"Big Brother" Effect: Impact of Birth Order and Gender on Learning Outcomes - Evidence from India

Bhanu Gupta*

December 12, 2016

Click here for the latest version

Please don't cite without permission

Abstract

How do learning outcomes differ across gender and are they affected by the sibling composition? Is there a difference in trends according to the societal preference for sons? What are the possible mechanisms that drive these patterns and has the introduction of Right to Education (RTE) led to any change in gender-bias? This study tries to answer some of these questions by looking at data from Annual Survey of Education Report (ASER-2014). I find that (i) later born children have lower learning outcomes as compared to the first borns, (ii) In the less biased states, being the eldest son improves the learning outcomes for the male, while having no elder brother is detrimental for the female. The results seem to be driven by the differential education spending by the households across the children. Finally, introduction of RTE has not led to changes in the observed gender-bias.

*Graduate Student, Department of Economics, University of Michigan at Ann Arbor. I would like to thank Professor Dean Yang for his constant guidance and suggestions during the entire project. I would also like to thank Professor Ugo Troiano who helped get the data. Finally, I would like to thank all my colleagues in Development sequence for their valuable suggestions. Any remaining errors are completely mine.

"For women no (sacramental) rite (is performed) with sacred texts, thus the law is settled; women (who are) destitute of strength and destitute of (the knowledge of) Vedic texts, (are as impure as) falsehood (itself), that is a fixed rule" - Manusmriti Chapter 9 V19.

"The State shall not discriminate against any citizen on grounds only of religion, race, caste, sex, place of birth or any of them"

" Nothing in this article shall prevent the State from making any special provision for women and children"

Article 15 (1) & (3) - Constitution of India

1 Introduction

Going to school is not an end-goal in itself. It is what you learn there, that really counts. This subtle distinction is many a times forgotten by policy makers. For instance, United Nations Millennium Development Goals call for "children everywhere, boys and girls alike, will be able to complete a full course of primary schooling". They also espouse to "eliminate gender disparity in primary and secondary education". In many developing countries like China and India, there is a deep-rooted bias against education and upliftment of women. With the growth of income levels and arrival of modern technology, it might seem that some of this bias is fading. For instance, for India, in 2014 adjusted net primary enrollment rate for females stood at ninety-nine percent as compared to ninety-two percent for boys[1]. However, it is possible that girls are still dis-advantaged when it comes to actual learning and not just enrollment in schools. In this study I intend to probe the question of gender disparity in learning outcomes at the household level, and try to tease out the mechanisms that drive it. The context will be rural India. Specifically, I ask the following questions: 1) What is the impact of birth order on learning outcomes of the children? 2) Does the birth order gradient vary according to gender? 3) How does sibling composition affect the gender-bias? How do these patterns vary across states? and 4) What are the mechanisms driving the patterns observed in the data?

To answer these question, I look at the data from Annual Status of Education Report - 2014. It is a pan-India survey data for all the rural districts of India with 600 households surveyed in each district of the 31 states and union territories. Various child, paternal and household characterisicts that can impact learning outcomes, are recorded. The key variable of interest is learning outcomes which are assessed by testing the math and reading skills of

[1]Source : United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics

the child in his own native tongue. To establish causality, I use multi-variate Fixed-effects regression approach by introducing mother fixed-effects, which allows me to run "within" regression and hence, control for any time-invariant mother characteristics, that can affect the learning outcomes of the children. I also divide the sample into less and more biased states, based on the adult-sex ratio, to understand the sociological variations in the trends. The approach used in this study follows closely the one used by Jayachandran and Pande 2015 (henceforth referred to as JP) in solving the India -Africa height disadvantage puzzle.

The key findings of the paper as follows. There is a steep birth order gradient in learning outcomes, with the second and later borns disadvantaged as compared to the first borns. When we interact the gender dummy with birth order, the results vary depending on the gender-bias prevalent in the society. In the more biased states, the birth order gradient is less steep for girls as compared to the birth order gradient for boys. In the less biased states, the birth order gradient doesn't vary across gender. Moreover, the gender composition of the siblings seem to be an important factor influencing the variation in the learning outcomes. In the more biased states, the boy have an advantage in learning if they happen to be the eldest born son. For the girls, having no elder brother is detrimental to their learning outcomes. This result is consistent with the argument, that most Indian families desire to have at least one son. No such patterns exist in the less biased states. I also look at the potential channels that are responsible for the differential learning levels. It seems that the households do spend more on their eldest son by sending them for private tuitions and spend less on their daughters in the absence of a male child. Potential reason can be that households want to conserve resources in anticipation for a son. Similar results hold when we look at private school enrollment which is more expensive than fee-exempt government schools. To test son-preference assumption of households, I do a simple test to see if households with more girls tend to be bigger or not. Indeed it is the case that families with more percentage of girls tend to be larger also. Moreover, if the first two borns are girls, then the household often exceed their desired family size to get another boy.

