possible explanations for these patterns. First, the results may be driven by the standardization method, as first pointed out by Thomas (1990). Indeed, when we use alternative standards (the 1990 British Standards for height-for-age, and weight-for-age and the 2000 CDC standards for weight-for-height[37]), the coefficients on male are substantially smaller, statistically insignificant, and sometimes they switch signs. It is possible that the standards affect the results because baseline differences in height and weight between boys and girls depend on the nutrition patterns of each country. Moestue (2009) also finds that boy-girl differences for Bangladesh depend on the standard used; Tarozzi (2008) also documents that the choice of standard significantly affects the results on anthropometrics by gender. Another possibility is that the effect of lower inputs for girls appears with a lag: girls are sturdier than boys at birth and this initial advantage is larger than the immediate effects of low parental inputs. Over time though, the effects of lower inputs eventually benefit boys. Indeed, if we look at older children (above age two) we find that boys have better anthropometrics than girls. We cannot resolve this puzzle here, but we note that anthropometric measures yield substantially different results from inputs.

Finally, we investigate whether living arrangements are affected by the gender of the youngest boy. Studies in the United States report that having a son reduces the probability of parents getting divorced (Katzev, Warner and Acock 1994; Morgan, Lye and Condran 1988; Mott 1994), and that daughters are less likely to live with their fathers (Dahl and Moretti 2004). However, there is little research for developing countries, where boy-girl discrimination is thought to be a greater concern. Table 10 reports our results. Panel A uses the DHS and asks

---

[37] There are no British standards available for weight-for-height. The 2000 CDC standards for height are not available for children under age two.

whether gender of the youngest affects the likelihood that different family members live together. We do not look at marital status as an outcome because all mothers are married in our samples. We do look at whether women report that their husbands live at home. We find that if the youngest is a boy, the husband is more likely to live at home. The effect is insignificant and very small, less than 1% though. There is no evidence that the gender of the youngest affects the likelihood of the mother being the household head's wife. That might occur if, for example, families move in with their parents after a boy is born. There is also no discernable effect of the gender of the youngest on the number of siblings living at home. Sisters are slightly more likely to live at home (though this is not significant) whereas brothers are less likely to live at home if the youngest is a boy (about 5% less, and this is statistically significant).

In Panel B we look at household composition in the TUS. Although the unit of observation is now the household, we find similar results. If the youngest is a boy, then there are more men over age 15 in the household (about 15% more) and this is statistically significant. There are also more women over age 15 (about 11%), and this effect is also statistically significant. Again there is no apparent effect on the number of total children under age 14, but they are more likely to be girls and significantly less likely to be boys (about 20%). Together, these results suggest that when the youngest is a boy, the family is more likely to retain daughters for caregiving, or to have another female adult move in to provide childcare and household help. In addition, the household is more likely to have adult males and is less likely to have male children.

VII-Conclusion

This study asks whether parents treat girls and boys differently in India. Although women in India lag behind men in many domains, there is equivocal evidence on whether these lower outcomes are the result of lower parental investments in girls, particularly because boys and girls live in household with different observed and unobserved characteristics as a result of son-biased stopping rules. We develop a novel empirical strategy to address this problem by looking at children while they are still very young, whose parents have not had a chance to respond to the gender of the last child. We then used our identification strategy to look at differential treatment along measures previously used in the literature. In addition, we examine whether families spend

27

more time with childcare when the baby is a boy than when the baby is a girl. Time investments have not been studied previously in the context of developing countries.

We find evidence that boys receive more investments than girls in rural India. Households with an infant boy under the age of one spend roughly 30 minutes more per day (about 15% more time) on childcare than households with an infant girl. This difference is even larger for one-child households: households with one boy under age six spend roughly 60 minutes more (30%) per day on childcare than households with one girl under age six. We also find suggestive evidence that the quality of childcare given to boys is higher. Moreover, we find that boys are more likely to be vaccinated, to be breastfed longer and to be given vitamin supplementation. In general we find these inputs to be at least 10% higher for boys.

We also investigate why parents may choose to invest less in girls. We find no evidence of greater needs among boys for all measures except possibly for breastfeeding and childcare time for which the evidence is mixed. We also look at whether girls receive less because families of recently-born girls anticipate that they will have to continue having more children. We find some evidence to suggest this is true for breastfeeding, but not for other outcomes. Thus, in general, we find that these explanations cannot account for the patterns we observe across all outcomes. We conclude that parents invest less in girls because these investments have lower returns (for which there is some evidence in the literature), or because they have a preference for sons.

References

Almond, D. and, L. Edlund. 2007. "Trivers-Willard at birth and one year: evidence from US natality data 1983-2001." *Proceedings of the Royal Society,* 274: 2491-2496.

Almond, D., and J., Currie. Forthcoming. "Human Capital Development Before Age Five." Handbook of Labor Economics, Volume 4.

Anderson, S. and D. Ray. 2010. "Missing Women: Age and Disease," *Review of Economic Studies* 77(4): 1262-1300.

Arnold, F., M. K. Choe, and T. K. Roy. 1998. "Son preference, the family-building process and child mortality in India." *Population Studies,* 52: 301-315.

Babu, G.R., J. Olsen, S. Jana, S. Nandy, M.N. Farid and Sadhana. 2011. "Evaluation Of Immunization Cards And Parental Recall Against Gold Standard For Evaluating Immunization Coverage." *The Internet Journal of Epidemiology*. Volume 9 Number 2.

Bhalotra, S. and T. Cochrane. 2010. "Where Are All The Young Girls Gone? Identifying Sex-Selection In India," unpublished manuscript.

Basu, A. M. 1989. "Is Discrimination in Food Really Necessary for Explaining Sex Differentials in Childhood Mortality?" *Population Studies* 43 (2), 193–210.

Beaton, George H., R. Martorell, K. A. Aronson, B. Edmonston, G. McCabe., A. C. Ross and B. Harvey. 1994. "Vitamin A supplementation and child morbidity and mortality in developing countries," *Food and Nutrition Bulletin*, Volume 15 (1993/1994), Number 4

Behrman, J. 1988. "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?" *Oxford Economic Papers* 40, no. 1:32–54.

Bhat, P. N. M. 1990. "Estimating Transitional Probabilities of Age Misstatement." *Demography*, 27(1): 149-63.

Borooah, Vani K. 2004. "Gender bias among children in India in their diet and immunization against disease." *Social Science and Medicine* 58(9): 1719 – 1731.

Becker G. S.1965. "A Theory of the Allocation of Time," *Economic Journal* 75:493-517.

Briend, A., B. Wojtyniak, and M. G. Rowland. 1988. "Breastfeeding, Nutritional State, and Child Survival in Rural Bangladesh." *British Medical Journal*, 296(6626): 879-882.

Chahnazarian, A. 1988. "Determinants of the sex ratio at birth: Review of recent literature." *Social Biology,* 35(3-4): 214-235.

Chen, L. C., E. Huq, and S. D'Souza. 1981. "Sex bias in the family allocation of food and health care in rural Bangladesh," *Population and Development Review,* 7: 55-70.

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

Coale, A. J. and P. Demeny. 1967. "Methods of Estimating Basic Demographic Measures from Incomplete Data." Manual IV, Population Studies No 42, United Nations, New York.

Dahl, G.B. and E. Moretti. 2004. "The Demand for Sons." *Review of Economic Studies,* 75(4): 1085-1120.

Das Gupta, M. 1987. "Selective discrimination against female children in rural Punjab, India." *Population and Development Review*, 13(1): 77-100.

Deaton, A. 1989. "Looking for Boy-Girl Discrimination in Household Expenditure Data." *World Bank Economic Review,* 3: 1-15.

Deaton, A. 1997. The analysis of household surveys: A microeconometric approach to development policy. Baltimore and London: John Hopkins University Press.

Deaton, A. 2003. " Health, Inequality, and Economic Development," *Journal of Economic Literature*, 41(1): 113-158.

Duflo, E. 2005. "Gender Equality in Development," MIT Working Paper.

Fedick, C. B., S. Pacholok and A. H. Gauthier. 2005. "Methodological Issues In The Estimation Of Parental Time – Analysis Of Measures In A Canadian Time-Use Survey," *International Journal of Time Use Research,* 2(1): 67-87

Gangadharan, Lata and Pushkar Maitra. 2003. "Testing for Son Preference in South Africa." *Journal of African Economies* 12(3): 371-416.

Ganatra, B. and S. Hirve. 1994. "Male bias in health care utilization for under-fives in a rural community in western India." *Bulletin of the World Health Organization,* 72(1): 101-104.

Guryan, J., E. Hurst and M.Kearney, 2008. "Parental Education and Parental Time with Children." *Journal of Economic Perspectives,* 22(3): 23-46.

Harriss, Barbara. 1995. "The Intrafamily Distribution of Hunger in South Asia." In *The Political Economy of Hunger: Selected Essays*, ed. Jean Drèze, Amartya Sen, and Athar Hussain. Oxford: Clarendon.

Jayachandran, S. and A. Lleras-Muney. "Life Expectancy and human capital investments: Evidence from Maternal Mortality Declines," *Quarterly Journal of Economics*, Vol. 124, No. 1: 349–397, February 2009.

Jayachandran, S. and Kuziemko, I. 2009. "Why Do Mothers Breastfeed Girls Less than Boys: Evidence and Implications for Child Health in India." NBER Working Paper 15041.

Jensen, R. 2005. "Equal Treatment, Unequal Outcomes? Generating Sex Inequality through Fertility Behavior." Unpublished Manuscript, Harvard University.

