Effects of birth order and sibling sex composition on human capital investment in children in India

著者	Makino Momoe
権利	Copyrights 日本貿易振興機構（ジェトロ）アジア経済研究所 / Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) http://www.ide.go.jp
journal or publication title	IDE Discussion Paper
volume	319
year	2012-01-01
URL	http://hdl.handle.net/2344/1108

INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 319

Effects of Birth Order and Sibling Sex Composition on Human Capital Investment in Children in India

Momoe MAKINO*

Abstract

The paper explores the effects of birth order and sibling sex composition on human capital investment in children in India using the Indian Human Development Survey (IHDS). Endogeneity of fertility is addressed using instruments and controlling for household fixed effects. Family size effect is also distinguished from the sibling sex composition effect. Previous literature has often failed to take endogeneity into account and shows a negative birth order effect for girls in India. Once endogeneity of fertility is addressed, there is no evidence for a negative birth order effect or sibling sex composition effect for girls. Results show that boys are worse off in households that have a higher proportion of boys specifically when they have older brothers.

Keywords: birth order, sibling sex composition, household resource allocation, India

JEL classification: J13, J16, O12, O53

---

* South Asian Studies Group, Area Studies Center, IDE,
(momoe_makino@ide.go.jp)

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

Effects of Birth Order and Sibling Sex Composition on Human Capital Investment in Children in India

Momoe Makino
Institute of Developing Economies, Japan
December 2011

Abstract

The paper explores the effects of birth order and sibling sex composition on human capital investment in children in India using the Indian Human Development Survey (IHDS). Endogeneity of fertility is addressed using instruments and controlling for household fixed effects. Family size effect is also distinguished from the sibling sex composition effect. Previous literature has often failed to take endogeneity into account and shows a negative birth order effect for girls in India. Once endogeneity of fertility is addressed, there is no evidence for a negative birth order effect or sibling sex composition effect for girls. Results show that boys are worse off in households that have a higher proportion of boys specifically when they have older brothers.

Keywords: Birth order; Sibling sex composition; Household resource allocation; India
JEL classification: J13; J16; O12; O53

1. Introduction

Human capital investment in children (spending on education for example) has attracted the interest of many economists (see Becker 1993) as it determines outcomes such as earnings later in life. Earnings as well as the level of education are affected by the difference in family background (Becker and Tomes 1976). As Ejrnæs and Pörtner (2004) point out, intra-household resource allocation by parents is equally important. Based on Philippine data, half of the difference in completed education is explained by difference within the family. Among factors that affect resource allocation of parents are birth order and sex composition of siblings, and these are the focus of this paper. Birth order and sex composition affect the intra-household resource allocation via parental preferences, difference in prices of human capital investment, and/or child endowments. Birth order and sex are among the few endowments of children that are easily observable by researchers (Behrman 1988a). Endowments include both genetic and environmental factors and due to their effects, are incorporated in modeling human capital investment in children (Becker and Tomes 1976, Behrman *et al* 1982). Behrman and Taubman (1986) explicitly treat birth order as an endowment and incorporate it into their model. For example, higher birth order children may be associated with lower birth weight as they are born to older mothers and thus will have a lower endowment. Further, higher birth order children with older sisters may be in better educational environments because their elder sisters take care of them. Whether parents make a larger investment in the more endowed (reinforcement strategy) or less endowed (compensation strategy) child remains an empirical question.

Despite extensive study, no strong conclusion has been reached concerning the effect of birth order on outcomes of children. One of the biggest obstacles to a firm conclusion is endogeneity of fertility. This is a problem in analysis of birth order effects since birth order is the realization of parental fertility decisions (Ejrnæs and Pörtner 2004). A related problem is the difficulty in separating birth order from family size effects. Obviously, higher birth order children are from larger families. Higher birth order may be associated with worse outcomes, but this may simply be the result of quantity-quality trade-offs (Becker and Lewis 1973). Sources that contribute to birth order effects are not clear yet, either. Lower birth order children may be better off because mothers with a fixed endowment of time spend more time with them. Conversely, higher birth order children may be better off because they face lower financial constraints in households where parents face borrowing constraints and increase their earnings during their life cycle. The birth order effect likely depends on environmental factors including cultural context. Hence, it is no surprise that previous literature indicates different birth order effects across countries. In studying the birth order effect, it is important to take cultural context into account.

This paper is especially concerned with how birth order and sex composition of siblings affect parental decision-making on the human capital investment in the Indian context. There are many studies that look at the sibling sex composition effect.[1] However, these studies do not satisfactorily control for endogeneity of fertility.[2] Rosenblum (2010), who has studied mortality in children as an outcome of human capital investment, addresses the endogeneity problem using sex of the first-born child as a proxy for sex composition and family size, but cannot distinguish sex composition from family size effects.

The model in this paper incorporates cultural contexts of India (such as patrilocality, the practice of dowries, and low participation of female labor in income-generating activities) into the budget constraint. The budget constraint in this model is based on the assumption that sons bring future income into the household depending on their human capital while daughters require dowry payment from the family. Only with these assumptions, is it possible to derive the theoretical implication that boys are better off while girls are worse off with a higher proportion of sisters and vice versa with a higher proportion of brothers, given the total number of children in the household. Interestingly, the degree of "son preference" does not play a role in deriving this result.

This paper explores birth order and sibling sex composition effects in detail using the Indian Human Development Survey (IHDS). Test scores of reading, math, and writing of children aged 8-11 are used as a proxy for human capital. It is not uncommon for research of this kind to use test scores to measure the level of human capital (Hanushek 1992, Kaestner 1997, Iacovou 2007). The possibility of the selection bias problem using test scores is addressed in Section 4 and Section 6.4. The test was administered by the enumerators of the IHDS. Three

[1] Previous studies on sibling sex composition effects include those of Butcher and Case (1994), Kaestner (1997), and Hauser and Kuo (1998) in the US; Parish and Willis (1993) in Taiwan; Garg and Morduch (1998) and Morduch (2000) in African countries; Das Gupta (1987), Muhuri and Preston (1991), Pande (2003) and Rosenblum (2010) in South Asian countries.

[2] Parish and Willis (1993) do not control for the endogeneity of family size in their main text. They only mention the within-family estimates in a footnote and in the appendix. Garg and Morduch (1998) use the number of children born alive as an instrument for the number of children and the number of sisters, but the number of children born alive obviously depends on parental choice.

approaches are used to deal with endogeneity. First, a 2SLS estimation is conducted controlling for the number of children. Second, effects of sex composition and household size on outcomes are viewed, exploiting the sex of the first-born child. Third, test scores of children within the household are examined by controlling for household fixed effects. This may be the strongest contribution of this paper to the literature on birth order and sibling sex composition effects because unobserved household heterogeneity is completely controlled. Effects by sex and sibling composition are also decomposed in order to examine detailed birth order effects. Concretely, how the existence of older (or younger) brothers (or sisters) affects the outcomes is examined. This decomposition is important since the birth order effect can be derived from a combination of cultural factors such as patrilocality, seniority, and so on. The results consistently show that boys are worse off when they have more brothers (especially older brothers) but are better off with more sisters (especially older sisters). In contrast, there is no strong evidence for birth order and sibling sex composition effects for girls when birth order effects are decomposed by sex and sibling composition, or household heterogeneity is completely controlled. This is different from results found by Das Gupta (1987), Pande (2003), and Rosenblum (2010), all of whom use Indian data. The difference may come from the fact that endogeneity of fertility is addressed and sex composition effects are distinguished from family size effects in this study. Results of the research reported in this paper indicate some possible mechanisms that vary human capital investment across birth order: (1) Boys with more older sisters may be more valued as they can be the only source of future income at the time of their birth, (2) Parents may invest more in the lower birth order boys whose return to human capital investment is higher from the parental point of view in the patrilocal society where lower birth order boys are especially relied upon to take care of their parents in the future.

2. Background

Human capital investment in children has been extensively studied since Becker (1993) first formalized it into a model. According to the wealth model (Becker and Tomes 1976, Becker 1991), when parents are wealthy enough to do both human capital investment and bequests, they invest in human capital of children until the return to education is equal to the market rate of return. This conclusion is the same as in the pure investment model with the perfect capital market. It is not applicable, however, when parents are poor or they face credit constraint.

Behrman *et al* (1982) consider the separate earnings and transfer model (SET model) in order to analyze the relationship between child endowments and human capital investment. How human capital investment in children is related to their endowments is an empirical question. One hypothesis is that parents invest their resources in order to maximize returns of the household. In this case, the most endowed child will get all the resources (pure investment or reinforcement strategy). Another hypothesis is that parents care about the equality of their children and invest their resources more in the less endowed child (compensation strategy).

Behrman and Taubman (1986) extend their SET model and explicitly incorporate the birth order effect into it. They treat birth order as an endowment of children that is easily observed by researchers. Despite further extensive empirical studies on the birth order effect, only a few deal convincingly with endogeneity of fertility (Ejrnæs and Pörtner 2004, Black *et al* 2005, Conley and Glauber 2006, Kantarevic and Mechoulan 2006). Children with more siblings are supposed to be worse off because scarce resources are diluted among siblings, and each share of the pie becomes smaller (quantity-quality trade-offs; see Becker and Lewis, 1973). However, parents who are more concerned with the earning capacity of their children in the future may decide to have fewer children. Birth order can be a parental choice as in the case of the number of children. In a society with strong "son preference", parents may decide not to have any more children when they have a boy ("son-preferring, differential stopping behavior (SP-DSB);" see Jensen, 2005). Alternatively, the last born may be the choice of parents according to the stopping rule when they have an abler child.

Most studies also confound birth order and family size effects. Higher birth order children naturally come from a family of larger size. The observed worse outcomes of higher birth order children may not be due to their birth order but because they have more siblings (quantity-quality trade-offs). Most studies typically include birth order dummies as explanatory variables (Behrman and Taubman 1986), and these may simply capture the family size effect. There appear to be only four studies that take into account this problem (Ejrnæs and Pörtner 2004, Black *et al* 2005, Iacovou 2007, Booth and Kee 2009). There is no study in the context of South Asia that separates the birth order from the family size effect.

Not surprisingly, the mechanism that generates the birth order effect is far from conclusive. The difference in intra-household resource allocation among siblings depends on child endowments, both genetic and environmental, as well as parental preferences. Parental preference and environmental factors can vary across countries and societies because they are dependent on inheritance practices and the level of development (Booth and Kee 2009). There are various hypotheses in the literature regarding sources of birth order effects. Those predicting a negative effect (higher birth order children have lower outcomes for example) include: (1) Higher birth order children are born to older mothers and tend to have lower birth weights, and this may lead to lower outcomes in the future (Ejrnæs and Pörtner 2004), (2) Time spent by mothers for childcare is more diluted for later-born children (Behrman and Taubman 1986, Behrman 1988a, Booth and Kee 2009), (3) Later-born children may face more financial constraints as parents are running out of funds for education (Behrman and Taubman 1986, Kantarevic and Mechoulan 2006, Iacovou 2007), (4) Older siblings can increase their intellectual ability by teaching their younger siblings (Behrman and Taubman 1986, Behrman 1988a, Blake 1989), (5) In a society where the eldest child takes responsibility for taking care of his/her parents when they become old, they invest more in the eldest child (Ejrnæs and Pörtner 2004). On the other hand, those predicting a positive effect include: (1) Mothers are better at taking care of children as they accumulate experience (Behrman and Taubman 1986, Behrman 1988a, Blake 1989, Ejrnæs and Pörtner 2004), (2) Older siblings take care of younger siblings (Parish and Willis 1993, Booth and Kee 2009), (3) Life-cycle predicts that parents increase their earnings and can better afford to provide education for later-born children (Behrman 1988a, Parish and Willis 1993, Ejrnæs and Pörtner 2004, Black *et al* 2005, Booth and Kee 2009) (4) Lower birth order children may be forced to work at a younger age in order to support the family (Parish and Willis 1993, Ejrnæs and Pörtner 2004, Booth and Kee 2009) (5) The last-born will do best because parents will stop having more children when they have a high-ability child (Ejrnæs and Pörtner 2004, Black *et al* 2005).