I also test for a potential confounding factor that could distort the coefficients i.e.. introduction of Right to Education. This law was implemented between 2010-13 across various Indian states and could distort results as some children would be more exposed to RTE provisions as compared to others. It seems that RTE has not led to any changes in the gender-bias in terms of learning outcomes. Finally, as robustness check I use reading level as proxy for learning outcomes rather than math outcomes. The results, are comparable to the math outcomes. I also look at effect of gender on the birth order gradient, as well as, effect of sibling composition on the learning outcomes in the most progressive states (having adult sex ratio of more than thousand females per thousand males). The results are largely similar, with the boys facing slight disadvantage in absence of elder brother. The result could be driven by Kerala as it follows a matrilineal system of inheritance.

This study contributes to the existing literature in several important ways. First, this paper challenges the assumption that the gender-bias is common for all the girls in the same household. Existing studies that focus on differences across gender in terms of various outcomes of interest like health (Basu 1989), intra-family food distribution (Chen 1981) etc don't consider that gender-bias can be a function of sibling composition. Some recent studies like JP, have challenged this assumption.This paper does the same in the context of education. Second, most studies that study gender-bias in education focus on enrollment data and education expenditure (Azam and Kingdon 2011; Sahoo 2015). As pointed before, the actual outcome of interest is not school enrollment but actual learning. Thus, this study by focussing on learning outcomes, improves upon the existing literature. Finally, this study contributes to the large literature of program evaluation particularly of large public-funded education expansion programs. I find that, at least in India, RTE has not lead to any change in the gender-bias in the learning outcomes. This calls for a rethink of the strategies implemented in the program to attain gender-parity objectives.

The rest of the paper is organized as follows: Section 2 has detailed discussion on the existing literature; Section 3 talks about the data and summary statistics; Section 4 presents the identification strategy and results; Section 5 discusses the results and possible mechanism; Section 6 runs robustness checks and Section 7 concludes with some policy prescriptions.

2 Literature Review

From a theoretical perspective, there is a long standing interest regarding the quality-quantity tradeoff that parents face while making decisions about the optimal family size and investment in the children. Seminal work in this regard was done by Becker and Lewis (1974), which postulated that holding quality constant, the cost of having more children increases as quality increases, and vice-versa. However, in real life households may choose to invest differentially in the quality of the children based on observable dimensions like gender. Investigating and documenting differential investment in the children in the same household has generated vast amount of literature. Another motivation to study intra-household resource allocation stems from need to understand wide-spread gender discrimination prevalent in many parts of the globe, particularly developing world, which can help in formulating policies to counter it, as gender parity is considered to be a welfare-maximizing objective by many societies across the globe.

There are several important studies that document patterns of intra-household allocation of education resources. Azam and Kingdon (2011) find that bias exists for females in the decision to enroll them or not but there is no bias as far as spending on education is concerned. They also find that "households spend less on girls than boys by sending sons to fee-charging private schools and daughters to the fee-free government-funded schools". Sahoo (2015) shows in the context of Uttar Pradesh, a state in India, that "there is an intra-household gender gap of 5.4 percentage points in private school enrollment among children aged 6 to 19 years".

On the other hand, there have been important studies that document gender bias at the the household level in developing countries. Jayachandran (2014) documents the pervasive character of gender bias in developing countries along with often cited explanations. Chen et al. (1981) find that "Son preference in parental care, intra-family food distribution, feeding practices, and utilization of health services" is responsible for higher female mortality among rural households of Bangladesh. Basu (1989) contends that health expenditure rather than food expenditure is responsible for higher mortality of female children. On the other hand, Deaton (1989) finds no gender-bias based on expenditure on "adult goods".

A common missing point in these studies is that they assume that gender discrimination is done on a consistent basis among all the females in the household. There are only a handful of studies that challenge this assumption and none in education, that I am aware of. Gupta (1987) finds that daughters of higher birth order are discriminated more than first borns in Punjab, northern state of India. Jayachandran and Pande (2015) while trying to solve for the puzzle of the height advantage that African children enjoy as compared to Indian children, find that Indian first borns are taller than African counterpart, only if it is a boy. They also find that having an older girl sibling is advantageous for boys and disadvantageous for girls.