Jensen, R. 2010. "Economic opportunities and Gender Differences in Human Capital: Experimental Evidence from India," UCLA mimeo

Katzev, A. R., R. L. Warner, and A. C. Acock. 1994. "Girls or Boys? Relationship of Child Gender to Marital Instability." *Journal of Marriage and the Family,* 56: 89-100.

Kingdon, G. 2005. "Where Has All the Bias Gone? Detecting Gender Bias in the Household Allocation of Educational Expenditure in India", *Economic Development and Cultural Change*, 53(2): 409-451.

Klein M. I., Bergel E., Gibbons L., Coviello S., Bauer G., Benitez A., Serra M. E., Delgado M. F., Melendi G. A., Rodríguez S., Kleeberger S. R., Polack F. P.. 2008. "Differential Gender response to respiratory infections and to the protective effect of breast milk in preterm infants." Pediatrics, 121(6):e1510-6.

Koenig, M. A., M. A. Khan, B. Wojtyniak, J. D. Clemens, J. Chakraborty, V. Fauveau, J. F. Phillips, J. Akbar, and U. S. Barua. 1990. "Impact of Measles Vaccination on Childhood Mortality in Rural Bangladesh," *Bulletin of the World Health Organization* 4: 441–447.

Lin, Ming-Jen, Jin-Tan Liu and Nancy Qiang. 2010. "More Missing Women, Fewer Dying Girls: The Impact of Abortion on Sex Ratios at Birth and Excess Female Mortality in Taiwan," NBER Working Paper w14541.

Lundberg, S., S.W. Pabilonia and J. Ward-Batts. 2007. "Time Allocation of Parents and Investments in Sons and Daughters." Working paper. Department of Economics, University of Washington.

Maccini S. and D. Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, Vol. 99, No. 3, June 2009, pp. 1006-1026.

Mahajan A. and A. Tarozzi. 2007. "Child Nutrition in India in the Nineties," *Economic Development and Cultural Change.* 55(3): 441-486.

Mammen, K. 2009. "Fathers' Time investments in children: do sons get more?" *Journal of Population Economics.*

Marcoux, A. 2002. "Sex differentials in undernutrition: A look at survey evidence." *Population and Development Review*, 28(2): 275-284.

Miguel, E. 2005. "Poverty and witch killing." *Review of Economics Studies* 72(4): 1153-1172.

Morgan, S. P., D.N. Lye, and G. A. Condran. 1988. "Sons, Daughters, and the Risk of Marital Disruption." *American Journal of Sociology,* 94: 110-29.

Mott, F.L. 1994. "Sons, Daughters and Fathers' Absence: Differentials in Father-Leaving Probabilities and in Home Environments." *Journal of Family Issues,* 15: 97-128.

Moestue, H. 2009. "Can anthropometry measure gender discrimination? An analysis using WHO standards to assess the growth of Bangladeshi children." *Public Health Nutrition* 12(8): 1085-1091.

Moulton, L. H., L. Rahmathullah, N. A. Halsey, R. D. Thulasiraj, J. Katz, J. M. Tielsch. 2005. "Evaluation of Non-Specific Effects of Infant Immunizations on Early Infant Mortality in a Southern Indian Population." *Tropical Medicine & International Health* 10(10): 947-955.

Oster, E. 2009. "Does increased access increase equality? Gender and child health investments in India." *Journal of Development Economics*, 89(1): 62-76.

Oster, E. and B. Millet. 2010 "Do Call Centers Promote School Enrollment? Evidence from India," Mimeo University of Chicago.

Pande, R. and N. Astone. 2007. "Explaining son preference in rural India: the independent role of structural versus individual factors." *Population Research and Policy Review* 26(1): 1-29.

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

Rahmathullah, L., J. M. Tielsch, R. D. ThulasirajD, J. Katz, C. Coles, S. Devi, R. John, K. Prakash, A. V. Sadanand, N. Edwin N, and C. Kamaraj. 2003. "Impact of Supplementing Newborn Infants with Vitamin A on Early Infant Mortality: Community Based Randomised Trial in Southern India." *British Medical Journal*, 327(7409): 254-256.

Retherford R. D. and T. K. Roy. 2003. *Factors affecting sex-selective abortion in India and 17 major states.* NFHS Subject Reports, No. 21. Mumbai: IIPS; and Honolulu: East-West Center.

Roth, A., M L Garly, H Jensen, J Nielsen, and P Aaby. 2006. "Bacillus Calmette-Guérin vaccination and infant mortality." *Expert Rev Vaccines* 5(2): 277–293.

Rose, E. 2000. "Gender Bias, Credit Constraints and Time Allocation in Rural India." The *Economic Journal* 110: 738-758.

Rosenzweig, Mark R., and Paul T. Schultz. 1982. "Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India." *American Economic Review* 72, no. 4:803–15.

Sen, A. and S. Sengupta. 1983. "Malnutrition of rural children and the sex bias." *Economic and Political Weekly,* 18: 855-864.

Sen, A. 1990. "More Than 100 Million Women Are Missing." The New York Review of Books (December 20, 1990).

Sommerfelt, A. E. and A. L. Piani. 1997. Childhood Immunization: 1990-1994. DHS Comparative Studies No. 22. Calverton, MD: Macro International Inc.

Sommerfelt, A. E. and F. Arnold. 1998. "Sex differentials in nutritional status of young children," in Too Young to Die: Genes or Gender?" pp. 133-153. New York: Population Division, United Nations.

Stinson, S. 1985. "Sex Differences in Environmental Sensitivity during Growth and Development," *American Journal of Physical Anthropology*, 28(S6): 123-147.

Subramanian S. and Deaton, A. S. 1991. "Gender Effects in Indian Consumption Patterns." *Sarvekshana*, 14: 1-12.

Tarozzi, A. 2008. "Growth Reference Charts and the Nutritional Status of Indian Children," *Economics and Human Biology*, 6(3), 455-468.

Thomas, D. 1990. "Intra-Household Resource Allocation: An Inferential Approach," *The Journal of Human Resources,* 25(4):635-664

United Nations Secretariat. 1988. "Sex differentials in life expectancy and mortality in developed countries: An analysis by age groups and causes of death from recent and historical data." *Population Bulletin of the United Nations,* 25: 65-107.

World Health Organization. 2000. "Effect of Breastfeeding on Infant and Child Mortality Due to Infectious Diseases in Less Developed Countries: A Pooled Analysis." *Lancet*, 55(9202): 451-455.

Yamaguchi, K. 1989. "A Formal Theory for Male-Preferring Stopping Rules of Chilbearing: Sex Differences in Birth Order and the Number of Siblings." *Demography,* 26(3): 451-65.

Yeung, W.J., J.F. Sandberg, P.E. David-Kean and S.L. Hofferth. 2001. "Children's Time with Fathers in Intact Families." *Journal of Marriage and the Family,* 63: 136-5.

FIGURE 1

Fraction male by age of child
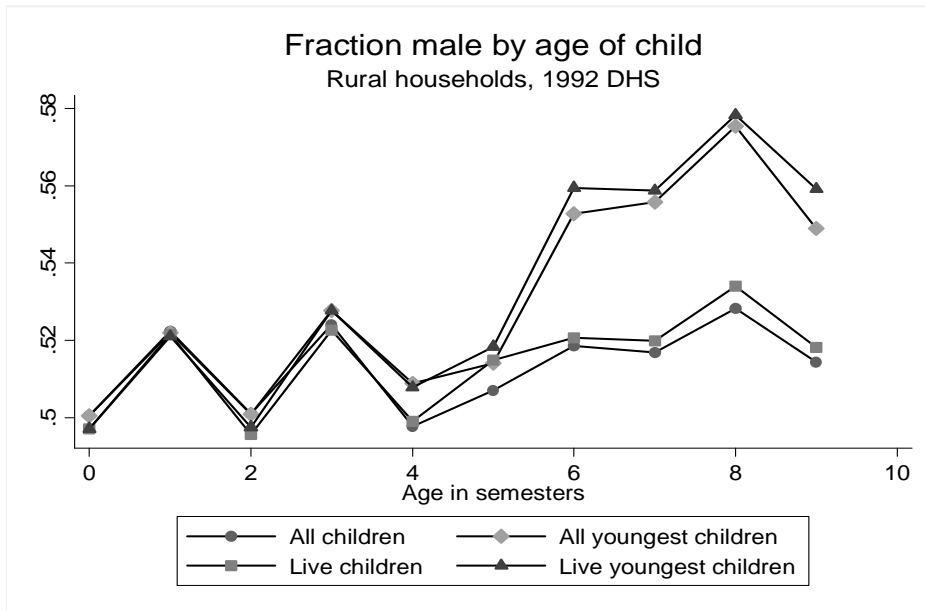Rural households, 1992 DHS