In the Indian context, the difference in human capital investment between sons and daughters is often called "son preference". It may reflect pure parental preference for sons as pointed out by Behrman (1988b),[3] or it may simply reflect the difference in returns to human capital investment between them (Rosenzweig and Schultz 1982). India is predominantly a patrilocal society where sons stay with their parents, and daughters leave their parents after

[3] The border between pure "son preference" and returns may be ambiguous, as pointed out by Behrman (1988b). He indicates that pure "son preference" of parents can be explained in terms of returns such that parents gain more pleasure from sons than from daughters. He also says, however, that such extreme definition of returns, which includes almost everything concerning parental "son preference", may not be useful.

marriage. Sons bring dowries into their family at the time of marriage, while parents have to pay dowries when their daughters get married. Das Gupta (1987) argues that underinvestment in daughters is not general across birth order. She shows, based on data in Rural Punjab in India, that the burden of excess child mortality falls selectively on higher birth order girls. Consistently with the findings of Das Gupta (1987), using a dataset of all of India, Pande (2003) shows that higher birth order girls have worse health outcomes. Both of them attribute the negative birth order effect for girls to parental preferences for sons. In a society with strong "son preference", girls who are born into households that already have girls are likely to be less valued. Rosenblum (2010) has shown that daughters in first-born girl households are worse off while those in the first-born boy households are better off. He attributes this to the fact that first-born girls make a larger family size (family size effect) and a higher proportion of girls in the household (sex composition effect). These results are different from those using data in Taiwan (Parish and Willis 1993) and in African countries (Garg and Morduch 1998, Morduch 2000) that show that the schooling or nutritional outcome of children is better when they have older sisters. The difference in results may be partly due to differences in the cultural context, and this reinforces the importance of taking cultural differences into account when birth order effects are studied. It may be partly because none of these studies, with the exception of Rosenblum (2010), satisfactorily control for endogeneity of fertility. Rosenblum exploits the "natural experiment" (sex of the first-born child) in order to exogenously control for the number of children and sex composition of the household. However, this method does not distinguish sex composition effects from family size effects.

3. Model of Investment in Children with Sibling Sex Composition

The model of the human capital investment in children used in this paper is a modification of Rosenzweig and Schultz (1982) as well as Rosenblum (2010). The outcome of human capital investment in their model is the mortality rate, but the number of households who have ever lost a child is limited (about 17 percent of households in the survey of this study). Therefore, their model is not general in the context of contemporary India. The basic assumption of the model is that parents care about the future welfare of their sons and daughters. The level of care between sons and daughters does not have to be equal and can be far apart. Since intra-household allocation between husband and wife is not an issue in this paper, a single decision maker in the household is assumed. Parents thus unitarily maximize the following utility function:

$$U = c(x) + \pi\rho_m c_m(H_m) + (1-\pi)\rho_f c_f(H_f) \tag{1}$$

where $c(x)$ is the current consumption of parents, and $c_i(H_i)\ i = m, f$ is the future consumption of children with the subscript *i* denoting male or female children. These consumption functions satisfy the standard positive, strictly concave, and continuously differentiable features ($c > 0, c' > 0, c'' < 0$). Future child consumption depends on the level of human capital, H_i which is invested by parents. Though women usually do not work outside the household in the Indian context, it seems reasonable to assume that future consumption of daughters is determined by their level of human capital due to assortative mating in the marriage market (Behrman *et al* 1995) and the fact that human capital positively affects the bargaining position in the conjugal family (Makino 2011). The discount factor, ρ_i, is not necessarily equal between sons and daughters, and this allows parental pure "son preference" in Indian society. Given the number of total children in the household, π is the continuous proportion of male children, and $(1-\pi)$ is that of female children. Note that while endogeneity of fertility is treated in the empirical analysis, it is not endogenized in the model here because the primary concern is in birth order and the sibling sex composition effects given the total number of children.[4]

Parents face budget constraints given by:

$$V + \pi R_m(H_m) - (1-\pi)R_f(H_f) = x + \pi H_m + (1-\pi)H_f \tag{2}$$

where V is parent income, which is exogenous in the model, and $R_i, i = m, f$ is the discounted contribution of male and female children to family resources. India is predominantly a patrilocal society, and sons live with their parents after they marry; thus, sons contribute to their parent's household resources while daughters do not. Earning capacity of sons depends on their level of

[4] The model can be easily manipulated in order to endogenize fertility by simply including the number of children and attaching it to the proportion of each sex siblings, as is done in the model of Rosenblum (2010). The inclusion does not change theoretical implication derived in the partial equilibrium framework.

human capital, H_m. This earning function of sons is positive, strictly concave, and continuously differentiable ($R_m > 0, R_m' > 0, R_m'' < 0$). $R_m(H_m)$ is interpreted to include contributions of wives and children. On the other hand, at the time of marriage, daughters leave their parents with dowries; thus there is a negative sign in front of their contribution R_f. Dowries are assumed to be non-increasing functions of the human capital of daughters ($R_f > 0, R_f' \le 0, R_f'' \ge 0$). Note that this assumption does not exclude the case where dowries are independent of the human capital of daughters ($R_f' = 0, R_f'' = 0$).[5] The price of parental consumption, x, and the price of human capital, H_i, are normalized to 1.

Maximizing the parent utility function (1) subject to their budget constraints (2) gives the following first order conditions:

$$[H_m] \quad c'R_m' - c' + \rho_m c_m' = 0 \tag{3a}$$

$$[H_f] \quad -c'R_f' - c' + \rho_f c_f' = 0 \tag{3b}$$

Solving (3a) and (3b) simultaneously yields the demand functions of human capital investment, $H_m = H_m(\pi, \rho_m, \rho_f, V)$ and $H_f = H_f(\pi, \rho_m, \rho_f, V)$. Plugging these demand functions back into the first order conditions (3a) and (3b), and differentiating them with respect to π yields:

$$\frac{\partial H_m}{\partial \pi} = \frac{1}{\Delta}\left[c''\left(-H_m + H_f + R_m + R_f\right)\left(\rho_f c_f'' - c'R_f''\right)\left(1 - R_m'\right)\right] < 0 \tag{4a}$$

$$\frac{\partial H_f}{\partial \pi} = \frac{1}{\Delta}\left[c''\left(-H_m + H_f + R_m + R_f\right)\left(\rho_m c_m'' + c'R_m''\right)\left(1 + R_f'\right)\right] > 0 \tag{4b}$$

---

[5] It is not yet conclusive whether dowries are prices or bequests. Economists following Becker's price model (1991) have interpreted dowries as prices that clear the marriage market. According to the price model of dowries, human capital of brides decreases the amount of dowries while human capital of grooms increases it. Contribution of sons $R_m(H_m)$, which is the increasing function of human capital, can be interpreted to include the dowries they receive at the time of marriage. Negative contribution of daughters, $R_f(H_f)$, is decreasing in human capital. However, there is little empirical evidence supporting the price model of dowries (See Zhang and Chan 1999, Brown 2009, Makino 2011). Even research supporting the price model shows that only women's literacy is important in determining the amount of dowries, not women's schooling (Behrman *et al* 1999). Based on these empirical findings, the model does not exclude the case where dowries are bequests and independent of the human capital of daughters ($R_f' = 0, R_f'' = 0$).

where

$$\Delta = c''(1-\pi)\left(c'R_m'' + \rho_m c_m''\right)(1+R_f')^2 - c'\pi\left(c'R_f'' - \rho_f c_f''\right)\left(1-R_m'\right)^2 - \left(c'R_f'' - \rho_f c_f''\right)\left(c'R_m'' + \rho_m c_m''\right) > 0$$

by the second order condition. The negative sign of (4a) means that, given the total number of siblings, a higher proportion of boys decreases the human capital investment in boys. The positive sign of (4b) means that a higher proportion of male siblings increases the human capital investment in girls. Intuitively, the rationale behind the results above is that when parents have a smaller proportion of boys, they have more incentive to invest in them in order to increase the contribution of sons to family resources so that they can bear the burden of having a higher proportion of girls. One may think that the signs of (4a) and (4b) are indeterminate, but it is possible to determine the signs of these terms as above with the following considerations: First, when the first order conditions (3a) and (3b) are examined, it is clear that $1-R_m' < 0$ and $1+R_f' > 0$. Second, a realistic assumption based on the previous literature is that $H_m < H_f + R_m + R_f$. This inequality means that the total cost[6] of human capital of sons is smaller than the sum of their total contribution to family resources, the amount of dowries, and the total cost of human capital of daughters. This assumption is maintained because human capital investment in children even in South Asian countries is more or less associated with its economic returns (Rosenzweig and Schultz 1982, Behrman 1988b, Foster and Rosenzweig 1996). It is also because dowries in India are reported to be several times above annual household income (Rahman and Rao 2004, Bloch *et al* 2004, Anderson 2007, Makino 2011) or to share a substantial portion of household assets (Rao 1993).[7] Note that, based on the previous studies, the inequality, $H_m < H_f + R_m + R_f$, holds despite the gender gap in the human capital investment, $H_m > H_f$, in Indian society.

The above theoretical implication is derived from one single assumption regarding the contribution of boys versus girls to households based on being a patrilocal society and the custom of dowries. In particular, the above implication is not affected by the level of "son preference" that parents have which is captured by the parameter ρ_i. For example, even if

[6] Remember that the cost of human capital is normalized to 1 in the budget constraint (2).

[7] As an example, Rosenblum (2010) derives the deterministic sign in his model of comparative statics only with the assumption of a "sufficiently high" amount of dowries.

parents do not care about the welfare of their daughters at all (in this case, $\rho_f = 0$), the signs of equations (2.4) and (2.5) are maintained.

4. Data

In order to explore the birth order and the sibling sex composition effects, the Indian Human Development Survey (IHDS) collected from November 1st 2004 to October 30th 2005 is used. The survey covers 41,554 households in 1,503 villages and 971 urban neighborhoods across India. The IHDS includes the complete birth history of "eligible women" defined as those ever married between the ages 15 and 49. Since the household roster of the IHDS records the relationship of each household member to the head of the household (such as wife, son, daughter, daughter-in-law, grandchild, nephew, niece, etc.), birth history of wives of heads of households can be exactly matched with their children. However, we cannot exactly know, for example, who the mother of the head's nephew is. Therefore the sample for this study is restricted to children of the heads of household. Since the interest of this study is in birth order effect, twins are also excluded from the sample as is common in the study of birth order effect (Black *et al* 2007).

The IHDS reports the test scores of children aged 8 to 11. It consists of reading, math, and writing tests, and these are used as a proxy for human capital. The reading test is recorded in five discrete variables: cannot read (=0), letter (=1), word (=2), paragraph (=3), and story (=4).[8] The math test is recorded in four discrete variables: cannot count (=0), number (=1), subtraction (=2), and division (=3). The writing test is recorded in two discrete variables: cannot write (=0) and write with two or less mistakes (=1). Test scores of reading and math are normalized so that the highest score is equal to 1. Table 1 reports the summary statistics of the normalized test scores.[9] Boys perform better than girls on all three exams. The difference between outcomes for boys and girls in reading and math tests is statistically significant at the .05 level of significance. The standard deviation is higher for girls in all three exams, and this implies that human capital investment is more diversified among girls.

Though the test score is just one intermediate aspect of human capital, it is not uncommon for research of this topic to use test scores to measure the level of human capital

[8] Since these discrete variables are not cardinal but rather ordinal, a robustness check is conducted with the ordered logit model. The results are not qualitatively different from the results given by the OLS model (Table 4).
[9] Complementarily, Appendix Table A.1 shows the distribution of the raw test scores.