In this study, I intend to introduce a different angle through which gender-bias can be studied within the context of education, by relaxing the assumption that gender bias is same for all the girls in the household. Specifically, I look at the gender composition of the siblings, as well as the birth order of the child, to determine the quantum of bias that the girl may face. Another contribution of this study is that it tries to document the variation within different states of India, which may vary in terms of perceptions of not only the role of women in the society but also the their returns to education. Most of the earlier studies either concentrate on one or two states, or do a pan-India analysis. They don't try to disentangle regional variations that could affect the learning outcome patterns across gender. I try to figure out the variations by dividing the states on the basis of whether they have favorable sex ratio (towards girls) or not. Like many studies before, I also try to find the mechanism through which the bias operates, particularly enrollment in private schools and private tuition. As part of robustness check, I also assess the impact of Right to Education on gender bias, and find that exposure to Right to Education doesn't change the gender bias in terms of learning outcomes. Some studies like Glick (2008) in it's meta-analysis find that decreasing the cost of education can lead to increased enrollment of girls in the school. However, getting children to school cannot be the main objective of a policy maker, rather they learn is the most crucial factor. In this study, I have tried to bridge this gap by looking at learning outcomes and not just enrollment and shown that mass education expansion programs do not mitigate the gender bias at the household level in terms of learning outcomes. This finding in turn has important policy implications.

3 Data Descriptives

To investigate the question of impact of birth order and gender on the learning outcomes of the young children, I rely on the data captured in the Annual Status of Education Report-2014 (ASER) and The Census of India (2011). According to ASER, the survey is done for all the rural districts of India with 600 households surveyed in each district of the 31 states and union territories. Within each district, 30 villages are chosen and 20 households per village are surveyed. Information is collected for children aged 3-16 years with learning outcomes recorded for all in the age group 5-16 years, within each household. Data is also recorded for certain paternal and household characteristics. For the purpose of this study, I am using the data for 2014, which happens to be the latest year for which data is available.

The key variable of interest is the learning outcome of the children aged 6-16 years. In India, children enter the first grade at age 6 in most of the states. The first method to assess the learning outcome is to ascertain the math outcomes. In the data five different levels of mathematical aptitude are recorded with the lowest having no knowledge of arithmetic, and the highest being able to do division of a three digit number by a single digit.[2]. To make sense of different levels of learning, I create a new variable, Standard Math Level (Std_Math) which is binary variable equalling 1 when the child has attained the least amount of math level according to the grade in which she is studying, and 0 when she hasn't attained that level.[3] Since there are only five levels of math outcomes in the data, the above method of constructing standardized variable will assign the dummy value equal to 1 to children above grade 3 if she can do division even though the math level may be less than the requisite level of the actual standard. This is a limitation of the data. However, there is enough variation

[2] Level 1 = No Arithmetic, 2 = Recognize numbers 1-9; 3 = Recognize numbers 11-99; 4 = two-digit subtraction; 5 = division

[3] Grade 1 and 2 children should be able to recognize numbers from 1 to 99; grade 3 students should be able to do subtraction; and any child above grade 3 should be able to do division.

in the dummy for all class standards[4] and I do control for child age and standard in the main regression. Similarly, I construct a variable for reading level. The details are given later.

The standardized math learning outcome variable, is valid only for children who actually go to school. The learning outcomes of children who are not in school may be different from in-school children. Most of the analysis in this study is intra-household. Thus, to compare the learning outcomes of children of the same mother, I restrict the sample to households that send all the children in the school-going age to school. A final point in the creation of relevant data is that initially the sample had 642911 child-level observations. I drop households that have either missing mother characteristics, child characteristics or twins. This reduces the sample to 304,722 which is 47.4 percent of the original data. Thus, I have to assume that data is Missing Completely at Random (MCAR) to infer that the results are unbiased. Moreover, in the regressions with mother fixed effects, the missing data should not result in biased coefficients, as I drop the entire household whenever the data is missing even for a single child.