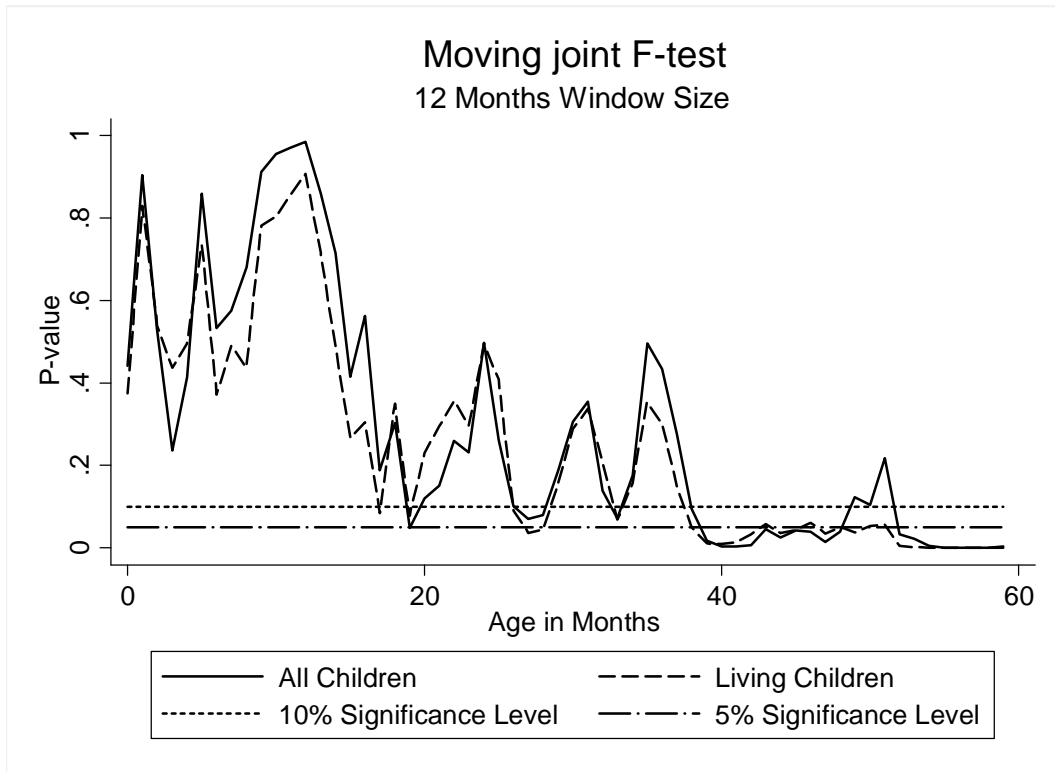FIGURE 2a: Do characteristics of the mother and the family predict gender?

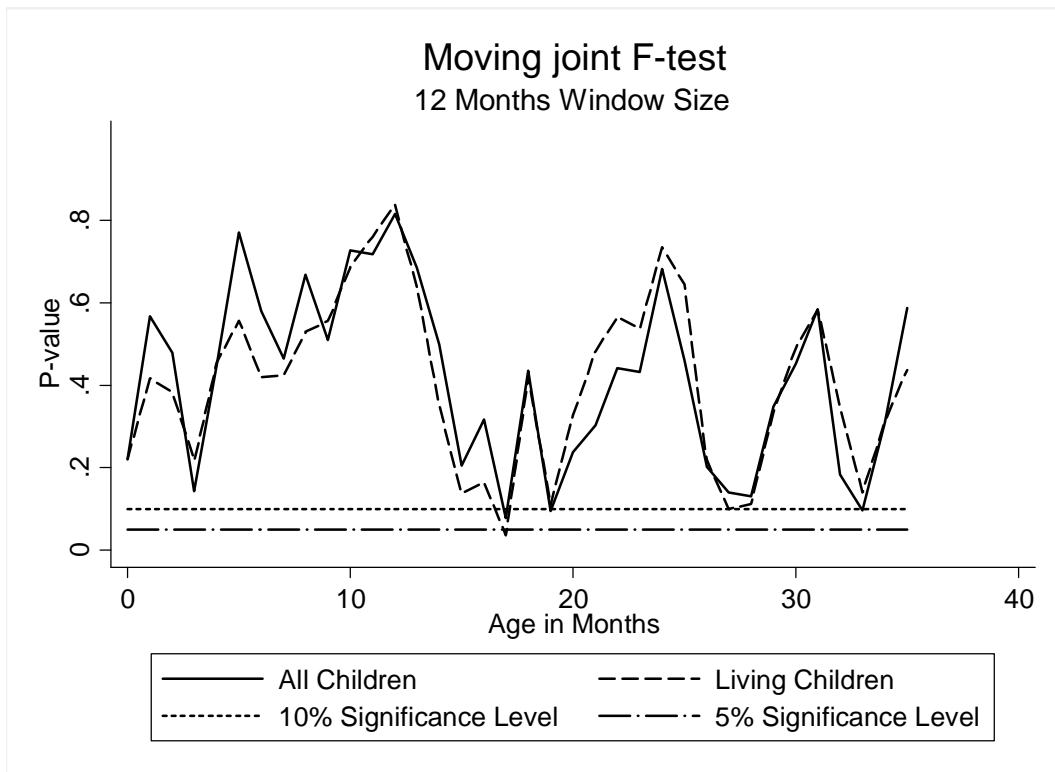Figure 2b: Predicting gender using mother, family and prenatal care use.

FIGURE 3: Childcare Time by gender, Indian Time Use Survey 1998 -1999
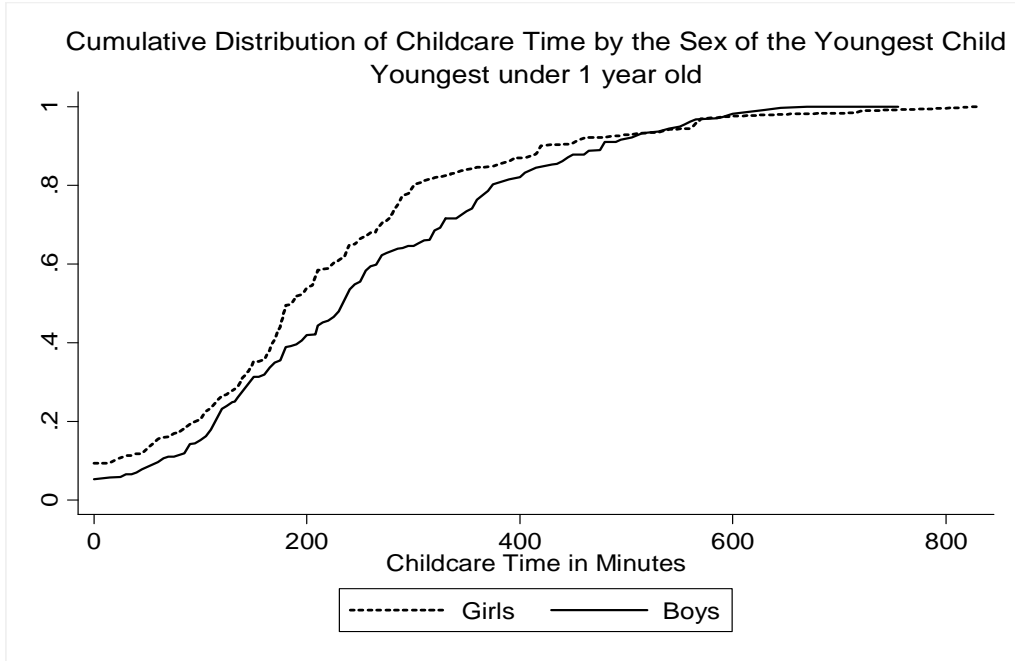
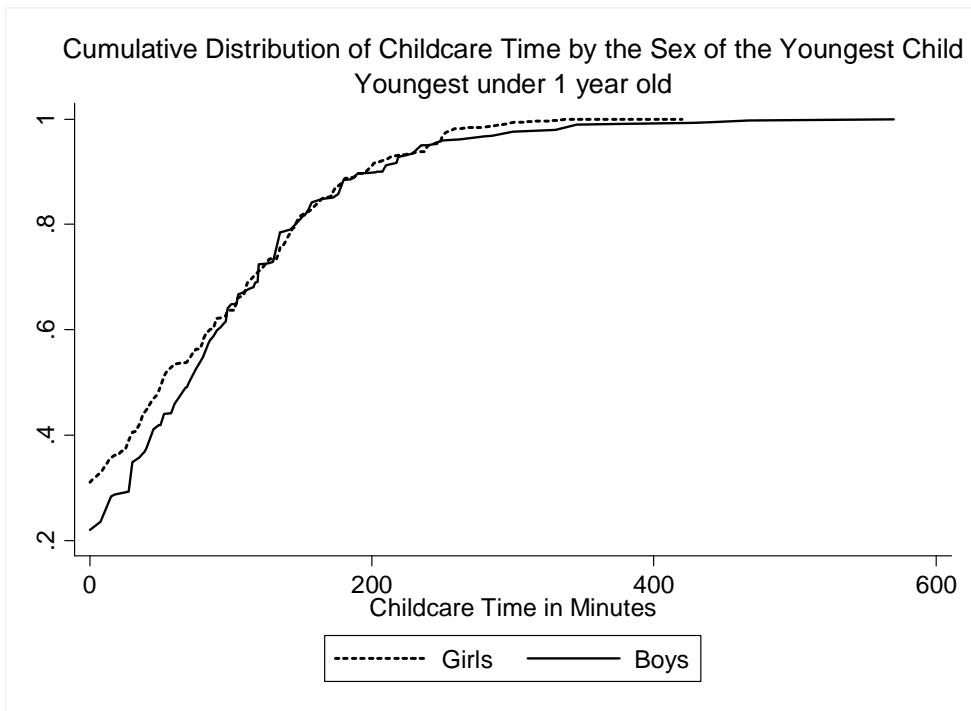**Cumulative Distribution of Childcare Time by the Sex of the Youngest Child**
Youngest under 1 year old

Childcare Time in Minutes

---------- Girls ———— Boys

FIGURE 4 Childcare Time by gender, South Africa Time Use Survey 2000

Cumulative Distribution of Childcare Time by the Sex of the Youngest Child
Youngest under 1 year old

Childcare Time in Minutes

---------- Girls ———— Boys

Appendix 1

Proposition. Let $E[S/N]$ be the desired number of sons of the family of the average child in a family with N children and $E[S/N,G]$ the desired number of sons of the family of the average girl in a family with N children (G is an indicator variable for whether the child is a girl). If families continue to have children until they achieve their desired number of sons S, then $E[S/N] \geq E[S/N,G]$.