(Hanushek 1992, Kaestner 1997, Iacovou 2007). In fact, there is evidence that early success in school leads to greater educational achievement, and the educational achievement of school-age children may provide a more direct examination of the sibling sex composition effect (Kaestner 1997).[10] There are some advantages to using the test scores of the IHDS. First, since the test is administered by enumerators of the survey, the scores are not subject to recall error or false report. Second, the test is taken by children aged 8 to 11 at the time of survey, and the cohort effect of children can be eliminated. Controlling for the cohort effect is important in studying the birth order effect (Blake 1989). Third, determining how to decide the cut-off age in order to extract the level of completed schooling of children is not of concern. Researchers often use the level of completed schooling as a measure of human capital investment, and this requires them to restrict their sample, such as over the age of 25 (see Black *et al* 2005), in order to ensure that all children have completed their education.[11] In doing so, however, it becomes more difficult to get a dataset with a complete list of siblings and their completed level of education (Jensen 2005).

One concern with using the test scores of children in this study may be a selection bias or a possible correlation between birth order and whether or not the test was taken. For example, suppose that girls are less likely to take the test. Then the girls who took the test may have higher unobserved ability than the girls who did not take the test due to the possibility that parents have more incentive to let their higher ability children take the test. The unobserved ability may also be correlated with birth order. If this selection bias is present, all estimation results are misleading. However, there seems to be no serious concern about it, as no evidence can be found that birth order or sex of children affect test-taken status.[12]

Table 2 shows the summary statistics of 8,253 children aged 8 to 11 who took all three exams. The average number of children in the household is 3.7. The average age of their mothers is 34, and their age at marriage is 17. Only 46 percent of their mothers are literate, and their average schooling of mothers is 3.4 years while that of fathers is 5.8 years. Thirty three percent of them live in urban areas, and 27 percent live in the villages where the practice of endogamous

[10] Kantarevic and Mechoulan (2006) suggest that some factors early in life contribute to the first-born premium shown in their empirical results.
[11] For an alternative method not requiring the complete years of schooling, see Ejrnæs and Pörtner (2004).
[12] See Appendix Table A.6. The estimated coefficients for birth order or sibling sex composition (i.e., number of brothers or sisters index) and sex of children are extracted from the estimation models of the birth order and the sibling sex composition effects on test-taken status of children, and are presented in Table A.6.

marriage is common.[13] The village-level variable is included since it is often argued, in the Indian context, that the status of women is higher and the socio-economic indicators of mothers and children are better in societies where endogamous marriage is common (Dyson and Moore 1983, Pande 2003, Rahman and Rao 2004). Caste, religion, and state characteristics are also controlled in the empirical analysis.

Table 2 also shows summary statistics by sex of children. The means of the reported socio-economic variables are not statistically different between boys and girls at the .05 level of significance, except for the number of children and the dummy variable for Sikh. The number of children in the household is significantly larger for girls than for boys. This is consistent with the idea of "SP-DSB" as argued by Jensen (2005). In a society with strong "son preference", parents may decide not to have children anymore when they have a boy. The "SP-DSB" implies that girls are more likely to be from a larger family than boys. This simple evidence underscores the importance of taking endogeneity of fertility into account in studying the birth order effect. Otherwise, the true family size effect may be interpreted incorrectly as the negative birth order effect for girls. Girls are less likely to be from a Sikh family. Sikhs may be more likely to use sex selective abortion in order to have an ideal number of boys.[14]

5. Estimation Strategy

Studies of birth order effect have to take endogeneity of fertility into account. The birth order effect should also not be confused with the family size effect. The 2SLS estimation strategy is thus first used in order to deal with endogeneity. The instruments for fertility include (1) the village-level accessibility to the clinic that offers several measures of family planning services at first birth, (2) the number of brothers and sisters of the father, and (3) the sex of the first-born child. Though the sex composition of children is often used as an instrument of fertility (see Angrist and Evans 1998), it is not an appropriate instrument in the context of India where the sex composition of children is not exogenous. "Missing women" is a well-known phenomenon in India (Sen 1990, Anderson and Ray 2010), and the ratio of males to females is skewed due to sex-selective abortion, girl infanticide, and higher mortality of girls. However, it

[13] The variable of endogamous marriage takes the value 1 if the marriage satisfies one of the following conditions: (1) Any member in the mother's family has been married into the father's family; (2) The mother is related to the father by blood; (3) The mother grew up in the same village/town as the father.
[14] The previous version of Pörtner (2010) confirms this point.

can be shown that first-birth has the normal male-female ratio, and thus the sex of the first-born is exogenous (Pörtner 2010, Rosenblum 2010). The sex ratio of the first-born in the dataset used in this study is also in the normal range and can be considered to be random.[15] One may question the exogeneity of the availability of village-level family planning services and the number of siblings of the father. In addition to the test of over-identifying restrictions, the sensitivity of the results is checked by using a different set of instruments in the next section. As for the measures concerning the availability of family planning services, five dummy variables are constructed taking the value 1 if each of the five contraception methods (oral pills, IUD insertion, male sterilization, female sterilization, and injection) was offered by the clinic in the village at the time when the mother was first pregnant. Alternatively, the index variable (0-5) summing up all the dummy variables is used in place of the five dummies to determine which is the better instrument. Table 3 shows the results of the first-stage regression with fertility as a dependent variable. The inclusion of five dummies seems to fit slightly better when R-squared values are compared. Five dummies are thus included in the main estimation. The use of the index variable (0-5) instead of the five dummies does not substantially change the results. In both specifications, the problem of a weak instrument is refuted with F-statistics showing the joint significance of the instruments. The village-level availability of family planning services at first birth significantly reduces the number of children. Also, parents are more likely to have additional children when their first-born is a girl, and this is consistent with the idea of "SP-DSB."

The birth order effect is separated from the effect of total number of household children using the relative birth order proposed by Ejrnæs and Pörtner (2004) and also the birth order index proposed by Booth and Kee (2009). The relative birth order is constructed as $(b-1)/(n-1)$, where b is the absolute birth order, and n is the number of children in the family. The birth order index is constructed as b/A, where b is the absolute birth order, and A is the average birth order of the family calculated as $(n+1)/2$. Both methods generate a variable for birth order which is less correlated with family size than the absolute birth order. The correlation between the total number of children and the absolute birth order is 0.727, while that between the

[15] The sex ratio of the first-born, the second and the third-born, the fourth-born and the higher in the dataset is 1.061, 1.084, and 1.107, respectively. The sex ratio is calculated for those younger than 12 years old in order to alleviate the recall error of mothers, as is pointed out by Pörtner (2010) and Rosenblum (2010). The sex ratio of the first-born in the dataset used in this study, 1.061, is considered normal, while the higher sex ratio of higher birth orders is consistent with the literature.

total number of children and the relative birth order, and the birth order index is 0.127 and 0.181 respectively. It seems, in this dataset, that the relative birth order is a better measure than the birth order index in terms of its relative independence from the total number of children. The relative birth order is thus used in the main estimation. Whether or not the use of the birth order index instead of the relative birth order affects the results is also checked.

Since there are only four options for ages of children (i.e., 8-11), dummy variables for ages of children instead of actual ages are included in order to completely remove the age effect. The results are not substantially different relative to which variable is used.

Sources of the birth order effect are also important, and these are explored by examining the sibling sex composition effect. The index of the number of older (or younger) brothers (or sisters) is used in place of the relative birth order. The number of older siblings is sometimes used as a measure of birth order (see Morduch 2000 for example), but it is not independent of the number of children in the family. Though it is not perfect, the best feasible measure is to treat the endogeneity of the number of children and the number of older (or younger) brothers (or sisters) with the instruments explained above. The theoretical implication in Sections 3 is that, controlling for the number of children, the number of brothers has a negative impact on outcomes for boys, while the number of sisters has a positive impact. Conversely, the impact of the number of sisters is negative and that of brothers is positive for girls.

It is also useful to investigate the mechanisms of the birth order effect by looking at, for example, birth interval or the existence of elder siblings who have completed their education. If maternal time is an important investment in human capital of children, a longer birth interval will have a positive effect on human capital investment. Or, if older siblings become an additional income source to the household, human capital investment will increase for children who have older siblings with higher levels of completed education. The problem of including the birth interval and the educational level of siblings aged 18 and older, however, is that these variables may also be endogenous. Therefore, they are not included in the main estimation but only checked relative to whether or not results change when these variables are included and what the coefficients of these variables are.

Secondly, the estimation conducted by Rosenblum (2010) is replicated and extended. He deals with both endogeneity of fertility and sex composition in the household by using a natural experiment, specifically the sex of the first-born. His estimation allows assessment of the impact of the sex of the first-born on outcomes (in his case, the number of children or the mortality of them) given by the following estimation equation:

$$Y_{ijk} = \gamma FirstBornGirl + \beta_1 X_j + \beta_2 State_k + e_{ijk} \qquad (5)$$

where $FirstBornGirl$ takes the value 1 when the sex of the first-born is a girl, X_j includes age and sex of the child i taking the tests and a set of socio-economic characteristics of the household j shown in Table 2, and $State$ is a set of state dummies. The included age variable is not the actual age of children but the dummy variables (ages = 9, 10, or 11 years old) as in the estimation above. Y_{ijk} is the outcome as measured by test scores of child i of birth order two or higher in household j in state k. The interest of this study is in the coefficient of $FirstBornGirl$, γ. The problem in estimation model (5) is that γ captures both the family size effect and the sibling sex composition effect. The family size effect predicts a negative sign for γ. On the other hand, the theoretical implication in Section 3 is that γ should be positive for outcomes of boys and negative for girls. In order to separate the sibling sex composition effect from the family size effect, the estimation model (5) is extended by controlling for the number of children using the set of instruments shown in Table 3, except for the sex of the first-born child.

Finally, family heterogeneity is controlled completely by including household fixed effects. This is the strength of this study, since by doing so, the birth order effect can be completely separated from the family size effect. Also, any correlation between household unobservables that can affect birth order and outcomes, such as test scores, is completely wiped out. There are 1,445 households in which more than two children aged 8 to 11 took the exams. The estimation equation is given by:

$$Y_{ij} = \gamma BirthOrder_i + \beta_1 ChildAgeDummies_i + \beta_2 ChildSex_i + Household_j + e_{ij} \qquad (6)$$

The coefficient of interest is γ; this captures the birth order effect within the household. Since the unobserved household characteristics as well as fertility are captured by the household fixed effects, in addition to the relative birth order and the birth order index, the absolute birth order is also used in order to check the robustness of the results in the next section. The use of absolute birth order enables us to see the economic significance of the birth order effect.

6. Results

6.1. OLS and 2SLS Estimation for Birth Order and Sibling Sex Composition Effects

Table 4 and Table 5 contain the OLS and 2SLS estimation results of the birth order effect, respectively. In the main estimation, relative birth order is used as a measure of birth order since it outperforms the birth order index in that it is less correlated with the number of children than the birth order index. 2SLS estimation results are reported using the birth order index in the Appendix (Table A.2), which essentially shows similar results to those in Table 5. Five dummies of the availability of each of the family planning services are used as instruments instead of the index variable (0-5) since the use of the five dummies fits better in the first-stage regression. The use of the index variable (0-5) does not largely affect the results (Appendix Table A.3). Over-identifying restrictions are not rejected (Sargan statistics 2.09-3.82 with p=0.80-0.95 depending on the test types), which suggests the validity of including all these instruments.[16] The 2SLS estimation shows that higher birth order girls do better on all the exams than lower birth order girls, while there is no birth order effect for boys. The positive birth order effect for girls is contrary to what Das Gupta (1987) and Pande (2003) found as evidence for parental pure "son preference". The difference may be because this study takes into account endogeneity of fertility and deals with separating birth order and sibling sex composition effects from the family size effect. Naturally, older children have a higher score. Boys perform better on all the exams than girls. The estimated coefficients of the household socio-economic characteristics show expected signs. Children have higher scores when their mothers and fathers have higher education, and their mother's age at marriage is higher. Children from urban areas and from higher castes have higher scores. Also, Muslim children have a lower score compared with Hindu children. The main difference between OLS and 2SLS estimation results is that the coefficient of the number of children becomes insignificant or even positive in the latter. One may argue that it is

[16] In addition to the test of over-identifying restrictions, estimation procedures are repeated by dropping one or two sets of instruments in order to check the sensitivity of the results to a different set of instruments. These results do not alter the main result with the complete sets of instruments, though the significance level decreases when sex of the first-born child is dropped (Appendix Table A.4).

counterevidence to quantity-quality trade-offs, but given the large standard errors, it is difficult to interpret anything meaningful regarding the estimated coefficient of the number of children.