One of the dependent variable of interest is the birth order of the child. The birth order is defined according to the age of the child. However, since the data is only for children from 3-16 years, therefore, the birth order according to the data is not a perfect measure of the actual birth order as there can be children in the household who are younger than 3 years and not captured in the data. This pitfall of the data doesn't affect the analysis as we are interested in the impact of birth order and gender, which should kick in only among the school-going children as they are the ones competing with each other for the resources. While, it would have been optimal to get the data for all the children in the household, we can still reliably estimate the effect of birth order and gender given the data for all the school going children. We have to make the assumption that new-borns and infants (less than 3 years old) don't compete for resources differentially among the school-going children (age 6 and above). To define the birth order, I restrict it to three categories - first born (oldest), second born and later born. Other child level characteristics include child age and the class standard. Other important covariates include parent and household level characteristics.[5]

If we look at the summary statistics in Table 1, the 48 percent of all children are girl with 1.08 girls per mother on an average. The average age of the child in the sample is 9.2 years. The average number of children per mother is 2.19 according to the sample, which is lower than the national Total Fertility Rate (TFR) of 2.3 in 2013[6] as many women are yet to reach the end of their child-bearing age. The average age of mothers in the sample is

[4]Even in grade 12, only 73 percent of the students are able to do division

[5]Paternal characteristics - Parent's age and schooling. Household characteristics include index of dummies indicating if the house has electricity, toilet and television using principal component analysis.

[6]Sample Registration System Statistical Report (2013)

32.9 years with 60 percent of them having attended school. Additionally, 56 percent of the children are first born, 30 percent are second born and rest later born.

4 Identification Strategy and Results

4.1 Birth Order

As mentioned before, Becker and Lewis (1974) in his analysis assumes that the parents invest equally in their children. This has proven to be a flawed assumption as several studies have shown empirically that parents invest differentially in their children not only in terms of investment in schooling (Sahoo (2015);Azam and Kingdon (2011)), but also in other human development related investments like health expenditure (Gupta(1987)), food allocation(Behrman(1988);Sen and Sengupta (1983)) which can lead to varied development outcomes like differences in height (Jayachandran and Pande (2015)). This study focusses on another key development indicator, that is, learning outcome.

First I assign each child having same mother a birth order with the oldest child getting lowest birth order. The dependent variable is standardized math learning outcome. In the multi-variate regression, I control for two main child level characteristics that can affect the learning outcomes that is child age and the school standard in which she is studying. However, just controlling for child-level characteristics doesn't allow us to compare different households as they can be very different from each other in terms of geography, paternal characteristics, household infrastructure etc. Thus, I add household level characteristics like paternal education and index of household infrastructure. However, as JP point out that households can still vary in terms of unobservables like fertility decisions. As the women have not reached the end of child-bearing age in the sample, therefore, I cannot control for total number of children in the house. Thus, following the strategy of JP, I rely on within regression to get reliable estimates by using mother fixed effects. Thus, my preferred specification is:

$$StdMath_{im} = \beta_1 SecondChild_{im} + \beta_2 LaterChild_{im} + \theta_1 ChildAge_{im} + \theta_2 SchoolClass_{im} + \eta_m + \epsilon_{im} \text{ - (1)}$$

where $StdMath_{im}$ represents the standardized math outcomes of child i of mother m. First child is the reference category. β_1 and β_2 are the main coefficients of interest with η_m and ϵ_{im} being the mother fixed effects and error terms respectively. I cannot include any household level characteristics here, because there will be no variation among the observations for the within-mother level regression. Table 2 column 1 shows results from regression controlling for only child level characteristics, whereas column 2 includes household level characteristics. It appears that both second and later born children have disadvantage in terms of learning outcome as compared to the first born. Other covariates, except father's age, have the expected sign with the paternal education, mother age and household infrastructure having positive impact on learning. The t-test for the difference in coefficients between the coefficients between second and later born dummy is significantly different from zero. Thus, there is a clear birth-order gradient when it comes to learning. However, I focus on results from column 3 as it controls for all the time-invariant mother-level characteristics due to inclusion of mother fixed effects. The learning gradient actually increases implying that the omitted variables in column 2 specification were biasing the gradient downwards.

4.2 Gender

4.2.1 Birth Order and Gender

To understand, the impact of interaction of gender of birth order, I now include an dummy for girl child, as well as, the interaction of girl child dummy with the the birth order. The new preferred specification is then,

$$StdMath_{im} = \beta_1 SecondChild_{im} + \beta_2 LaterChild_{im} + \gamma_1 Girl_{im} + \gamma_2 (Girl_{im} * SecondChild_{im}) + \gamma_3 (Girl_{im} * LaterChild_{im}) + \theta_1 ChildAge_{im} + \theta_2 SchoolClass_{im} + \eta_m + \epsilon_{im} \text{ - (2)}$$