Proof: It is sufficient to show that $p(S \leq s/N)$ first-order stochastically dominates $p(S \leq s/N,G)$. By definition, we have:

$$p(S=s/N)=p(S=s,G/N) + p(S=s,B/N),$$

$$p(S=s/N)=p(S=s/N,G)\, p(G/N)+p(B/S,N)\, p(S=s/N),$$

which can be rewriten as:

$$p(S=s/N)\, [1-p(B/\,S,N)] = p(S=s/N,G)p(G/N),$$

$$p(S=s/N)=p(S=s/N,G)*[p(G/N)/p(G/S,N)].$$

Finally, notice that the term between brackets is increasing in S since $p(G/S,N)$ is decreasing in S. Because $p(S=s/N)$ and $p(S=s/N,G)$ must be equal to one when summed over s, it must be that the term between brackets $[p(G/N)/p(G/S,N)]$ is lower than 1 for low values of S and greater than 1 for high values of S, and that $p(S \leq s/N)$ first-order stochastically dominates $p(S \leq s/N,G)$. QED.

Appendix 2

We consider a model in which there is heterogeneity in the way that households treat boys and girls. Formally, the model can be written as:

$$Z_{ih} = \delta_{0,h}(1 - B_{ih}) + \delta_{1,h}B_{ih} + \xi_{ih}, \tag{1}$$

where Z_{ih} is the investment in child i in household h, $\delta_{0,h}$ ($\delta_{1,h}$) is the average investment made by household h in girls (boys), B_{ih} is a dummy that is equal to 1 if child i in household h is a boy and ν_{ih} is an error term.

Unconditional

Suppose we run a regression of Z_{ih} on a constant and B_{ih}. In this case, the OLS bias is given by:

$$\frac{Cov\left(\xi_{ih}, B_{ih}\right)}{Var\left(B_{ih}\right)}$$

because the average boy and the average girl are in families with identical preferences for children – i.e., $E\left[\delta_{0,h}|B_{ih}\right] = E\left[\delta_{0,h}\right]$ and $E\left[\delta_{1,h}|B_{ih}\right] = E\left[\delta_{1,h}\right]$. However, because girls are on average in larger families than boys, the OLS estimate is biased if N_h is part of the error term ξ_{ih}.

Conditional on Family Size

We can rewrite (1) as:

$$Z_{ih} = E\left[\delta_{0,h}|N_h\right] + \left(E\left[\delta_{1,h}|N_h\right] - E\left[\delta_{0,h}|N_h\right]\right)B_{ih} + \nu_{ih} + \xi_{ih}, \tag{2}$$

where

$$\nu_{ih} = \delta_{0,h} - E\left[\delta_{0,h}|N_h\right] + \left[\left(\delta_{1,h} - E\left[\delta_{1,h}|N_h\right]\right) - \left(\delta_{0,h} - E\left[\delta_{0,h}|N_h\right]\right)\right]B_{ih}. \tag{3}$$

Suppose we run a regression of Z_{ih} on a constant and B_{ih}, controlling for family size N_h. In this case, the OLS bias is given by:

$$\frac{Cov\left(\nu_{ih}, B_{ih}|N_h\right)}{Var\left(B_{ih}|N_h\right)} + \frac{Cov\left(\xi_{ih}, B_{ih}|N_h\right)}{Var\left(B_{ih}|N_h\right)}.$$

Here we are interested in the bias that may arise from how families with different preferences treat boys and girls so we will concentrate in the first term. Let us write:

$$\frac{Cov\left(\nu_{ih}, B_{ih}|N_h\right)}{Var\left(B_{ih}|N_h\right)} = \frac{E\left[\left(\nu_{ih} - E\left[\nu_{ih}\right]\right)B_{ih}|N_h\right]}{Var\left(B_{ih}|N_h\right)}$$

and the law of iterated expectations implies:

$$\begin{aligned}
\frac{Cov\left(\nu_{ih}, B_{ih}|N_h\right)}{Var\left(B_{ih}|N_h\right)} &= \frac{E\left[E\left[\left(\nu_{ih} - E\left[\nu_{ih}\right]\right)B_{ih}|B_{ih}, N_h\right]\right]}{Var\left(B_{ih}|N_h\right)} = \\
&= \frac{pr\left(B_{ih} = 1|N_h\right)}{Var\left(B_{ih}|N_h\right)}E\left[\left(\nu_{ih} - E\left[\nu_{ih}\right]\right)|B_{ih} = 1, N_h\right]. \tag{4}
\end{aligned}$$

Furthermore, notice that we can rewrite $E\left[\nu_{ih}|N_h\right]$ as:

$$\begin{aligned}
E\left[\nu_{ih}|N_h\right] &= pr\left(B_{ih} = 1|N_h\right)E\left[\nu_{ih}|B_{ih} = 1, N_h\right] + \\
&\quad + pr\left(B_{ih} = 0|N_h\right)E\left[\nu_{ih}|B_{ih} = 0, N_h\right]. \tag{5}
\end{aligned}$$

Substituting (5) into (4) yields:

$$\frac{Cov\left(\nu_{ih}, B_{ih}|N_h\right)}{Var\left(B_{ih}|N_h\right)} = \kappa\left\{E\left[\nu_{ih}|B_{ih}=1, N_h\right] - E\left[\delta_{0,h}|B_{ih}=0, N_h\right] + E\left[\delta_{0,h}|, N_h\right]\right\} =$$

$$= \kappa\left\{\left(E\left[\delta_{1,h}|B_{ih}=1, N_h\right] - E\left[\delta_{1,h}|N_h\right]\right) - \left(E\left[\delta_{0,h}|B_{ih}=0, N_h\right] - E\left[\delta_{0,h}|N_h\right]\right)\right\}, (6)$$

where

$$\kappa = \frac{pr\left(B_{ih}=1|N_h\right)\left[1 - pr\left(B_{ih}=1|N_h\right)\right]}{Var\left(B_{ih}|N_h\right)}. \tag{7}$$

Therefore, the sign of the bias is determined by the expression between curly brackets in (6):

$$\left(E\left[\delta_{1,h}|B_{ih}=1, N_h\right] - E\left[\delta_{1,h}|N_h\right]\right) - \left(E\left[\delta_{0,h}|B_{ih}=0, N_h\right] - E\left[\delta_{0,h}|N_h\right]\right).$$

Notice that the same formula applies if one conditions on a vector X_{ih} that contains family size N_h.

TABLE 1: TESTING RANDOM ASSIGMENT (DHS 1992). MEAN DIFFERENCES BY GENDER.

Sample:	Youngest live child											
	Age 0-15 months (N=11,595)				Age 16-47 months (N=13,690)				Age 48-59 months (N=2,501)			
	const. (mean for females)	s.e.	coefficient on male	s.e	const. (mean for females)	s.e.	coefficient on male	s.e	const. (mean for females)	s.e.	coefficient on male	s.e
Child characteristics												
# of siblings ever born	2.088	[0.032]***	0.009	[0.046]	2.185	[0.032]***	0.062	[0.041]	2.69	[0.079]***	-0.02	[0.101]
# of brothers ever born	1.011	[0.020]***	-0.037	[0.027]	1.079	[0.018]***	0.003	[0.024]	1.432	[0.048]***	-0.103	[0.062]*
# of sisters ever born	1.076	[0.020]***	0.046	[0.029]	1.106	[0.021]***	0.058	[0.028]**	1.259	[0.050]***	0.083	[0.062]
Birth month	6.866	[0.055]***	-0.109	[0.077]	6.723	[0.049]***	0.024	[0.069]	6.762	[0.125]***	-0.17	[0.164]
Mother's characteristics												
Mother's age	24.716	[0.089]***	0.028	[0.125]	26.669	[0.091]***	0.147	[0.124]	30.1	[0.226]***	-0.208	[0.276]
Mother's ethnicity (scheduled caste omitted)												
Scheduled tribe	0.114	[0.005]***	-0.005	[0.007]	0.114	[0.006]***	-0.001	[0.007]	0.086	[0.010]***	0.01	[0.013]
Other	0.744	[0.007]***	0.005	[0.009]	0.753	[0.008]***	-0.003	[0.009]	0.773	[0.016]***	0.003	[0.019]
Mother's religion (other omitted)												
Hindu	0.818	[0.006]***	-0.002	[0.008]	0.832	[0.008]***	0	[0.007]	0.817	[0.015]***	0.031	[0.018]*
Muslim	0.132	[0.005]***	0.003	[0.008]	0.126	[0.008]***	-0.003	[0.006]	0.126	[0.014]***	-0.037	[0.015]**
Christian	0.021	[0.002]***	0	[0.002]	0.018	[0.002]***	-0.002	[0.002]	0.025	[0.005]***	-0.006	[0.006]
Mother's years of education	1.954	[0.053]***	-0.042	[0.074]	1.929	[0.059]***	-0.023	[0.065]	2.154	[0.131]***	-0.088	[0.176]
Mother born in urban area	0.063	[0.004]***	-0.005	[0.005]	0.063	[0.004]***	-0.002	[0.005]	0.072	[0.010]***	0.001	[0.012]
Mother's age first married	16.36	[0.043]***	0.041	[0.062]	16.223	[0.046]***	-0.02	[0.056]	16.092	[0.108]***	0.005	[0.137]
Mother's age at first birth	18.551	[0.045]***	0.054	[0.066]	18.459	[0.047]***	-0.032	[0.061]	18.534	[0.121]***	-0.288	[0.150]*
Mother speaks Hindi	0.483	[0.008]***	-0.016	[0.011]	0.422	[0.009]***	0.018	[0.011]*	0.361	[0.018]***	0.038	[0.025]
Pvalue of test that mother and family charateristic	**0.2803**				**0.6013**				**0.005**			
Prenatal characteristics[+]												
Any prenatal care?	0.607	[0.008]***	-0.004	[0.011]	0.582	[0.009]***	0.007	[0.010]				
# prenatal visits	2.245	[0.041]***	-0.033	[0.058]	2.304	[0.047]***	0.007	[0.055]				
Any tetanus shots?	0.587	[0.008]***	0.009	[0.011]	0.564	[0.009]***	0.006	[0.011]				
# tetanus shots	1.28	[0.019]***	0.026	[0.027]	1.3	[0.022]***	-0.002	[0.026]				
non-home delivery	0.157	[0.006]***	0.01	[0.008]	0.174	[0.006]***	-0.005	[0.008]		NA		
Pvalue of test that prenatal predict gender	**0.1937**				**0.6283**							
Test results from regressions that include all charateristics												
(3) Pvalue prenatal only (Joint Test)	**0.135**				**0.5104**							
(4) Pvalue excl. prenatal (Joint Test)	**0.2268**				**0.5405**				**NA**			
(5) Pvalue all (Joint Test)	**0.1435**				**0.661**				**0.005**			