Although the birth interval and the completed years of education of siblings aged 18 and older may be endogenous, how the main results change can be viewed by including these variables in order to explore the sources of the birth order effect. The inclusion of these variables does not affect the implications given by the main results in Table 5. The coefficient of birth interval shows that when the age difference from the immediate older sibling is larger, children have higher scores. Also, the completed years of education of siblings aged 18 and older have a significantly positive effect on their scores. The former may imply that a mother's time is an important investment in human capital of children. The latter may imply that the older siblings become an additional resource to the household for later born children.

The sibling sex composition effect is explored by decomposing siblings of children into older/younger brothers and sisters. Six-by-nine[17] different specifications are performed treating two endogenous variables (the index of the number of children and the number of older/younger sisters/brothers) with the same set of instruments used in 2SLS estimation in Table 5.[18] The over-identifying restrictions are not rejected in all 54 specifications, thus supporting the validity of the set of instruments. Relative birth order in Table 5 is replaced by the index of sibling sex composition. Other included exogenous variables are the same as in Table 5.

Estimated coefficients of the index of sibling sex composition (older/younger sisters/brothers index) on the test scores are extracted in Table 6. This shows that boys have a higher score when they have older sisters. The positive effect of having sisters on outcomes of boys is consistent with the theoretical implication. However, having younger sisters has no effect on these outcomes, and this may imply that boys who are born to households that already have girls are valued more, as parents would count on and invest in them more. On the other hand, having older brothers appears to have some negative effects on outcomes for boys. There is no significant effect of having older/younger sisters/brothers on outcomes for girls. No sibling sex composition effect for girls combined with the higher standard deviation of test scores for girls compared to those of boys implies that the variation in human capital in girls mainly comes from

[17] Estimation for both sexes/boys only/girls only on each reading, math, and writing test scores are conducted (three by three = nine specifications). Sibling sex composition is also decomposed into having sisters, brothers, older sisters, older brothers, younger brothers, and younger sisters (six specifications).
[18] 2SLS estimations are also performed with one endogenous variable (that is the number of children). The results are not qualitatively different from those in the main estimation with two endogenous variables.

differences across families. The result of no sibling sex composition effect for girls is different from that of Garg and Morduch (1998), which shows that both boys and girls are better off with the existence of older sisters. The difference of results in this study may come from the different cultural contexts and the fact that more convincing instruments are used in order to treat the endogeneity of fertility (remember that they use the number of children born alive as an instrument, which is obviously endogenous). The results in Table 6 combined with Table 5 imply that while there is a positive birth order effect for girls, it is not derived from sibling sex composition. Later born girls are equally better off than earlier born girls when they have older brothers and when they have older sisters.

6.2. Replication and Extension of Rosenblum (2010)

The estimation given by (5) is conducted, and this is essentially a replication of Rosenblum (2010). Table 7 extracts the OLS estimation coefficient, γ, in the estimation equation (5); this is the effect of having a first-born girl on the test scores of children of birth order two or higher. Having a first-born girl has a significantly positive effect on the reading test for boys and a negative effect on that for girls, and this is consistent with Rosenblum (2010). Since the variable “first-born girl” captures both the family size effect and the sex composition effect, the OLS estimation used by Rosenblum (2010) cannot differentiate the sex composition effect from the family size effect.

Rosenblum’s empirical analysis is extended by including the number of children treated with the set of instruments used in Table 5, except for “first-born girl.” The 2SLS estimation result for γ is also extracted in Table 7. When the sex composition effect is separated from the family size effect, it is clear that “first-born girl” has a larger effect on all the test scores for boys and is significantly positive on the reading and math scores. It implies that, given the number of children in the household, a higher proportion of girls has a positive effect on outcomes for boys. Although sibling sex composition has no significant effect on outcomes for girls, as is consistent with the results in Table 6, the estimated coefficients are negative for all the test scores, and the magnitude of the effects are larger than that for boys. The large standard errors of the estimated coefficients for girls seem consistent with the larger standard deviation of girls’ test scores and the previous results implying the large variation in outcomes for girls across families. Note that the number of children has a negative effect on boys' scores though it is insignificant. This shows the importance of separating out the sex composition effect from the family size effect, because the family size effect and the sibling sex composition effect are both generated by the "first-born girl" and may work in the opposite direction. The significantly negative effect of "first-born girl" on the outcomes of girls shown by Rosenblum (2010) and replication in this study of his estimation (shown in the first row of Table 7) may be simply the family size effect because his estimation confounds both.

6.3. OLS Estimation with Household Fixed Effects

Finally, household heterogeneity is controlled. The estimation results based on (6) are shown in Table 8. Any measure of birth order has a significantly negative coefficient, especially for reading and math scores. This means that higher birth order children perform worse on these exams. The coefficient of the absolute birth order indicates that increasing birth order by one leads to 0.05 point lower scores on all the exams. This is economically significant based on the fact that the highest score is normalized to 1, and that the difference in the mean score between boys and girls is 0.02 in the reading and the writing tests, and 0.04 in the math test.

Because the number of observations becomes small when household fixed effects are included, and because two children from the same household included in the regression are not necessarily the same sex, the estimation cannot be conducted by gender as in previous estimation strategies (Section 6.1 and 6.2). Instead, birth order is made to interact with the sex of children in order to further explore the birth order and the sibling sex composition effects.[19] When the interaction term is included, there is some evidence of the negative birth order effect only for boys (Table 9). The result shows that the negative birth order effect in Table 8 is mainly derived from boys. Looking further into the by-gender birth order effect (Table 10), some evidence can be seen for a negative by-gender birth order effect for boys. Also, there is some evidence of a positive by-gender birth order effect for girls (see the coefficient of the absolute by-gender birth order in the math score). This is consistent with the results shown in Table 5. The math score usually includes a significant result when compared with reading or writing scores, and this may

[19] Ozer-Balli and Sørensen (2011) warn against just putting interaction terms into a regression equation without considering robustness and implications for the interpretation of results. Results in this study are robust with the demeaned interaction terms following their suggestion. Results with the simple interaction terms are thus reported.

reflect the fact that math education is more important in Indian society when compared to other subjects. The analysis controlling for the household heterogeneity further supports the view that higher birth order boys are worse off, especially when they have older brothers.

6.4. Check for Selection Bias

As mentioned in Section 4, one may question a selection bias in this study, that means the possible correlation between child birth order and the status of whether or not they took the test.[20] For example, if girls are less likely to take the test, then girls who take the test may have better unobserved ability than girls who do not take the test. The unobserved ability may be correlated with birth order. In order to check whether or not the birth order or the sex of children systematically affects the status of test-taken, estimation procedures above are repeated by replacing the test score with a dummy variable taking the value 1 when a child took all the tests. Selection bias seems not to be a concern in this study as there appears to be no evidence that the sex or the birth order of children affects test-taken status (Appendix Table A.6).

7. Conclusion

Birth order and sibling sex composition effects are explored using test scores reported in the Indian Human Development Survey (IHDS). The simple model of human capital investment in children in the context of India suggests that boys are better off with a higher proportion of girls and worse off with a higher proportion of boys, and vice versa for girls, given the total number of children in the household. Sibling sex composition effects can exist independently of the level of the “son preference” of parents. It can be the result of one single assumption of the contribution of boys versus girls to household resources based on patrilocal society and the practice of dowries. In accord with the theoretical implication, boys are found to be worse off when they are in a family with a higher proportion of boys. In addition, empirical analysis consistently shows strong evidence of the negative by-gender birth order effect for boys. Boys

[20] It has been suggested that using current enrollment status of children instead of test scores can avoid the problem of selection bias because by doing so, all children aged 8-11 can be included. However, the use of current enrollment status of children is not useful for the objectives of this study since there seems to be little intra-household variation in current enrollment status of children (Appendix Table A.5). An exception is evidence of the negative by-gender birth order effect for boys, which is consistent with the result shown in Table 10. Estimation procedures are repeated by replacing the test score with the enrollment status of children, and coefficient estimates of birth order and sex of children are reported in Appendix Table A.5.

are worse off when they have older brothers. Higher birth order girls are better off than lower birth order girls in the most simple 2SLS estimation. Decomposing the birth order effect into the sibling sex composition effect and completely controlling for household heterogeneity, there appears to be no strong evidence of birth order and sibling sex composition effects for girls. When both results are combined, a possible interpretation is that higher birth order girls are equally better off with older brothers and with older sisters.

Detailed examination of birth order and sibling sex composition effects implies mechanisms that derive the birth order effect. Specifically, it suggests the importance of the level of contribution to household income. In a patrilocal society, boys can bring income into the household, and their wives and children are an additional labor force for the household, while girls do not bring income to the household. On the contrary, girls require dowries at the time of marriage to take with them. When the proportion of boys in the family is lower, parents have more incentives to invest in sons in order to increase the contribution of sons to their family resources so that they can bear the burden of having a higher proportion of girls. With a strong by-gender birth order effect for boys and of weak birth order and sex composition effects for girls, sibling rivalry is seen to be severe among boys while it is not strongly observed among girls. Among sons, parents tend to care for the eldest most, perhaps because the eldest son is most likely to stay with his parents after marriage and is thus most likely to contribute to the parent's household. Parents may also weigh the eldest son's role of leading a ritual or family ceremony. Other mechanisms deriving the birth order effect include time allocation of mothers and additional income sources brought by older siblings.

Boys are better off with older sisters, but welfare of older sisters is not affected by the existence of younger brothers. Boys who are born into households that already have girls may be more valued as they can be a scarce source of future income. Older sisters may take care of their younger brothers without sacrificing their own welfare. Unlike Das Gupta (1987) and Pande (2003), no evidence is found in this study that girls who are born into households that already have girls are less preferred or that higher birth order girls are the most deprived among siblings. Rather, there is some, though weak, evidence of positive birth order effect for girls. Since the endogeneity of fertility is dealt with more accurately in this study than in previous studies, these results provide new evidence for the birth order and sibling sex composition effects in India.

These results do not necessarily imply that girls are not disadvantaged. The consistently higher scores for boys on all exams suggest that the absolute level of human capital investment is lower for girls. Besides, the reason why the outcomes for girls are not strongly affected by sibling sex composition as predicted by the theoretical model is not answered in this study. Further investigating the mechanism of how the patrilocal society, the difference in income generating opportunities between males and females, and the practice of dowries generate the differences in human capital investment within the household is left for future research.

References

Anderson, Siwan. 2007. "The Economics of Dowry and Brideprice." *Journal of Economic Perspectives*, Vol.21, No.4, 151-174.

----------, and Debraj Ray. 2010. "Missing Women: Age and Disease." *Review of Economic Studies*, Vol.77: 1262-1300.

Angrist, Joshua D. and William N. Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review*, Vol.88, No.3: 450-477.

Becker, Gary S. 1991. *Treatise on the Family*. Enl. Ed. Cambridge, Massachusetts: Harvard University Press.

---------- 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Third. Ed. Chicago, Illinois: University of Chicago Press.

----------, and H. Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, Vol.81, Issue 2, Pt.2: S279-S288.