In this specification, γ_1 captures the impact of being a girl, while $\gamma_1 + \gamma_2$ compares the learning outcome of the girl versus a boy, conditional on being a second child. γ_2 can also be interpreted as the incremental effect on learning outcomes of a girl as she becomes a second child. Similar interpretations holds for $\gamma_{3.}$

Table 3 column 3 shows that there is differential in the birth order gradient for the second child according to gender. The learning gradient for the girls in less steep than boys for the second child but it is statistically same as boys, as far as later child is concerned. This result is a bit confusing and therefore, needs deeper probe. In India, there is a clear geographical divide in terms of attitudes towards girls both in terms of their investment in their education and their status. Mostly southern states do well in terms of gender equality, whereas the northern states exhibit a more biased attitude towards girls (Murthi et al (1995)). Thus, I divide the state into less biased and more biased states.[7]

[7]The less biased states have the adult sex ratio, defined as number of females per thousand males, more than the all-India average of 949 females per thousand males. This also the reason for using the term less biased, rather than unbiased as these states are less unequal towards women but far from granting them same status as men. More biased states are the ones where the sex ratio is worse than the national average.

The cutting of sample according to gender bias throws up some interesting trends. The birth order gradient for the girls becomes less steep for the more biased states (Table 3 column 4), whereas the gender-aspect of gradient vanishes for the less biased states (Table 3 column 5). Important point to note is that the coefficient for the girl dummy also reduces and becomes less significant for the less biased versus more biased states. In order to directly compare the coefficients of less and more biased states, I interact all the coefficients in Equation 2, with the dummy for biased states. The results are reported in the Table A1. However, as pointed before, birth order gradient being less steep for girls in the more biased states is a bit surprising that motivates the next step in the analysis.

4.2.2 Big Brother Effect

Jayachandran and Pande (2015) while trying to solve the India-Africa height puzzle, find that "an Indian son with an older sibling is taller than his African counterpart if and only if he is the eldest son" and "height deficit is largest for daughters with no older brothers." In order to fully comprehend the trends thrown up in table 3, particularly less steep learning gradient of girls in the more biased states, I use the strategy used in JP and run separate regressions for the more and less biased states. The estimating equation now becomes:

$StdMath_{im} = \beta_1 SecondChild_{im} + \beta_2 LaterChild_{im} + \gamma_1 Girl_{im} + \delta_1(NoElderBro_{im}) + \delta_2(Girl_{im} * NoElderBro_{im}) + \theta_1 ChildAge_{im} + \theta_2 SchoolClass_{im} + \eta_m + \epsilon_{im}$ - (3)

$NoElderBro$ represents dummy for whether the child had an elder brother or not. Thus, a first born boy or girl will also have this dummy equal to 1 as they don't have any elder sibling. For the later borns, this dummy equals one only when they don't have an elder brother. There is also an interaction term of the elder brother and girl child dummy to differentiate the impact of having no elder brother between genders. In table 4 column 3, δ_1 is positive and significant, while δ_2 is negative and significant when we consider the entire sample. Thus, having an no elder brother has different impact for boys and girls. The coefficients' sign and significance remain the same for the more biased states whereas, both the coefficients are not significant for the less biased states. Again to differentiate the coefficients between more and less biased states, I run a regression interacting every covariate with the dummy for more biased state. The results are reported in the Table A2.

Finally, to disentangle the effect of having no elder brother on the birth order gradient I add a triple interaction term of gender, birth order and dummy for elder brother to Equation (2). Thus, the specification becomes:

$StdMath_{im} = \beta_1 SecondChild_{im} + \beta_2 LaterChild_{im} + \gamma_1 Girl_{im} + \gamma_2(Girl_{im} * SecondChild_{im})$

$+ \gamma_3(Girl_{im} * LaterChild_{im}) + \rho_1(Girl_{im} * SecondChild_{im} * NoElderBro_{im}) + \rho_2(Girl_{im} * LaterChild_{im} * NoElderBro_{im}) + \theta_1 ChildAge_{im} + \theta_2 SchoolClass_{im} + \eta_m + \epsilon_{im}$ - (4)

In the above equation, the coefficients ρ_1 and ρ_2 are the of interest. As can be seen from table 5 (column 2), for the more biased states, girls at the second and later birth order tend to do better if they have an elder brother (γ_2 and γ_3 are both positive and significant). On the other hand, if they don't have an elder brother they tend to do worse (ρ_1 and ρ_2 are negative, though ρ_2 is not significant). For the less-biased states (Table 5 column 3) none of the coefficients are significant.