Standard errors (in brackets) are computed taking survey design into account. Coefficients reported from separate linear regressions, where each characteristic is regressed on a dummy for male and a constant. The p-value for the joint test comes from regressing the youngest child's gender on all the charateristics (except number of all siblings, since that is collinear to the number of borthers and number of sisters) and testing whether they are jointly significant. + Prenatal characteristics only available for children 0-47 months old. In the 16-59 age group the number of observations for these variables is 13,690.   ** $p<0.05$, * $p<0.1$

TABLE 2. DESCRIPTIVE STATISTICS, TIME USE SURVEY (1998-1999). RURAL AREAS.

	HHs with youngest below age 1		HHs with youngest ages 0-1		HHs with youngest ages 2-5	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Percentage of all households	0.04		0.15		0.30	
Time Use:						
Time spent on child care (minutes per day)	236.62	159.28	196.90	152.19	107.19	129.05
Time spent on child care by female members	192.15	134.80	166.23	132.76	88.87	108.79
Time spent on child care by male members	44.47	82.20	30.67	64.70	18.31	51.21
Time spent on physical care	165.22	125.96	137.89	121.83	73.06	94.29
Time spent supervising children	55.50	114.87	48.96	105.50	24.26	76.10
Time spent instructing children	5.25	27.37	4.10	23.10	5.24	26.62
Time spent taking children to places	10.65	67.78	5.94	42.81	4.62	39.92
Time spent on exclusive child care	132.44	153.63	95.55	137.45	57.43	105.52
Household characteristics:						
Household size	4.68	1.87	4.54	1.83	3.95	1.54
Male youngest	0.46	0.50	0.51	0.50	0.55	0.50
Scheduled tribe	0.24	0.43	0.23	0.42	0.21	0.41
Scheduled caste	0.14	0.35	0.18	0.38	0.20	0.40
Hindu	0.91	0.28	0.91	0.29	0.92	0.28
Per capita expenditure	393.31	175.58	393.83	188.92	408.61	196.36
Land owned and possessed	4.85	7.97	4.51	8.25	3.89	9.13
Observations	562		1,947		3,815	

Notes: Weighted statistics for households in each sample. The statistics in the first two columns are for households where the youngest child is under 1, columns (3) and (4) for households where the youngest child is under 1, and the last two columns for households where the youngest is between 2 and 5 years of age.

TABLE 3. EFFECT OF CHILD GENDER ON HOUSEHOLD CHILD CARE TIME, TIME USE SURVEY (1998-1999)

Model:	OLS	OLS	Logit	OLS	Tobit
Dependent variable:	Number of minutes per day, including 0s	Number of minutes per day, including 0s	Any care? (Beta reported)	Number of minutes per day>0	Number of minutes per day
	(1)	(2)	(3)	(4)	(5)
Panel A: Youngest kids under 1 year old					
Male=1	32.772	30.018	0.613	24.226	36.309
	[17.669]*	[17.511]*	[0.397]	[17.344]	[18.855]*
Controls?	no	yes	no	no	no
Obs	562	562	562	516	562
Mean Y	236.62	236.62	0.93	255.51	236.62
Panel B: Youngest kids under 2 years old					
Male=1	18.689	16.602	-0.052	21.951	18.647
	[8.643]**	[8.593]*	[0.182]	[8.629]**	[9.509]**
Controls?	no	yes	no	no	no
Obs	1947	1947	1947	1747	1947
Mean Y	196.90	196.90	0.90	219.17	196.90
Panel C: Youngest kids 2-5 years old					
Male=1	-1.745	-2.056	-0.013	-1.888	-2.205
	[5.086]	[5.094]	[0.089]	[5.991]	[6.905]
Controls?	no	yes	no	no	no
Obs	3815	3815	3815	2765	3815
Mean Y	107.19	107.19	0.72	149.79	107.19

Robust standard errors in brackets. The dependent variable in all columns except (3) is the number of minutes per day spent with child care by all household members. The dependent variable in column (3) is an indicator variable for positive childcare time. Panel A reports results for households whose youngest child is under 1 year old, panel B for those whose youngest is under 2 and panel C for those whose youngest is between 2 and 5 years old. The controls include dummies for household caste (2 dummies), a dummy for whether the household was Hindu and the area of the land owned and possessed by the household. Survey weights are used in estimation. *** p<0.01, ** p<0.05, * p<0.1

TABLE 4. HETEROGENEITY IN CHILDCARE TIME, TIME USE SURVEY (1998-1999).

	Childcare		Households w/ ONLY 1 child under 6	Physical Care		Supervising		Exclusive Care	Urban Households, childcare
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Youngest kids under 1 year old									OLS
Male = 1	32.772	44.224	59.855	30.139	5.4	20.217	55.231	59.261	-5.709
	[17.669]*	[27.700]	[30.143]**	[13.444]**	[23.359]	[12.220]*	[16.674]***	[24.936]**	[34.546]
Male* (# other children under 6)		-9.082			20.59		-28.721	-27.47	
		[17.920]			[14.671]		[11.020]***	[15.135]*	
# Other children under 6		8.627			-10.209		18.098	-3.689	
		[11.917]			[8.372]		[7.163]**	[9.636]	
Constant		210.49	204.183	151.245	164.184	46.131	23.193	123.841	265.737
		[20.368]***	[21.463]***	[8.622]***	[15.016]***	[8.397]***	[7.541]***	[16.318]***	[27.939]***
Observations	562	562	151	562	562	562	562	562	204
Panel B: Youngest kids under 2 years old									
Male = 1	18.689	49.654	50.777	14.673	20.442	10.643	35.414	37.278	-15.867
	[8.643]**	[14.179]***	[15.969]***	[6.597]**	[11.150]*	[6.077]*	[9.874]***	[12.577]***	[16.173]
Male* (# other children under 6)		-25.206			-4.725		-20.231	-23.943	
		[9.042]***			[7.181]		[6.003]***	[7.496]***	
# Other children under 6		12.439			1.606		8.315	-3.696	
		[5.904]**			[4.471]		[4.061]**	[5.001]	
Constant	187.332	171.739	168.946	130.381	128.368	43.516	33.092	95.753	235.373
	[6.143]***	[9.715]***	[10.742]***	[4.346]***	[7.393]***	[3.723]***	[5.369]***	[8.097]***	[11.733]***
Observations	1947	1947	481	1947	1947	1947	1947	1947	677

Robust standard errors in brackets. Panel A reports results for households whose youngest child is under 1 year old, panel B for those whose youngest is under 2. The dependent variable in columns (1) and (2) is the number of minutes per day spent with child care. In column (3), the dependent variable is childcare time and the sample is further restricted to households with only one child under 6. The dependent variable in columns (4) and (5) is the amount of time spent taking physical care of children (e.g., washing, dressing and feeding). The dependent variable in columns (6) and (7) is the amount of time spent supervising children. The dependent variable in column (8) is the amount of time spent exclusively on childcare -- i.e., the caretaker was not multi-tasking. The dependent variable in column (9) is the amount of time spent on childcare, as in columns (1) and (2). The variable "# Other children under 6" excludes the youngest child. Its mean is equal to 1.23 children. All columns are estimated using OLS. Survey weights are used for estimation. *** p<0.01, ** p<0.05, * p<0.1

TABLE 5: EFFECT OF CHILD GENDER ON PARENTAL INPUTS, CHILDREN 0-15 MONTHS OLD. DHS (1992)

Dependent variable:		Was child ever breastfed?		# months breastfed	log(# months breastfed)	# months breastfed		Vitamin A supplement?		Did mother have vaccination card at interview?	
Model:	controls ?	OLS	Logit (beta reported)	censored regression	censored regression	Accelerated Failure Time model	Proportional Hazard Model	OLS	Logit (beta reported)	OLS	Logit (beta reported)
				(# censored obs: 10,689)							
Male = 1	no	0.006	0.134	1.802	0.289	0.412	-0.41	0.014	0.133	0.043	0.215
		[0.004]	[0.100]	[0.574]***	[0.092]***	[0.129]***	[0.129]***	[0.007]**	[0.063]**	[0.009]***	[0.046]***
Male = 1	yes	0.006	0.138	1.937	0.313	0.438	-0.434	0.013	0.139	0.041	0.225
		[0.004]	[0.100]	[0.564]***	[0.092]***	[0.130]***	[0.131]***	[0.006]*	[0.069]**	[0.009]***	[0.049]***
Obs		11609		11073	11073			11248		11616	
Mean of Y		0.953		7.677	1.803			0.117		0.275	