----------, and Nigel Tomes. 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy*, Vol.84, No.2, Pt.2: S143-S162.

Behrman, Jere R. 1988a. "Nutrition, Health, Birth Order and Seasonality: Intrahousehold Allocation among Children in Rural India." *Journal of Development Economics*, Vol.28: 43-62.

----------. 1988b. "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?" *Oxford Economic Papers*, Vol. 40: 32-54.

----------, Robert A. Pollak, and Paul Taubman. 1982. "Parental Preferences and Provision for Progeny." *Journal of Political Economy*, Vol.90, No.1: 52-73.

----------, and Paul Taubman. 1986. "Birth Order, Schooling, and Earnings." *Journal of Labor Economics*, Vol.4, No.3, Part2: S121-S145.

----------, Nancy Birdsall and Anil Deolalikar. 1995. "Marriage Markets, Labor Markets, and Unobserved Human Capital: An Empirical Exploration for South-Central India." *Economic Development and Cultural Change*, Vol.43, No.3: 585-601.

----------, Andrew D. Foster, Mark R. Rosenzweig, and Prem Vashishtha. 1999. "Women's Schooling, Home Teaching, and Economic Growth." *Journal of Political Economy*, Vol.107, No.4: 682-714.

Black, Sandra E. Paul J. Devereux, and Kjell G. Salvanes. 2005. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *Quarterly Journal of Economics*, Vol.120, Issue2: 669-700.

----------. 2007. "Older and Wiser? Birth Order and IQ of Young Men." *NBER Working Paper Series*, No.13237.

Blake, Judith. 1989. *Family Size and Achievement*. Berkeley, California: University of California Press.

Bloch, Francis, Vijayendra Rao, and Sonalde Desai. 2004. "Wedding Celebrations as Conspicuous Consumption: Signaling Social Status in Rural India." *Journal of Human Resources*, Vol.39, No.3: 675-695.

Booth, Alison L. and Hiau Joo Kee. 2009. "Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment." *Journal of Population Economics*, Vol.22: 367-397.

Brown, Phillip H. 2009. "Dowry and Intrahousehold Bargaining: Evidence from China." *Journal of Human Resources*, Vol.44, No.1: 25-46.

Butcher, Kristen F. and Anne Case. 1994. "The Effect of Sibling Sex Composition on Women's Education and Earnings." *Quarterly Journal of Economics*, Vol.109, No.3: 531-563.

Conley, Dalton and Rebecca Glauber. 2006. "Parental Educational Investment and Children's Academic Risk." *Journal of Human Resources*, Vol.41, No.4: 723-737

Das Gupta, Monica. 1987. "Selective Discrimination against Female Children in Rural Punjab, India." *Population and Development Review 13*, No.1: 77-100.

Dyson, Tim and Mick Moore. 1983. "On Kinship Structure, Female Autonomy, and Demographic Behavior in India." *Population and Development Review*, Vol.9, No.1: 35-60.

Ejrnæs, Mette and Calus C. Pörtner. 2004. "Birth Order and the Intrahousehold Allocation of Time and Education." *Review of Economics and Statistics*, Vol.86, No.4: 1008-1019.

Foster, Andrew D. and Mark R. Rosenzweig. 1996. "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution." *American Economic Review*, Vol.86, No.4: 931-953.

Garg, Asish and Jonathan Morduch. 1998. "Sibling Rivalry and the Gender Gap: Evidence form Child Health Outcomes in Ghana." *Journal of Population Economics*, Vol.11: 471-493.

Hanushek, Eric A. 1992. “The Trade-Off between Child Quantity and Quality.” *Journal of Political Economy*, Vol.100, No.1: 84-117.

Hauser, Robert M. and Hsiang-Hui Daphne Kuo. 1998. “Does the Gender Composition of Sibships Affect Women’s Educational Attainment?” *Journal of Human Resources*, Vol.33, No.3: 644-657.

Iacovou, Maria. 2007. “Family Size, Birth Order, and Educational Attainment.” *Marriage & Family Review*, Vol.42, Issue 3: 35-57.

Jensen, Robert. 2005. “Equal Treatment, Unequal Outcomes? Generating Gender Inequality through Fertility Behavior.” Mimeo. Harvard University, J. F. Kennedy School of Government.

Kaestner, Robert. 1997. “Are Brothers Really Better? Sibling Sex Composition and Educational Achievement Revisited.” *Journal of Human Resources*, Vol.32, No.2: 250-284.

Kantarevic, Jasmin and Stephane Mechoulan. 2006. “Birth Order, Educational Attainment, and Earnings: An Investigation Using the PSID.” *Journal of Human Resources*, Vol.41, No.4: 757-777.

Makino, Momoe. 2011. “Dowries, Female Bargaining, and Children’s Health.” Mimeo. University of Washington.

Morduch, Jonathan. 2000. “Sibling Rivalry in Africa.” *American Economic Review Papers and Proceedings*, Vol.90, No.2: 405-409.

Muhuri, Pradip K. and Samuel H. Preston. 1991. “Effects of Family Composition on Mortality Differentials by Sex Among Children in Matlab, Bangladesh.” *Population and Development Review*, Vol.17, No.3: 415-434.

Ozer-Balli, Hatice and Bent E. Sørensen. 2011. “Interaction Effects in Econometrics.” Mimeo. University of Houston.

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

Parish, William L. and Robert J. Willis. 1993. “Daughters, Education, and Family Budgets.” *Journal of Human Resources*, Vo.28, No.4: 863-898.

Pörtner, Claus C. 2010. “Sex Selective Abortions, Fertility and Birth Spacing.” Mimeo. University of Washington.

Rahman, Lupin and Vijayendra Rao. 2004. "The Determinants of Gender Equity in India: Examining Dyson and Moore's Thesis with New Data." *Population and Development Review*. Vol.30, Issue 2: 239-268.

Rao, Vijayendtra. 1993. "The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India." *Journal of Political Economy*. Vol.101, No.4: 666-677.

Rosenblum, Daniel. 2010. "The Effect of Fertility Decisions on Excess Female Mortality in India." Mimeo. Dalhousie University.

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

Sen, Amartya. 1990. "More Than 100 Million Women are Missing." *New York Review of Books*, Vol.37, No.20.

Zhang, Junsen and William Chan. 1999. "Dowry and Wife's Welfare: A Theoretical and Empirical Analysis." *Journal of Political Economy*, Vol.107, No.4: 786-808.

Table 1
Summary Statistics of Test Scores for Children Aged 8-11

Test Variables	mean	mean (boys only)	mean (girls only)	t-test (girl-boy)
Reading	0.648	0.656	0.638	-2.494
	(0.334)	(0.329)	(0.340)	
Number of observations	8227	4352	3875	
Math	0.516	0.536	0.495	-5.409
	(0.342)	(0.338)	(0.345)	
Number of observations	8190	4331	3859	
Writing	0.686	0.696	0.676	-1.906
	(0.464)	(0.460)	(0.468)	
Number of observations	8158	4305	3853	

Standard deviations are in parentheses.
Test scores are normalized to 1 = the highest score.

Table 2
Summary Statistics (for children aged 8-11 who have taken three tests)

Variables	mean	number of observation	mean (boys only)	mean (girls only)	t-test (girl-boy)
Children's attributes:					
Number of children	3.705	8123	3.573	3.852	7.770
	(1.626)		(1.620)	(1.619)	
Child's age	9.468	8123	9.459	9.479	0.833
	(1.059)		(1.060)	(1.057)	
Child's sex (boy =1)	0.528	8123			
	(0.499)				
Socioeconomic characteristics:					
Mother's school years	3.432	8097	3.432	3.432	-0.001
	(4.376)		(4.417)	(4.330)	
Mother's literacy	0.458	8118	0.454	0.462	0.737
	(0.498)		(0.498)	(0.499)	
Father's school years	5.830	8099	5.741	5.930	1.768
	(4.812)		(4.853)	(4.765)	
Mother's age	34.078	8120	34.144	34.004	-1.208
	(5.192)		(5.243)	(5.135)	
Mother's age at marriage	16.885	8107	16.923	16.843	-1.013
	(3.551)		(3.585)	(3.513)	
Age difference (Father-Mother)	5.188	8118	5.155	5.224	0.932
	(3.352)		(3.290)	(3.421)	
Urban (yes =1)	0.333	8123	0.326	0.341	1.499
	(0.471)		(0.469)	(0.474)	
Village: Endogamous marriage (yes =1)	0.269	8123	0.266	0.272	1.154
	(0.239)		(0.237)	(0.241)	
Caste variables:					
Same caste (yes =1)	0.941	8123	0.944	0.938	-1.010
	(0.235)		(0.230)	(0.240)	
Brahmin (yes =1)	0.051	8123	0.048	0.055	1.563
	(0.220)		(0.213)	(0.229)	
Highcaste (yes =1)	0.139	8123	0.144	0.134	-1.334
	(0.346)		(0.351)	(0.340)	
Scheduled caste (yes =1)	0.234	8123	0.227	0.242	1.602
	(0.423)		(0.419)	(0.428)	
Scheduled tribe (yes =1)	0.076	8123	0.078	0.072	-1.011
	(0.264)		(0.269)	(0.259)	
Other backward caste (yes =1)	0.400	8123	0.402	0.398	-0.416
	(0.490)		(0.490)	(0.489)	
Religion variables:	0.786	8123			
Hindu (yes =1)	(0.410)		0.784	0.789	0.464
	0.145	8123	(0.411)	(0.408)	
Muslim (yes =1)	(0.352)		0.143	0.147	0.534
	0.024	8123	(0.350)	(0.354)	
Christian (yes =1)	(0.152)		0.023	0.024	0.261
	0.026	8123	(0.150)	(0.153)	
Sikh (yes =1)	(0.159)		0.030	0.022	-2.111
	0.007	8123	(0.170)	(0.147)	
Buddhist (yes =1)	(0.085)		0.007	0.008	0.296
	0.012	8123	(0.083)	(0.087)	
Other religions except for Hindu (yes =1)	(0.110)		0.013	0.011	-0.965
	(0.112)		(0.115)	(0.104)	

Standard deviations are in parentheses.

Table 3

First Stage Regression

Variables	Number of Children	Number of Children	Number of Sons	Number of Sons	Number of Daughters	Number of Daughters
Instruments						
Village Family Planning Available (Oral pills)	-0.0146		-0.0615		0.0388	
	(0.0947)		(0.0712)		(0.0625)	
Village Family Planning Available (IUD insertion)	-0.0124		0.0118		-0.0109	
	(0.0978)		(0.0694)		(0.0665)	
Village Family Planning Available (Male sterilization)	-0.160**		-0.0724		-0.0735	
	(0.0652)		(0.0555)		(0.0559)	
Village Family Planning Available (Female sterilization)	0.0256		0.0331		-0.0146	
	(0.0736)		(0.0583)		(0.0630)	
Village Family Planning Available (Injection)	-0.00577		-0.0127		0.00129	
	(0.0536)		(0.0363)		(0.0451)	
Village Family Planning Available (Index 0-5)		-0.0339***		-0.0198***		-0.0121
		(0.00994)		(0.00752)		(0.00791)
Number of father's sister	0.0141	0.0143	0.00775	0.00753	0.00506	0.00543
	(0.0115)	(0.0115)	(0.00719)	(0.00723)	(0.00795)	(0.00794)
Number of father's brother	0.00384	0.00375	0.00306	0.00309	-0.000295	-0.000423
	(0.00951)	(0.00954)	(0.00655)	(0.00655)	(0.00726)	(0.00727)
First born girl (=1)	0.329***	0.327***	-0.811***	-0.812***	1.149***	1.148***
	(0.0326)	(0.0325)	(0.0257)	(0.0257)	(0.0301)	(0.0300)
Observations	7,875	7,875	7,747	7,747	7,755	7,755
R-squared	0.431	0.431	0.324	0.324	0.354	0.353
F-statistics (test of weak instruments)	14.6	27.7	130.2	251.9	185.9	368.1

Cluster (village)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%).