5 Discussion of Results and Possible Mechanism.

Till now, we have found that (i) birth order matters for learning outcome, with higher birth order detrimental for learning; (ii) the birth order gradient varies across gender but only in more biased states (iii) having an elder brother or not matters and that too differentially across gender in the more biased states only, and (iv) variation in the birth order gradient across gender is largely driven by presence or absence of elder brother.

In the above analysis, I tried to disentangle the effects of birth order, gender and sibling composition in a systematic and rigorous manner. The effect of having an elder brother or not is strong. As pointed out in JP, this result is consistent with the idea that in Hindu religion, the eldest son has special place as he not only inherits the paternal property but also carries out the last rites of the parents after their demise. Thus, we can expect that parents will try to invest more in their eldest son. In table 4, we find that having no elder brother is beneficial for the boy as he then becomes the eldest son and may get more resources from the parents . On the other hand, having no elder brother results in poorer learning outcomes for girls as parents tend to conserve their resources in anticipation for the boy. This line of thought leads to following testable hypotheses:

1) Households will spend lower on the education of girls with no elder brother.

2) Households will exceed their optimal fertility rate, if they haven't had a son yet.

To test the the first hypothesis I check if the households decide to send their children to private tuition, based on birth order and sibling composition. One would expect that if the households really want to discriminate against the girl child, conditional on having elder brother, then they would not spend extra money on them by sending them for private tuitions. Another way to test if households are investing differentially in the children is to see if they send their children to the private schools. In India private schools tend to be more expensive than the public schools, even after the advent of so-called Low Fee Private

(LFP) schools. Härmä (2009) estimates the full cost to parents in sending their kids to private schools in the range of \$22.42-26.44 per annum as compared to cost of \$3 per annum in government schools. Table 6 (column 2 and 3) show results for whether the parents send their children for private tuition or not after controlling for mother-level time invariant characteristics. It seems that households in more biased states do discriminate against girls who don't have elder brother. On the other hand, boys who don't have elder brother enjoy more expenditure on their education. In the less biased states, having no elder brother is not significant for boys , whereas it is negative for girls in these states too. The coefficient is however, lower in magnitude. Evidence from decision to send their children to private school is more noisy (Table 7; column 2 and 3) with the coefficient for boys having no elder brother insignificant for boys, and negatively significant for girls in both the category of states. This could be result of overall change in preferences of households towards sending their children to private schools, as they are perceived to be better (33 percent of the children in the sample are going to private schools, whereas the figure was only 20 percent in 2007). Thus, it is possible that more boys are being sent to the private schools, regardless of whether they are first born or not, and girls are continued to be discriminated against[8].

To test the second hypothesis I run two different regressions at the household level and test if the number of children in the household depend on a) the proportion of females in the house, and b) event when both the first and second child are girls. In the first case, one can expect that households would go for higher number of children as the proportion of females increase if there is son-preference in the society. The choice of second kind of specification is motivated by the fact, that the households in the sample have on average 2.19 children per mother as compared to the national average of 2.3 children per mother. These numbers can be taken as proxy for the desired number of children that a household desires. If a woman has first two children as girls, and she desires at least one son, then we should see that among all the households that have at least two children, the ones with having first two as girls, will have a higher probability of exceeding the optimum amount. Consequentially, a rational household will try to conserve resources for the expected boy and this can explain the fact that the girls with no elder brother have worse learning outcomes. Thus, for the second specification I run a regression for a subsample of households that have at least 2 children, and try to ascertain if the households having first two children as girls have different number of children than households who have at least a boy among the first two borns.

Table 8 documents the results for the two different specifications. In the first three columns I report the total number of children as a function of proportion of girls, as well

[8]For instance, Sahoo (2015) finds intra-household gender gap of 5.4 percentage points in private school enrollment among children aged 6 to 19 year

as, other relevant covariates. The number of children increase as the proportion of girls increase both for the pooled sample (column 1), as well as, for individually for more and less biased states (column 2 and 3). As expected, the households in the less biased states base their decision of having children less on the girls already present as compared to more biased states[9]. Column 4-6 represent the second specification where I just take the subsample of households having at least two children. It is evident that households go for more children if the first two happen to be girls. This particularly true for the households in the more biased states as compared to the ones in the less biased states. The results, of both these regressions suggest that girls live in larger households and it seems that households tend to have more children till they get at least one boy. In the pursuit if that objective, they often exceed their target number of children, which could result in less flow of resources towards females if they don't have an elder brother.[10]