Standard errors [in brackets] are computed taking survey design into account. Child ever breastfed is equal to zero if mother reports that child was not breastfed or if breastfeeding duration was less than a month. Each coefficient corresponds to a separate estimation, and survey weights are used. The number of observations for each age group varies from outcome to outcome because there are a few missing values. Controls include all variables in Table 1: # of brothers, # of sisters, birth month, mother's age, mother's caste (2 dummies), mother's religion (3 dummies), mother's years of education, whether mother was born in rural area, mother's age at first marriage, mother's age at first birth, mother speaks Hindi, prental care use, number fo prental care visits, tetanus shot, number of tetanus shots, and home delivery. *** p<0.01, ** p<0.05, * p<0.1

TABLE 6: EFFECT OF CHILD GENDER ON VACCINATIONS, DHS 1992

	controls?	BCG	DPT 1st dose	DPT 2nd dose	DPT 3rd dose	Polio 1st dose	Polio 2nd dose	Polio 3rd dose	Measles
Panel A: Youngest kids 0-15 months old. OLS									
Male = 1	no	0.036	0.048	0.035	0.031	0.049	0.037	0.032	0.02
		[0.011]***	[0.011]***	[0.010]***	[0.009]***	[0.011]***	[0.010]***	[0.009]***	[0.008]***
Male = 1	yes	0.033	0.047	0.033	0.03	0.047	0.036	0.031	0.019
		[0.010]***	[0.010]***	[0.009]***	[0.009]***	[0.010]***	[0.009]***	[0.009]***	[0.007]***
Panel B: Youngest kids 0-15 months old. LOGIT (beta reported)									
Male = 1	no	0.144	0.195	0.152	0.162	0.198	0.159	0.164	0.161
		[0.043]***	[0.042]***	[0.044]***	[0.048]***	[0.042]***	[0.043]***	[0.047]***	[0.060]***
Male = 1	yes	0.17	0.23	0.175	0.182	0.232	0.185	0.184	0.165
		[0.051]***	[0.050]***	[0.050]***	[0.053]***	[0.051]***	[0.049]***	[0.052]***	[0.064]***
Obs		11591	11591	11587	11587	11609	11605	11605	11520
Mean of Y		0.448	0.47	0.354	0.262	0.472	0.368	0.271	0.147
Panel C: Youngest kids 9-15 months old. OLS									
Male = 1	no	0.036	0.058	0.055	0.052	0.058	0.051	0.054	0.038
		[0.017]**	[0.017]***	[0.017]***	[0.016]***	[0.017]***	[0.017]***	[0.016]***	[0.015]**
Obs		4815	4808	4806	4806	4822	4818	4818	4759
Mean of Y		0.573	0.613	0.525	0.434	0.617	0.545	0.453	0.313
Panel D: Youngest kids 0-15 months old <u>with vaccination card</u> OLS									
Male = 1	no	-0.017	0.005	-0.004	0.023	0.006	-0.005	0.027	0.016
		[0.014]	[0.009]	[0.019]	[0.021]	[0.010]	[0.019]	[0.021]	[0.019]
Obs		3338	3338	3338	3338	3338	3338	3338	3338
Mean of Y		0.869	0.947	0.747	0.574	0.94	0.75	0.574	0.271

Standard errors [in brackets] are computed taking survey design into account. Each coefficient corresponds to a separate estimation, and survey weights are used. The number of observations for each age group varies from outcome to outcome because there are a few missing values. Controls include all variables in Table 1: # of brothers, # of sisters, birth month, mother's age, mother's caste (2 dummies), mother's religion (3 dummies), mother's years of education, whether mother was born in rural area, mother's age at first marriage, mother's age at first birth, mother speaks Hindi, prental care use, number fo prental care visits, tetanus shot, number of tetanus shots, and home delivery. *** p<0.01, ** p<0.05, * p<0.1

TABLE 7: ANTICIPATED FAMILY SIZE, YOUNGEST CHILDREN 0-15 MONTHS OLD WITH OLDER SIBLINGS. DHS 1992

	want more children?	Ever breastfed?	Vitamin A supplement	Vaccination card?	BCG	DPT 1st dose	DPT 2nd dose	DPT 3rd dose	Polio 1st dose	Polio 2nd dose	Polio 3rd dose	Measles	# months breastfed	log(# months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(14)	(15)
Panel A: Predicting desire for more children														
Male = 1	-0.134													
	[0.012]***													
First Stage:														
First born is male	-0.082													
	[0.013]***													
Panel B: OLS controlling for anticipated family size														
Male = 1		0.006	0.012	0.034	0.037	0.049	0.044	0.039	0.050	0.042	0.035	0.018	2.974	0.486
		[0.005]	[0.007]	[0.010]***	[0.013]***	[0.013]***	[0.012]***	[0.011]***	[0.013]***	[0.012]***	[0.011]***	[0.009]**	[0.715]***	[0.119]***
More kids = 1		0.007	-0.010	-0.005	-0.016	-0.012	-0.003	-0.006	-0.022	-0.013	-0.013	-0.015	2.737	0.454
		[0.006]	[0.008]	[0.012]	[0.014]	[0.014]	[0.013]	[0.011]	[0.014]	[0.013]	[0.011]	[0.009]*	[0.844]***	[0.140]***
Panel C: Benchmark for sample of children with older siblings														
Male = 1		0.005	0.013	0.034	0.039	0.051	0.044	0.040	0.053	0.043	0.037	0.020	2.680	0.437
		[0.005]	[0.007]*	[0.010]***	[0.012]***	[0.013]***	[0.012]***	[0.011]***	[0.013]***	[0.012]***	[0.011]***	[0.008]**	[0.711]***	[0.117]***
Panel C: IV controlling for anticipated family size														
						IV (First born male is the instrument)							**Reduced Form^^**	
Male = 1		-0.015	0.007	0.038	0.043	0.057	0.036	0.045	0.076	0.052	0.047	0.029	2.704	0.440
		[0.011]	[0.015]	[0.021]*	[0.023]*	[0.024]**	[0.022]*	[0.019]**	[0.024]***	[0.022]**	[0.019]**	[0.016]*	[0.710]***	[0.117]***
More kids = 1		-0.149	-0.042	0.024	0.031	0.045	-0.058	0.041	0.177	0.068	0.075	0.070	1.780	0.285
		[0.072]**	[0.097]	[0.134]	[0.147]	[0.151]	[0.140]	[0.120]	[0.155]	[0.143]	[0.126]	[0.102]	[0.773]**	[0.126]**
Obs	8499	8484	8214	8490	8468	8474	8473	8473	8486	8483	8483	8421	8092	8092
Mean of Y	0.409	0.954	0.106	0.251	0.417	0.440	0.324	0.234	0.442	0.338	0.245	0.130	7.631	1.795

Notes: Robust standard errors [in brackets]. Each sub-panel (separated by horizontal lines) corresponds to a separate linear regression of the dependent variable listed in the column on the independent variables listed in the rows. The sample is restricted to children who had (alive or dead) older siblings. Survey weights are used for estimation. In Panel A: 1-we regress wanting more children on youngest gender (coefficien -0.134), and 2-we regress wanting more children on the gender of the first born (coefficient -0.82). ^^ this column reports the OLS regression of breastfeeding duration regressed on gender of the youngest and gender of the first born (we do not include wanting more children). *** p<0.01, ** p<0.05, * p<0.1

TABLE 8. DO BOYS NEED MORE? MEAN GENDER DIFFERENCES IN PARENTAL INPUTS

Dependent variable:	Model:	Coefficient on I(male=1)	s.e.	N	Mean Y	% effect	Coefficient on I(male=1)	s.e.	Coefficient on I(male=1)*I(north=1)	s.e.	N	Mean Y	% effect (North vs. South)
							DHS CHILDREN AGES 0-15 MONTHS						
		1998 SOUTH AFRICAN DHS					**1992 INDIAN DHS (NORTH VERSUS SOUTH COMPARISON)**						
Ever breastfed?	OLS	0.024	[0.033]	723	0.875	3%	0.004	[0.010]	-0.001	[0.012]	7329	0.951	0%
# months breastfed	cens. reg.	-0.409	[1.218]	632	6.941	-6%	2.345	[0.997]**	-0.688	[1.267]	6979	7.649	-9%
log(# months breastfed)	cens. reg.	-0.085	[0.234]	632	1.663	-5%	0.362	[0.153]**	-0.118	[0.192]	6979	1.794	-7%
Vaccination card?	OLS	-0.05	[0.035]	732	0.835	-6%	0.035	[0.020]*	0.017	[0.024]	7326	0.284	6%
BCG	OLS	0.001	[0.011]	729	0.978	0%	0.009	[0.020]	0.04	[0.025]	7310	0.494	8%
DPT 1st dose	OLS	0.017	[0.029]	723	0.864	2%	0.024	[0.020]	0.029	[0.025]	7309	0.508	6%
DPT 2nd dose	OLS	-0.038	[0.037]	715	0.722	-5%	-0.002	[0.023]	0.052	[0.027]*	7305	0.384	14%
DPT 3rd dose	OLS	-0.071	[0.040]*	715	0.577	-12%	-0.006	[0.023]	0.059	[0.026]**	7305	0.288	20%
Polio 1st dose	OLS	0.002	[0.032]	720	0.849	0%	0.035	[0.020]*	0.015	[0.025]	7319	0.511	3%
Polio 2nd dose	OLS	-0.043	[0.039]	716	0.705	-6%	0.021	[0.022]	0.027	[0.026]	7315	0.4	7%
Polio 3rd dose	OLS	-0.05	[0.040]	716	0.554	-9%	0.004	[0.023]	0.047	[0.026]*	7315	0.3	16%
Measles	OLS	0.098	[0.040]**	718	0.348	28%	0.008	[0.020]	0.016	[0.022]	7261	0.168	10%
Vitamin A	OLS	na	na	na	na	na	-0.015	[0.017]	0.045	[0.019]**	7096	0.133	34%