Other socio-economic variables, caste variables, religion variables, state dummies, and a constant term are included in the RHS as in Table 5.

Table 4

Birth Order Effects on Children's (age 8-11) Test Scores (OLS)

	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only
Variables	Reading	Reading	Reading	Math	Math	Math	Writing	Writing	Writing
Socio-economic variables:									
Number of children	-0.0238***	-0.0225***	-0.0238***	-0.0255***	-0.0255***	-0.0245***	-0.0245***	-0.0280***	-0.0202***
	(0.00338)	(0.00469)	(0.00447)	(0.00333)	(0.00460)	(0.00434)	(0.00493)	(0.00670)	(0.00651)
Relative birth order (b-1)/(n-1)	0.0225**	0.00241	0.0469***	-0.00289	-0.00914	0.00474	0.0136	-0.000943	0.0322
	(0.0104)	(0.0139)	(0.0158)	(0.0105)	(0.0146)	(0.0161)	(0.0153)	(0.0203)	(0.0234)
Child's age =9	0.104***	0.0881***	0.123***	0.0972***	0.0912***	0.106***	0.118***	0.0877***	0.154***
	(0.0103)	(0.0139)	(0.0144)	(0.00931)	(0.0130)	(0.0129)	(0.0146)	(0.0201)	(0.0209)
Child's age =10	0.171***	0.171***	0.170***	0.161***	0.165***	0.157***	0.171***	0.156***	0.187***
	(0.00928)	(0.0129)	(0.0131)	(0.00842)	(0.0119)	(0.0122)	(0.0134)	(0.0190)	(0.0192)
Child's age =11	0.226***	0.234***	0.220***	0.219***	0.236***	0.203***	0.210***	0.185***	0.241***
	(0.0111)	(0.0144)	(0.0162)	(0.0105)	(0.0144)	(0.0154)	(0.0159)	(0.0214)	(0.0221)
Child's sex (boy =1)	0.0154**			0.0370***			0.0210**		
	(0.00698)			(0.00710)			(0.00995)		
Mother's school years	0.00563***	0.00471**	0.00686***	0.00769***	0.00604***	0.00915***	0.00636***	0.00748**	0.00558*
	(0.00169)	(0.00224)	(0.00235)	(0.00176)	(0.00230)	(0.00256)	(0.00227)	(0.00292)	(0.00327)
Mother's literacy	0.0541***	0.0435**	0.0682***	0.0385***	0.0282	0.0548***	0.0632***	0.0388	0.0887***
	(0.0144)	(0.0189)	(0.0203)	(0.0145)	(0.0190)	(0.0207)	(0.0194)	(0.0251)	(0.0280)
Father's school years	0.00987***	0.0114***	0.00793***	0.0104***	0.0122***	0.00832***	0.0111***	0.0127***	0.00910***
	(0.00101)	(0.00136)	(0.00143)	(0.00102)	(0.00137)	(0.00138)	(0.00152)	(0.00198)	(0.00211)
Mother's age	-0.00112	-0.00395	0.00265	0.00809	0.00440	0.0128	-0.00580	-0.00364	-0.00434
	(0.00794)	(0.0107)	(0.0116)	(0.00805)	(0.0106)	(0.0115)	(0.0118)	(0.0158)	(0.0177)
Mother's age^2	2.92e-05	6.10e-05	-1.36e-05	-5.51e-05	-2.06e-05	-0.000101	0.000110	9.48e-05	7.47e-05
	(0.000110)	(0.000147)	(0.000161)	(0.000111)	(0.000146)	(0.000160)	(0.000162)	(0.000216)	(0.000245)
Mother's age at marriage	0.00320**	0.00229	0.00389**	0.00236*	0.00140	0.00332*	0.00303	0.00122	0.00493*
	(0.00140)	(0.00188)	(0.00196)	(0.00139)	(0.00181)	(0.00197)	(0.00206)	(0.00265)	(0.00292)
Age difference (Father-Mother)	0.00318***	0.00315*	0.00301**	0.00510***	0.00305*	0.00698***	0.00393**	0.00496**	0.00281
	(0.00116)	(0.00168)	(0.00153)	(0.00119)	(0.00172)	(0.00157)	(0.00175)	(0.00246)	(0.00237)
Urban (yes =1)	0.0439***	0.0250**	0.0613***	0.0578***	0.0399***	0.0756***	0.0366***	0.0190	0.0536***
	(0.00949)	(0.0124)	(0.0125)	(0.00974)	(0.0130)	(0.0131)	(0.0137)	(0.0175)	(0.0181)

Village: Endogamous marriage	-0.0379*	-0.0344	-0.0419	-0.0286	-0.0232	-0.0338	-0.0467	-0.0692*	-0.0230
	(0.0199)	(0.0250)	(0.0264)	(0.0199)	(0.0257)	(0.0254)	(0.0286)	(0.0371)	(0.0363)
Caste variables:									
Same caste (yes =1)	0.0328**	0.0568**	0.0169	0.0163	0.0302	0.0142	0.0239	0.0441	0.0102
	(0.0166)	(0.0241)	(0.0220)	(0.0164)	(0.0224)	(0.0225)	(0.0242)	(0.0312)	(0.0324)
Brahmin (yes =1)	0.0687***	0.0698**	0.0598*	0.0756***	0.0767**	0.0684**	0.0322	0.0315	0.0223
	(0.0248)	(0.0330)	(0.0330)	(0.0259)	(0.0344)	(0.0344)	(0.0363)	(0.0473)	(0.0477)
Highcaste (yes =1)	0.0416*	0.0483	0.0336	0.0334	0.0301	0.0365	0.0511	0.0815*	0.0137
	(0.0227)	(0.0299)	(0.0306)	(0.0232)	(0.0303)	(0.0311)	(0.0327)	(0.0421)	(0.0430)
Scheduled caste (yes =1)	-0.00716	-0.0114	-0.00570	-0.0315	-0.0358	-0.0301	-0.00424	0.0157	-0.0338
	(0.0221)	(0.0287)	(0.0297)	(0.0224)	(0.0287)	(0.0295)	(0.0321)	(0.0413)	(0.0413)
Scheduled tribe (yes =1)	-0.00286	0.0184	-0.0331	-0.0449*	-0.0156	-0.0838**	-0.0310	0.0400	-0.120**
	(0.0258)	(0.0327)	(0.0367)	(0.0254)	(0.0323)	(0.0343)	(0.0372)	(0.0477)	(0.0506)
Other backward caste (yes =1)	0.0186	0.0303	0.00484	0.00665	0.00926	0.00166	0.00862	0.0451	-0.0367
	(0.0204)	(0.0262)	(0.0279)	(0.0208)	(0.0266)	(0.0271)	(0.0295)	(0.0379)	(0.0379)
Religion variables:									
Muslim (yes =1)	-0.00772	-0.0152	-0.00147	-0.0442**	-0.0510**	-0.0382*	-0.0391	-0.0260	-0.0535
	(0.0179)	(0.0236)	(0.0240)	(0.0182)	(0.0242)	(0.0226)	(0.0261)	(0.0330)	(0.0352)
Christian (yes =1)	-0.00454	0.00693	-0.0210	-0.00118	0.0121	-0.0215	-0.0493	-0.0149	-0.105*
	(0.0299)	(0.0452)	(0.0359)	(0.0296)	(0.0433)	(0.0391)	(0.0368)	(0.0513)	(0.0543)
Sikh (yes =1)	-0.00697	0.0292	-0.0531	0.00429	0.0414	-0.0463	0.00350	0.0407	-0.0473
	(0.0298)	(0.0365)	(0.0452)	(0.0272)	(0.0372)	(0.0386)	(0.0429)	(0.0551)	(0.0685)
Buddhist (yes =1)	0.0294	0.104	-0.0427	0.00440	0.0194	-0.00706	-0.0629	0.0420	-0.140
	(0.0453)	(0.0699)	(0.0613)	(0.0379)	(0.0611)	(0.0493)	(0.0626)	(0.0954)	(0.0966)
Other religions except for Hindu (yes =1)	-0.0325	-0.0593	0.00732	0.00242	-0.00917	0.0286	0.0189	0.00603	0.0541
	(0.0412)	(0.0520)	(0.0537)	(0.0404)	(0.0431)	(0.0586)	(0.0618)	(0.0675)	(0.0864)
Constant	0.455***	0.298	0.298	0.325**	0.235	0.151	0.667***	0.645**	0.691**
	(0.143)	(0.186)	(0.190)	(0.138)	(0.208)	(0.189)	(0.209)	(0.283)	(0.317)
Observations	7,670	4,029	3,641	7,634	4,009	3,625	7,606	3,985	3,621

Cluster (village)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). State dummies and a constant term are included in the RHS.

Table 5

Birth Order Effects on Children's (age 8-11) Test Scores (2SLS)

	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only
Variables	Reading	Reading	Reading	Math	Math	Math	Writing	Writing	Writing
Socio-economic variables:									
Number of children	0.0212	0.0493*	0.0318	0.00176	0.00654	0.0410	0.0474	0.0232	0.0859
	(0.0272)	(0.0300)	(0.0425)	(0.0262)	(0.0280)	(0.0431)	(0.0383)	(0.0411)	(0.0633)
Relative birth order (b-1)/(n-1)	0.0424***	0.0304	0.0731***	0.00869	0.00192	0.0351	0.0453*	0.0199	0.0811**
	(0.0160)	(0.0187)	(0.0252)	(0.0160)	(0.0189)	(0.0259)	(0.0231)	(0.0258)	(0.0384)
Child's age =9	0.108***	0.0935***	0.131***	0.0991***	0.0928***	0.114***	0.124***	0.0879***	0.169***
	(0.0110)	(0.0147)	(0.0159)	(0.00992)	(0.0133)	(0.0144)	(0.0156)	(0.0207)	(0.0233)
Child's age =10	0.179***	0.182***	0.182***	0.167***	0.170***	0.172***	0.184***	0.163***	0.209***
	(0.0109)	(0.0146)	(0.0156)	(0.00994)	(0.0129)	(0.0153)	(0.0156)	(0.0207)	(0.0233)
Child's age =11	0.235***	0.250***	0.233***	0.225***	0.243***	0.218***	0.227***	0.197***	0.265***
	(0.0130)	(0.0169)	(0.0187)	(0.0125)	(0.0163)	(0.0182)	(0.0187)	(0.0237)	(0.0273)
Child's sex (boy =1)	0.0265***			0.0442***			0.0398***		
	(0.00993)			(0.00976)			(0.0144)		
Mother's school years	0.00859***	0.00923***	0.0107***	0.00985***	0.00855***	0.0142***	0.0113***	0.0106***	0.0138**
	(0.00253)	(0.00300)	(0.00409)	(0.00244)	(0.00282)	(0.00429)	(0.00342)	(0.00378)	(0.00602)
Mother's literacy	0.0459***	0.0316	0.0567**	0.0310**	0.0216	0.0359	0.0460**	0.0270	0.0608*
	(0.0158)	(0.0204)	(0.0239)	(0.0156)	(0.0196)	(0.0252)	(0.0213)	(0.0265)	(0.0340)
Father's school years	0.0106***	0.0128***	0.00860***	0.0108***	0.0127***	0.00907***	0.0121***	0.0136***	0.00986***
	(0.00111)	(0.00156)	(0.00150)	(0.00112)	(0.00152)	(0.00146)	(0.00166)	(0.00224)	(0.00223)
Mother's age	-0.00699	-0.0142	-0.00485	0.00474	0.000676	0.000954	-0.0174	-0.0124	-0.0218
	(0.00965)	(0.0122)	(0.0151)	(0.00962)	(0.0117)	(0.0154)	(0.0142)	(0.0173)	(0.0233)
Mother's age^2	2.03e-05	5.67e-05	-1.74e-05	-6.42e-05	-3.37e-05	-6.07e-05	0.000125	0.000108	0.000117
	(0.000113)	(0.000154)	(0.000172)	(0.000113)	(0.000148)	(0.000174)	(0.000169)	(0.000219)	(0.000269)
Mother's age at marriage	0.00761**	0.00889***	0.00974*	0.00489	0.00424	0.0101**	0.0103**	0.00645	0.0159**
	(0.00311)	(0.00341)	(0.00500)	(0.00304)	(0.00324)	(0.00512)	(0.00439)	(0.00467)	(0.00746)
Age difference (Father-Mother)	0.00291**	0.00266	0.00254	0.00477***	0.00260	0.00644***	0.00358**	0.00514**	0.00181
	(0.00119)	(0.00179)	(0.00157)	(0.00120)	(0.00173)	(0.00162)	(0.00179)	(0.00252)	(0.00250)
Urban (yes =1)	0.0459***	0.0253*	0.0680***	0.0580***	0.0393***	0.0822***	0.0441***	0.0238	0.0680***
	(0.0101)	(0.0131)	(0.0142)	(0.0101)	(0.0131)	(0.0149)	(0.0147)	(0.0180)	(0.0208)