6 Robustness of the Results

6.1 Eliminating the possible confounding factor - RTE.

One of the potential confounding factors in the analysis could be introduction of RTE (Right to Education Act), which brought down the cost of education in many states. It was implemented across India states from 2010 to 2013 in a phased manner. Thus, it will be true that older children will be more exposed to benefits/pitfalls of RTE than the younger cohort.[11]

To test if the exposure to RTE is driving the results, I create a new variable which capture the number of years child was exposed to RTE once it was implemented.[12]Then, I run the regression given by the equation (4), but also include the variable that captures the exposure to RTE. Results are shown in Table 9. Column 1 shows that the result of pooled sample and the results are similar to Table 4. Not surprisingly, exposure to RTE has a negative coefficient. There have been studies which document that introduction of RTE has led to

[9]Coefficient of proportion of girls already present for less biased states is 0.099, whereas for more biased states it is 0.188, almost two times that of less biased states

[10]These results should be analyzed carefully, because most households have not completed their fertility cycle as, on average, women have not reached the end of child-bearing age. Also, I do not observe new born babies below the age of 3 years, which could potentially bias the results.

[11]This relationship is actually more nuanced. The RTE affects only children between 6-14 years. So if a child is born in a state that implements RTE in 2010 and is aged between 14 - 10 years in 2014, she will be exposed to RTE for 4 years. If she is 15 years old in 2014, then she will be exposed to RTE only for 3 years. On the other hand, if she is 9 years old, then she will be exposed again for 3 years.

[12]I define implementation of RTE as the year when the states notify the RTE rules in the official gazette.

decline of education standards in the country (Gupta and Prakash 2015). Column 2 and 3 show similar trends, where having no elder brother is better for boys and worse for girls, in the more biased states, whereas having elder brother or not doesn't matter for both boys and girls in the less biased states. Thus, it can be concluded that introduction of RTE is not confounding the results reported so far.

6.2 Alternate way to capture learning outcomes.

Instead of using math proficiency as instrument for learning outcomes, I can also use reading proficiency as measure of learning. Again I create a dummy for standardized learning outcome for english, which takes a value of 1 if the child has attained the level which she should have according to the standard in which she is in[13]. In table 10, I report the main results from the key regression specifications for more and less biased states. In column 1 and 2, I replicate the regression specified in Equation (2) for more and less biased states respectively, with reading level as the dependent variable. The result for the more biased states, is similar to the math level regressions, however the coefficient for birth order gradient for females born later is not significant, though the sign is same. In a similar vein, birth order gradient for second born girls is significant for the less biased states, whereas it was insignificant for the math outcomes. These slight differences in the trends between math and reading level outcomes can be due to presence of differential spillovers within a household among siblings, as math and reading require different kind of skill sets.

Column 3 and 4, use reading outcomes to test equation (3), which investigates the differential impact of having an elder brother across genders. Column 3 reports results for more biased states. Similar to math outcomes, results indicate that it is better for boys to be the eldest boy in the household (i.e have no elder brother) and it is worse for girls if they don't have an elder brother. This result is exactly similar to the math outcomes, which is very encouraging. For the less biased states (column 4), having an elder brother doesn't matter for boys, but is worse for girls. This result, does diverge a bit for girls in the less biased states, when we compare it to math outcomes (table 4), but it can be result of using two different instruments (math vs reading) in capturing learning outcomes.

6.3 Restricting the sample to not-biased states.

So far I have divided the states among more and less biased states. There are, however,

[13]Original reading in the data is : level 1 = cannot read anything; level 2 = identify letters; level 3 = can read words; level 4 = can read Standard 1 text; level 5 = can read Standard 2 level text.

five states and union territories that have adult sex ratio greater than thousand.[14]I test equation (2) and (3) using a sub-sample of children from these 5 states. The result is reported in table 11. As expected, the birth order gradient for the girls is not different from the boys (column 1). Also, the girls are not disadvantaged if they are don't have an elder brother. For the boys, having no elder brother matters actually negatively affects their learning outcomes, thought the coefficient is barely significant (the coefficient has p- value of 0.097). Interestingly, the sign of the coefficient is also negative. This implies that having elder brother may be slightly beneficial for boys in these states - which may be a consequence of positive spillovers of learning outcomes among children of the same gender. Another way to think about this result, is that having a girl as elder sibling is worse for the later-born boys. This result could be driven by the state of Kerala, which follows matrilineal form of inheritance, and hence girls may be favored than boys.