2000 SOUTH AFRICA TIME USE SURVEY (TUS) CHILDREN UNDER 1

	Model:	Coefficient on I(male=1)	s.e.	N	Mean Y	% effect
Childcare, mins. per day >= 0	OLS	1.184	[12.055]	521	99.6	1%
Childcare, mins. per day >= 0	Tobit	9.874	[16.353]	521	99.6	10%
Childcare, mins. per day > 0	OLS	-15.233	[13.303]	386	136.0	-11%
Any care?	Logit (Beta reported)	0.471	[0.265]*	521	0.733	13%

Standard errors [in brackets] are computed taking survey design into account. Vitamin A supplementation information is not available for South African DHS. Indian results only include Northern (Gujarat, Rajasthan, Uttar Pradesh, Madhya Pradesh, Punjab, Haryana) and Southern states (Kerala, Tamil Nadu, Andhra Pradesh, Karnataka, Maharashtra). Definition of North and Sounth India follows Dyson and Moore 1983. Results reported do not include demographic controls. All the Indian regressions include a dummy for northern states *** p<0.01, ** p<0.05, * p<0.1

TABLE 9: EFFECT OF CHILD GENDER ON ANTHROPROMETRIC MEASURES, AGES 0-15 MONTHS, DHS 1992

Dependent variable:	controls?	Height-for-age Z score		Stunted = 1 (height-for-age Z score< 2 s.d. below reference median)		Weight-for-age Z score		Underweight=1 (weight-for-age Z score< 2 s.d. below reference median)		Weight-for-height Z score		Wasted=1 (Weight-for-height Z score< 2 s.d. below reference median)	
		DHS	UK	DHS	UK	DHS	UK	DHS	UK	DHS	CDC	DHS	CDC
Male = 1	no	-0.218	-0.074	0.057	0.017	-0.162	0.013	0.04	-0.005	-0.063	0.028	0.036	0.001
		[0.043]***	[0.049]	[0.015]***	[0.015]	[0.033]***	[0.039]	[0.012]***	[0.013]	[0.038]*	[0.050]	[0.011]***	[0.014]
Male = 1	yes	-0.225	-0.085	0.059	0.02	-0.165	0.008	0.041	-0.004	-0.067	0.02	0.036	0.002
		[0.043]***	[0.048]*	[0.015]***	[0.015]	[0.032]***	[0.038]	[0.012]***	[0.013]	[0.037]*	[0.050]	[0.011]***	[0.014]
Obs		6396	6396	6396	6396	8550	8550	8550	8550	6411	6411	6411	6411
Mean of Y		-1.3	-1.353	0.323	0.361	-1.51	-2.026	0.381	0.525	-0.727	-1.253	0.137	0.309

Standard errors [in brackets] are computed taking survey design into account. Each coefficient corresponds to a separate estimation. Controls include all variables in Table 3: # of brothers, # of sisters, birth month, mother's age, mother's caste (2 dummies), mother's religion (3 dummies), mother's years of education, whether mother was born in rural area, mother's age at first marriage, mother's age at first birth, and whether mother speaks Hindi. The other measures are standardized using the UK (1990) standards or the 2000 CDC standards. The UK standards are not available for height for age. CDC standards for height are not available for children under 2. Survey weights are used in estimation. For children under 2 the standards use length rather than height, which is measured while lying instead of standing. *** p<0.01, ** p<0.05, * p<0.1

TABLE 10: EFFECT OF CHILD'S GENDER ON LIVING ARRANGEMENTS, YOUNGEST CHILDREN 15 MONTHS AND YOUNGER. RURAL HOUSEHOLDS.

Panel A: Effect of gender on family living arrangements in the DHS 1992. Youngest children 0-15 months old

Dependent variable:	Husband lives home?		Is mother the wife of the household head?		# of other sibs living at home		# of sisters living at home		# of brothers living at home	
	OLS	Logit (beta reported)	OLS	Logit (beta reported)	OLS	Negative binomial (IRR reported)	OLS	Negative binomial (IRR reported)	OLS	Negative binomial (IRR reported)
Male=1	0.008	0.084	-0.009	-0.038	-0.004	0.998	0.034	1.04	-0.037	0.953
	[0.007]	[0.075]	[0.011]	[0.043]	[0.034]	[0.021]	[0.024]	[0.029]	[0.021]*	[0.026]*
Obs	11517	11517	11624	11624	11627	11627	11627	11627	11627	11627
Mean of Y	0.89	0.89	0.461	0.461	1.644	1.644	0.86	0.86	0.784	0.784

Panel B: Effect of gender on household composition TUS (1998-1999). Youngest children under 1 year old

Dependent variable:	# Men 15 and older		# Women 15 and older		# Children 14 and younger		# Girls 14 and younger		# Boys 14 and younger	
	OLS	Poisson (IRR reported)	OLS	Poisson (IRR reported)	OLS	Poisson (IRR reported)	OLS	Poisson (IRR reported)	OLS	Poisson (IRR reported)
Male=1	0.209	1.153	0.158	1.111	-0.149	0.917	0.046	1.054	-0.196	0.793
	[0.082]**	[0.063]***	[0.072]**	[0.053]**	[0.137]	[0.073]	[0.101]	[0.121]	[0.102]*	[0.098]*
Obs	562	562	562	562	562	562	562	562	562	562
Mean of Y	1.46	1.46	1.49	1.49	1.73	1.73	0.88	0.88	0.86	0.86

The standard errors [in brackets] are computed taking survey design into account in the DHS and in the TUS they allow for heteroskedasticity. Each coefficient corresponds to a separate estimation, where the dummy for the youngest child's gender is the only covariate. In the TUS (Panel B) we estimated Poisson rather than negative binomial models because some of the negative binomial models in the TUS did not converge. Survey weights are used in estimation. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX TABLE 1: EFFECT OF CHILD GENDER ON HOUSEHOLD CHILDCARE TIME BY DEMOGRAPHIC GROUP, TIME USE SURVEY (1998-1999)

Dependent variable:	Childcare by females 15 and older				Childcare by males 15 and older				Childcare by females 14 and younger			
	(1)		(2)		(3)		(4)		(5)		(6)	
	Any care?	Any care?	minutes of care≥0	minutes of care≥0	Any care?	Any care?	minutes of care≥0	minutes of care≥0	Any care?	Any care?	minutes of care≥0	minutes of care≥0
Model:	Probit	Probit	OLS	OLS	Probit	Probit	OLS	OLS	Probit	Probit	OLS	OLS
Panel A: Youngest kids under 1 year old												
Male = 1	0.036	0.073	39.949	42.586	0.04	0.031	6.798	5.581	0.032	0.106	-3.184	67.028
	[0.029]	[0.051]	[12.833]***	[22.277]*	[0.052]	[0.082]	[6.405]	[8.416]	[0.126]	[0.196]	[26.099]	[40.359]*
Male * (# other children under 6)		-0.033		-2.312		0.011		1.42		-0.058		-55.313
		[0.032]		[14.883]		[0.052]		[4.941]		[0.107]		[28.478]*
# other children under 6		0.019		0.047		0.027		3.077		0.046		37.022
		[0.022]		[9.071]		[0.035]		[3.259]		[0.070]		[19.390]*
Constant			160.771	160.711			33.938	30.042			68.769	20.052
			[8.843]***	[15.049]***			[3.980]***	[5.441]***			[20.923]***	[19.749]
Observations	560	560	560	560	555	555	555	555	117	117	117	117
Panel B: Youngest kids under 2 years old												
Male = 1	-0.002	0.062	14.507	33.564	0.02	0.075	4.208	10.104	0.086	0.2	15.334	67.198
	[0.017]	[0.032]**	[6.696]**	[11.627]***	[0.026]	[0.042]*	[2.887]	[4.486]**	[0.066]	[0.113]*	[14.516]	[27.503]**
Male * (# other children under 6)		-0.051		-15.668		-0.045		-4.752		-0.082		-36.441
		[0.020]**		[7.515]**		[0.028]		[2.876]*		[0.062]		[14.866]**
# other children under 6		0.022		3.253		0.022		3.327		0.027		18.503
		[0.014]		[4.999]		[0.020]		[2.035]		[0.044]		[9.141]**
Constant			148.327	144.247			24.458	20.281			48.055	22.041
			[4.600]***	[7.984]***			[1.871]***	[2.840]***			[8.148]***	[10.417]**
Observations	1936	1936	1936	1936	1907	1907	1907	1907	408	408	408	408