Village: Endogamous marriage	-0.0473**	-0.0439*	-0.0561**	-0.0314	-0.0256	-0.0467*	-0.0608**	-0.0785**	-0.0482
	(0.0205)	(0.0264)	(0.0284)	(0.0204)	(0.0262)	(0.0280)	(0.0302)	(0.0383)	(0.0414)
Caste variables:	0.0308*	0.0618**	0.00689	0.0154	0.0309	0.00484	0.0255	0.0485	0.00336
Same caste (yes =1)	0.0308*	0.0618**	0.00689	0.0154	0.0309	0.00484	0.0255	0.0485	0.00336
	(0.0169)	(0.0248)	(0.0230)	(0.0166)	(0.0231)	(0.0232)	(0.0260)	(0.0321)	(0.0374)
Brahmin (yes =1)	0.0648**	0.0730**	0.0470	0.0731***	0.0736**	0.0609*	0.0290	0.0311	0.00879
	(0.0256)	(0.0360)	(0.0336)	(0.0267)	(0.0353)	(0.0363)	(0.0377)	(0.0484)	(0.0502)
Highcaste (yes =1)	0.0457**	0.0609*	0.0281	0.0336	0.0309	0.0309	0.0522	0.0871**	0.00173
	(0.0233)	(0.0323)	(0.0312)	(0.0236)	(0.0311)	(0.0321)	(0.0339)	(0.0433)	(0.0459)
Scheduled caste (yes =1)	-0.0181	-0.0212	-0.0294	-0.0374	-0.0404	-0.0545	-0.0250	0.00313	-0.0746
	(0.0238)	(0.0312)	(0.0339)	(0.0240)	(0.0298)	(0.0347)	(0.0345)	(0.0427)	(0.0500)
Scheduled tribe (yes =1)	-0.0173	0.00643	-0.0623	-0.0541**	-0.0218	-0.116***	-0.0558	0.0250	-0.167***
	(0.0279)	(0.0358)	(0.0413)	(0.0270)	(0.0334)	(0.0402)	(0.0395)	(0.0489)	(0.0596)
Other backward caste (yes =1)	0.0155	0.0375	-0.0120	0.00497	0.0134	-0.0152	0.00212	0.0490	-0.0638
	(0.0209)	(0.0281)	(0.0292)	(0.0212)	(0.0272)	(0.0290)	(0.0305)	(0.0386)	(0.0425)
Religion variables:									
Muslim (yes =1)	-0.0495*	-0.0808**	-0.0540	-0.0666**	-0.0795**	-0.0949**	-0.100**	-0.0690	-0.148**
	(0.0296)	(0.0351)	(0.0448)	(0.0296)	(0.0347)	(0.0457)	(0.0421)	(0.0485)	(0.0684)
Christian (yes =1)	-0.0182	-0.0183	-0.0404	-0.0103	-1.55e-05	-0.0423	-0.0735*	-0.0285	-0.136**
	(0.0329)	(0.0488)	(0.0386)	(0.0318)	(0.0470)	(0.0420)	(0.0413)	(0.0563)	(0.0593)
Sikh (yes =1)	-0.00478	0.0407	-0.0593	0.00582	0.0460	-0.0506	0.00890	0.0498	-0.0528
	(0.0303)	(0.0394)	(0.0454)	(0.0277)	(0.0382)	(0.0417)	(0.0446)	(0.0576)	(0.0728)
Buddhist (yes =1)	0.0334	0.124*	-0.0458	0.00528	0.0257	-0.0101	-0.0564	0.0592	-0.145
	(0.0476)	(0.0696)	(0.0694)	(0.0387)	(0.0626)	(0.0546)	(0.0655)	(0.0954)	(0.109)
Other religions except for Hindu (yes =1)	-0.0442	-0.0765	-0.0125	0.00233	-0.0140	0.0242	0.00439	-0.0109	0.0428
	(0.0452)	(0.0642)	(0.0528)	(0.0418)	(0.0461)	(0.0578)	(0.0638)	(0.0724)	(0.0830)
Constant	0.544***	0.457**	0.503**	0.316**	0.381*	0.241	0.841***	0.810***	0.849**
	(0.162)	(0.197)	(0.228)	(0.161)	(0.213)	(0.226)	(0.220)	(0.292)	(0.333)
Observations	7,540	3,963	3,577	7,504	3,943	3,561	7,477	3,919	3,558

Cluster (village)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). State dummies and a constant term are included in the RHS.

Table 6

Sibling Sex Composition Effects on Children's Test Scores (2SLS with two endogenous variables = number of children, number of older/younger brothers/sisters index)

	Both Sexes	**Boys Only**	**Girls Only**	**Both Sexes**	**Boys Only**	**Girls Only**	**Both Sexes**	**Boys Only**	**Girls Only**
Variables	Reading	Reading	Reading	Math	Math	Math	Writing	Writing	Writing
Number of sisters index: number of sisters/(n-1)	0.000216	0.0854	-0.0698	0.0470	0.0979*	-0.00870	0.0208	0.120	-0.00539
	(0.0434)	(0.0532)	(0.0605)	(0.0431)	(0.0567)	(0.0602)	(0.0646)	(0.0790)	(0.0894)
Number of brothers index: number of brothers/(n-1)	4.06e-05	-0.0811	0.0685	-0.0515	-0.101*	0.0101	-0.0211	-0.126	0.00931
	(0.0444)	(0.0543)	(0.0613)	(0.0441)	(0.0579)	(0.0612)	(0.0660)	(0.0811)	(0.0909)
Number of older sisters index: number of older sisters/(n-1)	0.0122	0.0822*	-0.105	0.0543	0.0936**	-0.00945	0.0312	0.105	-0.0599
	(0.0451)	(0.0439)	(0.0924)	(0.0453)	(0.0477)	(0.0910)	(0.0673)	(0.0651)	(0.137)
Number of younger sisters index: number of younger sisters/(n-1)	-0.0499	-0.292	-0.108	0.176	-0.223	0.00213	-0.0632	-0.312	0.0628
	(0.196)	(0.187)	(0.135)	(0.199)	(0.191)	(0.135)	(0.286)	(0.275)	(0.201)
Number of older brothers: number of older brothers/(n-1)	-0.00576	-0.104	0.0380	-0.0461	-0.127*	0.00295	-0.0166	-0.149	0.0124
	(0.0380)	(0.0682)	(0.0381)	(0.0378)	(0.0717)	(0.0381)	(0.0564)	(0.102)	(0.0566)
Number of younger brothers: number of younger brothers/(n-1)	0.0376	-0.248	-0.129	0.177	-0.342	-0.0380	0.0384	-0.363	-0.0743
	(0.297)	(0.187)	(0.118)	(0.305)	(0.209)	(0.118)	(0.454)	(0.276)	(0.173)

Cluster (village)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%).

The table extracts only the estimated coefficients of the measure of sibling sex composition in each of the six-by-nine estimation models. Relative birth order in Table 5 is replaced by the index of sibling sex composition. Other included RHS variables are the same as in Table 5.

Table 7

Effects of 1st Born Girl on Test Scores of Children of Birth Order 2 or Higher

	Both Sexes	**Boys Only**	**Girls Only**	**Both Sexes**	**Boys Only**	**Girls Only**	**Both Sexes**	**Boys Only**	**Girls Only**
Variables	Reading	Reading	Reading	Math	Math	Math	Writing	Writing	Writing
OLS estimation: replication of Rosenblum (2010)									
1st born girl	0.00403	0.0250**	-0.0195*	0.00283	0.0153	-0.0100	0.00946	0.0222	-0.00541
	(0.00848)	(0.0112)	(0.0115)	(0.00820)	(0.0109)	(0.0119)	(0.0125)	(0.0163)	(0.0178)
2SLS estimation: endogenous variable = number of children									
Number of children	-0.00631	-0.0548	0.0208	-0.0241	-0.101	0.0526	0.0549	-0.0190	0.0853
	(0.0527)	(0.0773)	(0.0558)	(0.0501)	(0.0806)	(0.0574)	(0.0797)	(0.114)	(0.0835)
1st born girl	0.00607	0.0360**	-0.0314	0.0102	0.0334*	-0.0383	-0.00871	0.0260	-0.0495
	(0.0189)	(0.0173)	(0.0309)	(0.0179)	(0.0182)	(0.0315)	(0.0300)	(0.0265)	(0.0470)

Cluster (village)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). In the 2SLS estimation model, the instruments for the number of children are the availability of family planning services and the number of father's siblings.

The table extracts only the estimated coefficients of the 1st born girl and the number of children.

Table 8
Children's Test Scores Controlling for HH Heterogeneity

Variables	Reading	Reading	Reading	Math	Math	Math	Writing	Writing	Writing
Socio-economic variables:									
Child's age=9	0.0645***	0.0828***	0.0740***	0.0551***	0.0611***	0.0512***	0.0620**	0.0877***	0.0819***
	(0.0193)	(0.0173)	(0.0180)	(0.0200)	(0.0180)	(0.0188)	(0.0305)	(0.0288)	(0.0298)
Child's age=10	0.144***	0.197***	0.172***	0.140***	0.161***	0.134***	0.150***	0.221***	0.204***
	(0.0305)	(0.0181)	(0.0243)	(0.0299)	(0.0180)	(0.0238)	(0.0472)	(0.0266)	(0.0356)
Child's age=11	0.148***	0.222***	0.187***	0.165***	0.191***	0.153***	0.129**	0.228***	0.204***
	(0.0412)	(0.0241)	(0.0326)	(0.0407)	(0.0238)	(0.0317)	(0.0653)	(0.0375)	(0.0499)
Child's sex (boy =1)	0.00683	0.00558	0.00521	0.0497***	0.0457***	0.0463***	0.0355*	0.0357*	0.0347*
	(0.0123)	(0.0125)	(0.0125)	(0.0126)	(0.0127)	(0.0127)	(0.0185)	(0.0185)	(0.0186)
Absolute birth order (b)	-0.0656**			-0.0559**			-0.0642		
	(0.0279)			(0.0273)			(0.0443)		
Relative birth order (b-1)/(n-1)		-0.0265			-0.0860***			0.0226	
		(0.0313)			(0.0314)			(0.0454)	
Birth order index (b/(n+1)/2)			-0.0868*			-0.152***			-0.0197
			(0.0495)			(0.0484)			(0.0732)
Constant	0.723***	0.531***	0.629***	0.544***	0.432***	0.569***	0.744***	0.516***	0.565***
	(0.0913)	(0.0289)	(0.0680)	(0.0895)	(0.0292)	(0.0669)	(0.145)	(0.0425)	(0.101)
Observations	2713	2681	2683	2711	2679	2682	2696	2664	2667

Cluster (household)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). The model includes the household fixed effects.

Table 9
Children's Test Scores Controlling for HH Heterogeneity and Including Interaction Terms

Variables	Reading	Reading	Math	Math	Writing	Writing
Socio-economic variables:						
Child's age=9	0.0834***	0.0745***	0.0629***	0.0527***	0.0889***	0.0827***
	(0.0173)	(0.0180)	(0.0179)	(0.0187)	(0.0288)	(0.0299)
Child's age=10	0.198***	0.173***	0.165***	0.137***	0.224***	0.206***
	(0.0181)	(0.0243)	(0.0180)	(0.0237)	(0.0266)	(0.0357)
Child's age=11	0.224***	0.189***	0.198***	0.158***	0.232***	0.207***
	(0.0241)	(0.0325)	(0.0238)	(0.0317)	(0.0377)	(0.0502)
Child's sex (boy =1)	0.0292	0.0458	0.106***	0.155***	0.0758**	0.0960
	(0.0242)	(0.0385)	(0.0249)	(0.0385)	(0.0361)	(0.0588)
Relative birth order (b-1)/(n-1)	0.00231		-0.0125		0.0714	
	(0.0378)		(0.0394)		(0.0567)	
Int(relative birth order*sex)	-0.0432		-0.110***		-0.0733	
	(0.0358)		(0.0382)		(0.0561)	
Birth order index (b/(n+1)/2)		-0.0623		-0.0862		0.0174
		(0.0520)		(0.0525)		(0.0808)
Int(birth order index*sex)		-0.0382		-0.102***		-0.0577
		(0.0331)		(0.0339)		(0.0524)
Constant	0.515***	0.603***	0.391***	0.499***	0.488***	0.526***
	(0.0313)	(0.0700)	(0.0322)	(0.0703)	(0.0468)	(0.108)
Observations	2681	2683	2679	2682	2664	2667

Cluster (household)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). The model includes the household fixed effects.

The table does not include the specification with absolute birth order as a measure of birth order. The interaction term between the absolute birth order and sex of children is dropped due to the multicolinearity.

Table 10
By-gender Birth Order Effect Controlling for HH Heterogeneity

Variables	Reading	Reading	Math	Math	Writing	Writing
Socio-economic variables:						
Child's age=9	0.0867***	0.0911***	0.0757***	0.0711***	0.0821***	0.0879**
	(0.0168)	(0.0204)	(0.0173)	(0.0205)	(0.0273)	(0.0352)
Child's age=10	0.202***	0.210***	0.194***	0.191***	0.201***	0.230***
	(0.0126)	(0.0174)	(0.0125)	(0.0176)	(0.0194)	(0.0272)
Child's age=11	0.231***	0.258***	0.241***	0.234***	0.201***	0.251***
	(0.0174)	(0.0240)	(0.0172)	(0.0233)	(0.0262)	(0.0389)
Child's sex (boy =1)	0.00947	0.0521*	0.103***	0.0924***	0.0743*	0.0805*
	(0.0257)	(0.0272)	(0.0260)	(0.0268)	(0.0394)	(0.0429)
Absolute by-gender birth order (b)	-0.00992		0.0181*		-0.0104	
	(0.0101)		(0.0104)		(0.0163)	
Int(absolute by-gender birth order*sex)	-0.00299		-0.0327**		-0.0270	
	(0.0137)		(0.0137)		(0.0217)	
Relative by-gender birth order (b-1)/(n-1)		0.0404		0.0303		0.0674
		(0.0276)		(0.0302)		(0.0433)
Int(relative by-gender birth order*sex)		-0.0359		-0.0546*		-0.0658
		(0.0279)		(0.0296)		(0.0455)
Constant	0.533***	0.440***	0.332***	0.314***	0.566***	0.450***
	(0.0238)	(0.0284)	(0.0237)	(0.0290)	(0.0373)	(0.0437)
Observations	2720	2196	2718	2195	2703	2180

Cluster (household)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). The model includes the household fixed effects.

The table does not include the specification with by-gender birth order index as a measure of birth order. The interaction term between the by-gender birth order index and sex of children is dropped due to the multicolinearity.

Table A.1
Distribution of Raw Test Scores for Children Aged 8-11

Test Variables		number of observation	number of observation (boys only)	number of observation (girls only)
Reading	total	8227	4352	3875
	cannot read=0	797	378	419
	letter=1	1106	578	528
	word=2	1644	886	758
	paragraph=3	1813	971	842
	story=4	2867	1539	1328
Math	total	8190	4331	3859
	cannot count=0	1461	677	784
	number=1	2614	1386	1228
	subtraction=2	2273	1231	1042
	division=3	1842	1037	805
Writing	total	8158	4305	3853
	cannot write=0	2558	1310	1248
	write w/ 2 or less mistakes=1	5600	2995	2605

Table A.2

Using Birth Order Index instead of Relative Birth Order (2SLS, compare with the fist two rows in Table 5)

	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only
Variables	Reading	Reading	Reading	Math	Math	Math	Writing	Writing	Writing
Number of children	0.0191	0.0369	0.0324	-0.00311	-0.00141	0.0409	0.0425	0.0148	0.0871
	(0.0268)	(0.0276)	(0.0462)	(0.0256)	(0.0257)	(0.0466)	(0.0383)	(0.0383)	(0.0690)
Birth order index (b/(n+1)/2)	0.0482***	0.0280*	0.0751***	0.00998	-0.00216	0.0364	0.0426**	0.0177	0.0784**
	(0.0142)	(0.0168)	(0.0246)	(0.0142)	(0.0170)	(0.0249)	(0.0213)	(0.0240)	(0.0381)

Cluster (village)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%).

The table extracts only the estimated coefficients of the number of children and the birth order index. The estimation includes the same RHS variables as in Table 5.

Table A.3

Including Index Variable (0-5) instead of Five Dummies of Family Planning Services in Instruments (2SLS, compare with the first two rows in Table 5)

	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only
Variables	Reading	Reading	Reading	Math	Math	Math	Writing	Writing	Writing
Number of children	0.0181	0.0424	0.0238	0.000352	0.00109	0.0419	0.0450	0.0257	0.105
	(0.0272)	(0.0283)	(0.0488)	(0.0259)	(0.0262)	(0.0495)	(0.0388)	(0.0392)	(0.0743)
Relative birth order (b-1)/(n-1)	0.0479***	0.0292*	0.0718***	0.0111	-0.00158	0.0367	0.0433**	0.0201	0.0852**
	(0.0143)	(0.0170)	(0.0251)	(0.0143)	(0.0171)	(0.0255)	(0.0213)	(0.0242)	(0.0398)

Cluster (village)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%).

The table extracts only the estimated coefficients of the number of children and the relative birth order. he estimation includes the same RHS variables as in Table 5.

Table A.4

Sensitivity of Birth Order Effects to a Different Set of Instruments (coefficients of relative birth order are extracted)

	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only	Both sexes	Boys only	Girls only
Instruments	Reading	Reading	Reading	Math	Math	Math	Writing	Writing	Writing
Complete 3 sets (coefficient estimates extracted from Table 5)	0.0424***	0.0304	0.0731***	0.00869	0.00192	0.0351	0.0453*	0.0199	0.0811**
	(0.0160)	(0.0187)	(0.0252)	(0.0160)	(0.0189)	(0.0259)	(0.0231)	(0.0258)	(0.0384)
Only family planning, sex of first born	0.0507***	0.0301*	0.0780***	0.0107	1.92e-05	0.0318	0.0473**	0.0194	0.0821**
	(0.0142)	(0.0168)	(0.0255)	(0.0142)	(0.0170)	(0.0251)	(0.0213)	(0.0241)	(0.0393)
Only sex of first born, number of father's siblings	0.0477***	0.0296*	0.0772**	0.0157	0.000610	0.0602*	0.0405*	0.0189	0.0988**
	(0.0146)	(0.0171)	(0.0303)	(0.0148)	(0.0173)	(0.0338)	(0.0217)	(0.0242)	(0.0490)
Only family planning, number of father's siblings	0.0396	0.00320	0.0680**	-0.00997	-0.0404	0.0252	0.0422	-0.0171	0.0595
	(0.0266)	(0.0358)	(0.0309)	(0.0253)	(0.0365)	(0.0313)	(0.0390)	(0.0497)	(0.0465)
Only sex of first born	0.0434***	0.0349*	0.0808***	0.0181	0.0107	0.0516	0.0466*	0.0248	0.103**
	(0.0167)	(0.0194)	(0.0312)	(0.0169)	(0.0195)	(0.0320)	(0.0246)	(0.0269)	(0.0483)

Cluster (village)-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%).

Table A.5

Birth Order and Sibling Sex Composition Effect on Children's Enrollment Status (coefficient estimates of birth order measure and children's sex are extracted)

Variables	**Both Sexes**	**Boys Only**	**Girls Only**
Compare with Table 5: (N = 9,450)			
Relative birth order	0.00963	0.00944	0.00958
	(0.00653)	(0.00778)	(0.0105)
Children's sex	0.00365		
	(0.00426)		
Compare with Table 6: (N = 9,566)			
Number of sisters index	-0.00728	0.0140	-0.0168
	(0.0194)	(0.0241)	(0.0282)
Children's sex	0.00734		
	(0.00781)		
Number of brothers index	0.00760	-0.0209	0.0163
	(0.0202)	(0.0257)	(0.0285)
Children's sex	0.00758		
	(0.00820)		
Compare with Table 7: (N = 7,256)			
Girl1	-0.00698	0.000720	-0.0110
	(0.00839)	(0.00795)	(0.0135)
Children's sex	0.00939		
	(0.00766)		
Compare with Table 8: (N = 4,028)			
Relative birth order	-0.0153		
	(0.0101)		
Children's sex	-0.00175		
	(0.00538)		
Compare with Table 9:			
Relative birth order	-0.00805		
	(0.0126)		
Int(relative bo*sex)	-0.0111		
	(0.0140)		
Children's sex	0.00428		
	(0.00974)		
Compare with Table 10:			
By-gender relative birth order	-0.00869		
	(0.0103)		
Int(by-gender relative bo*sex)	-0.0195*		
	(0.0114)		
Children's sex	0.00313		
	(0.0128)		

Cluster-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%).

The tables in which these coefficients should be compared are provided.

Table A.6

Birth Order and Sibling Sex Composition Effect on Children's Test-Taken Status (coefficient estimates of birth order measure and children's sex are extracted)

Variables	**Both Sexes**	**Boys Only**	**Girls Only**
Compare with Table 5: (N = 10,070)			
Relative birth order	0.0171	0.00386	0.0462
	(0.0217)	(0.0249)	(0.0345)
Children's sex	0.00865		
	(0.0120)		
Compare with Table 6: (N = 10,189)			
Number of sisters index	0.000581	0.0380	-0.0309
	(0.0515)	(0.0666)	(0.0683)
Children's sex	0.0114		
	(0.0207)		
Number of brothers index	0.00177	-0.0460	0.0402
	(0.0533)	(0.0706)	(0.0690)
Children's sex	0.0120		
	(0.0214)		
Compare with Table 7: (N = 7,754)			
Girl1	-0.00864	-0.0142	0.00334
	(0.0202)	(0.0183)	(0.0340)
Children's sex	0.0173		
	(0.0188)		
Compare with Table 8: (N = 4,359)			
Relative birth order	0.0397		
	(0.0369)		
Children's sex	0.00700		
	(0.0154)		
Compare with Table 9:			
Relative birth order	0.0314		
	(0.0469)		
Int(relative bo*sex)	0.0128		
	(0.0467)		
Children's sex	0.000141		
	(0.0290)		
Compare with Table 10:			
By-gender relative birth order	-0.0336		
	(0.0343)		
Int(by-gender relative bo*sex)	0.0274		
	(0.0371)		
Children's sex	0.0285		
	(0.0323)		

Cluster-robust standard errors are in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%).

The tables in which these coefficients should be compared are provided.