Robust standard errors in brackets. The dependent variable in columns under (1), (3) and (5) is an indicator for whether household members of a given demographic group reported spending time taking care of children. The dependent variable in columns under (2), (4) and (6) is the number of minutes per day spent with child care by all household members of a given demographic group. Panel A reports results for households whose youngest child is under 1 year old, panel B for those whose youngest is under 2. Columns under (1) to (6) show results for rural households. The variable "# Other children under 6" excludes the youngest child. Its mean is equal to 1.23 children. Survey weights are used for estimation. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX TABLE 2: EFFECT OF FAMILY SIZE ON PARENTAL INPUTS, YOUNGEST CHILDREN 0-15 MONTHS OLD WITH OLDER SIBLINGS. DHS 1992

	1st Stage: effect of first born being male on family size	Ever breastfed?	Vitamin A supplement	Vaccination card?	BCG	DPT 1st dose	DPT 2nd dose	DPT 3rd dose	Polio 1st dose	Polio 2nd dose	Polio 3rd dose	Measles	# months breastfed	log(# months)
Panel A OLS														
Family Size		-0.001	-0.012	-0.035	-0.049	-0.047	-0.043	-0.035	-0.045	-0.042	-0.033	-0.022	0.193	0.028
		[0.002]	[0.002]***	[0.003]***	[0.004]***	[0.004]***	[0.003]***	[0.003]***	[0.004]***	[0.003]***	[0.003]***	[0.002]***	[0.240]	[0.038]
Panel B: IV esimates using first-born male as an instrument														
						IV (First born male is the instrument)							**Reduced Form**	
Family Size	-0.132	-0.089	-0.022	0.016	0.023	0.037	-0.031	0.034	0.119	0.048	0.055	0.058	1.717	0.240
	[0.039]***	[0.048]*	[0.057]	[0.086]	[0.092]	[0.097]	[0.087]	[0.077]	[0.106]	[0.093]	[0.082]	[0.067]	[0.776]**	[0.125]*
F-test	11.44													
Obs	8574	8558	8284	8564	8543	8549	8548	8548	8561	8558	8558	8496	8163	8163
Mean of Y	3.283	0.954	0.106	0.251	0.416	0.440	0.325	0.235	0.442	0.339	0.245	0.130	7.641	1.797

Robust standard errors [in brackets]. Each coefficient corresponds to a separate linear regression of the dependent variable listed in the column on the independent variable listed in the rows. The sample is restricted to children who had (alive or dead) older siblings. the regressions include no controls. Survey weights are used for estimation. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Dependent variable:	Model:	Youngest live children ages 0-15			First born only (among youngest live children ages 0-15)			Live and dead children (would be ages 0-15 months at the time of the survey)			Bounds to account for mortality (information imputed for kids who died too young to receive the investment)	
		Coefficient on I(male=1), [s.e]	Mean Y	% effect	Coefficient on I(male=1), [s.e]	Mean Y	% effect	Coefficient on I(male=1), [s.e]	Mean Y	% effect	Lower Bound	Upper Bound
Ever breastfed?	OLS	0.006 [0.004]	0.954	1%	0.008 [0.010]	0.95	1%	-0.004 [0.005]	0.915	0%	-0.007	-0.002
# months breastfed	cens. reg.	1.33 [0.448]***	7.626	17%	0.514 [0.911]	7.78	7%	1.104 [0.405]***	7.555	15%	-0.954	1.283
log(# months breastfed)	cens. reg.	0.199 [0.069]***	1.797	11%	0.077 [0.139]	1.82	4%	0.176 [0.068]***	1.784	10%	-0.13	0.176
Vitamin A?	OLS	0.016 [0.006]**	0.133	12%	0.018 [0.014]	0.15	12%	0.016 [0.006]**	0.126	13%	-0.039	0.077
Vaccination card?	OLS	0.043 [0.008]***	0.287	15%	0.066 [0.021]***	0.34	19%	0.037 [0.008]***	0.271	14%	0.036	0.039
BCG	OLS	0.037 [0.009]***	0.457	8%	0.025 [0.022]	0.54	5%	0.03 [0.009]***	0.432	7%	0.03	0.03
DPT 1st dose	OLS	0.048 [0.009]***	0.476	10%	0.043 [0.022]**	0.56	8%	0.047 [0.009]***	0.45	10%	0	0.091
DPT 2nd dose	OLS	0.034 [0.009]***	0.364	9%	0.008 [0.021]	0.44	2%	0.034 [0.009]***	0.343	10%	-0.015	0.083
DPT 3rd dose	OLS	0.028 [0.008]***	0.27	10%	0.007 [0.020]	0.34	2%	0.028 [0.008]***	0.254	11%	-0.023	0.082
Polio 1st dose	OLS	0.051 [0.009]***	0.477	11%	0.04 [0.022]*	0.56	7%	0.051 [0.009]***	0.451	11%	0.003	0.094
Polio 2nd dose	OLS	0.038 [0.009]***	0.374	10%	0.018 [0.021]	0.45	4%	0.038 [0.009]***	0.353	11%	-0.011	0.086
Polio 3rd dose	OLS	0.031 [0.008]***	0.277	11%	0.018 [0.020]	0.35	5%	0.031 [0.008]***	0.26	12%	-0.02	0.085
Measles	OLS	0.018 [0.007]***	0.149	12%	0.02 [0.017]	0.20	10%	0.018 [0.007]***	0.14	13%	-0.04	0.081

Each coefficient corresponds to a separate estimation, and survey weights are not used. No controls are included. The number of observations for each age group varies from outcome to outcome because there are a few missing values. Cens. reg. is a censored regression. Upper bounds assume that all dead girls would have not received inputs (for dummy variables) or would have been given the 25th percentile of the girls' outcomes distribution. For boys we assume that had they lived they would all have been given inputs (for dummye variables) or given the 75th percentile of boys' outcome distribution. For upper bounds we assume the opposite. We use this rule to imput investments for dead children for whom the information from mothers' reports are missing or for children who died too young to have receive the investment (before 2 months for polio/DPT 1st dose, before 3 momths for polio/DPT 2nd dose, before 4 momths for polio/DPT 3rd dose, before 6 months for vitamin A and before 9 months for measles). The urban sample was constructed using the same restrictions as our main estimation sample--we dropped twins and individuals with missing predetermined covariates.

APPENDIX TABLE 4: ADDITIONAL RESULTS ON EFFECT OF GENDER ON PARENTAL INVESTMENTS. CHILDREN AGES 0-15 MONTHS

Dependent variable:	Model:	Youngest live children ages 0-15				All live children ages 16-47 months				Urban areas (youngest live children ages 0-15)			
		Coefficient on I(male=1), [s.e]		Mean Y	% effect	Coefficient on I(male=1), [s.e]		Mean Y	% effect	Coefficient on I(male=1), [s.e]		Mean Y	effect
Ever breastfed?	OLS	0.006	[0.004]	0.954	1%	0	[0.002]	0.983	0%	0	[0.009]	0.951	0%
# months breastfed	cens. reg.	1.33	[0.448]***	7.626	17%	1.567	[0.291]***	21.81	7%	0.301	[0.665]	7.431	4%
log(# months breastfed)	cens. reg.	0.199	[0.069]***	1.797	11%	0.073	[0.016]***	2.965	2%	0.017	[0.110]	1.78	1%
Vitamin A?	OLS	0.016	[0.006]**	0.133	12%	0.02	[0.007]***	0.217	9%	0.023	[0.015]	0.191	12%
Vaccination card?	OLS	0.043	[0.008]***	0.287	15%	0.03	[0.007]***	0.19	16%	-0.003	[0.019]	0.402	-1%
BCG	OLS	0.037	[0.009]***	0.457	8%	0.045	[0.008]***	0.555	8%	-0.001	[0.019]	0.663	0%
DPT 1st dose	OLS	0.048	[0.009]***	0.476	10%	0.053	[0.008]***	0.602	9%	-0.005	[0.018]	0.636	-1%
DPT 2nd dose	OLS	0.034	[0.009]***	0.364	9%	0.046	[0.008]***	0.536	9%	0.006	[0.020]	0.536	1%
DPT 3rd dose	OLS	0.028	[0.008]***	0.27	10%	0.037	[0.008]***	0.464	8%	0.021	[0.021]	0.426	5%
Polio 1st dose	OLS	0.051	[0.009]***	0.477	11%	0.055	[0.008]***	0.606	9%	0.009	[0.019]	0.641	1%
Polio 2nd dose	OLS	0.038	[0.009]***	0.374	10%	0.047	[0.008]***	0.555	8%	0.006	[0.020]	0.548	1%
Polio 3rd dose	OLS	0.031	[0.008]***	0.277	11%	0.034	[0.008]***	0.482	7%	0.02	[0.020]	0.438	5%
Measles	OLS	0.018	[0.007]***	0.149	12%	0.039	[0.008]***	0.401	10%	-0.004	[0.018]	0.217	-2%

Each coefficient corresponds to a separate estimation, and survey weights are not used. No controls are included. The number of observations for each age group varies from outcome to outcome because there are a few missing values. Cens. reg. is a censored regression. Upper bounds assume that all dead girls would have not received inputs (for dummy variables) or would have been given the 25th percentile of the girls' outcomes distribution. For boys we assume that had they lived they would all have been given inputs (for dummye variables) or given the 75th percentile of boys' outcome distribution. For upper bounds we assume the opposite. We use this rule to imput investments for dead children for whom the information from mothers' reports are missing or for children who died too young to have receive the investment (before 2 months for polio/DPT 1st dose, before 3 momths for polio/DPT 2nd dose, before 4 momths for polio/DPT 3rd dose, before 6 months for vitamin A and before 9 months for measles). The urban sample was constructed using the same restrictions as our main estimation sample--we dropped twins and individuals with missing predetermined covariates.