7 Conclusion and Policy Prescription

To summarize, the key take away of this study is that gender bias is prevalent in India as far as educational outcomes are concerned. The gender bias, contrary to what is assumed in most studies, is not the same for all the children in the household. It varies according to both the birth order and composition of the siblings. It seems that the most important factor is the birth of eldest son. Specifically, if the boy is the eldest son, then they tend to have higher learning outcomes. On the other hand, if a girl doesn't have an elder brother, then she has lower learning outcomes. The results vary according to geographical regions. Suggestive evidence points out that trends in the learning outcomes may be driven by differential resource allocation by the households.

There are important implication of this analysis. It proves that the assumption of common gender bias (across all children in the household) is flawed. Thus, we need to take this fact into consideration while developing policies to combat gender-bias. For instance, many states in India incentive greater enrollment of girls by giving them tuition waivers or scholarship or other incentives like free bicycles. It is possible to improve the efficacy of these schemes by identifying the most disadvantaged children based on observable covariates like sibling composition and birth order and direct resources more towards these children.

Finally, the advantage of this study is that focusses on the "real" variable of interest, that is learning outcomes. Most of the work till now has been on various inputs like school enrollment etc. As the quality of the data gets better, we need to shift focus from inputs to outcomes. This will also need re-thinking in the way policies are designed. Large education

[14]The five states are: Chhattisgarh, Goa, Kerala, Pondicherry and Uttaranchal.

-expansion schemes often try to increase enrollment, without adequately focussing on final outcome. I have shown that neglecting these outcomes, led to persistence of inequality in the society, even in the presence of mass-expansion of school education.

References

[1] Azam, Mehtabul, and Geeta Gandhi Kingdon. "Are girls the fairer sex in India? Revisiting intra-household allocation of education expenditure." World Development 42 (2013): 143-164.

[2] Gupta, Monica Das. "Selective discrimination against female children in rural Punjab, India." Population and development review (1987): 77-100.

[3] Behrman, Jere R. "Intra-household allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion?." Oxford Economic Papers 40, no. 1 (1988): 32-54.

[4] Sen, Amartya, and Sunil Sengupta. "Malnutrition of rural children and the sex bias." Economic and political weekly (1983): 855-864.

[5] Murthi, Mamta, Anne-Catherine Guio, and Jean Dreze. "Mortality, fertility, and gender bias in India: A district-level analysis." Population and development review (1995): 745-782.

[6] Härmä, Joanna. "Can choice promote Education for All? Evidence from growth in private primary schooling in India." Compare 39, no. 2 (2009): 151-165.

[7] Sahoo, Soham. "Intra-Household Gender Disparity in School Choice: Evidence from Private Schooling in India." Available at SSRN (2015).

[8] Gupta, Bhanu and Nishith Prakash , "The Impact of Change in Primary Completion Rate on Learning Outcomes in India", Background paper prepared for the Education for All Global Monitoring Report (2015).

[9] Becker, Gary S., and H. Gregg Lewis. "Interaction between quantity and quality of children." In Economics of the family: Marriage, children, and human capital, pp. 81-90. University of Chicago Press, 1974.

[10] Glick, Peter. "What policies will reduce gender schooling gaps in developing countries: Evidence and interpretation." World Development 36, no. 9 (2008): 1623-1646.

[11] Jayachandran, Seema, and Rohini Pande. Why are Indian children so short?. No. w21036. National Bureau of Economic Research, 2015.

[12] Jayachandran, Seema. The roots of gender inequality in developing countries. No. w20380. National Bureau of Economic Research, 2014.

[13] Deaton, Angus. "Looking for boy-girl discrimination in household expenditure data." The World Bank Economic Review 3, no. 1 (1989): 1-15.

[14] Chen, Lincoln C., Emdadul Huq, and Stan d'Souza. "Sex bias in the family allocation of food and health care in rural Bangladesh." Population and development review (1981): 55-70.

[15] Basu, Alaka Malwade. "Is discrimination in food really necessary for explaining sex differentials in childhood mortality?." Population studies 43, no. 2 (1989): 193-210.

[16] ASER (2014), Annual Status of Education Report, Pratham Resource Centre, New Delhi.

[17] The Census of India (2011).

Table 1: Summary Statistics

	(1)				
	count	mean	sd	min	max
SchoolClass	246121	5.23681	2.992877	1	12
Math level	246121	.3846116	.4865044	0	1
Girl_number	304722	1.086256	.9449538	0	8
Child_number	304722	2.1982	1.013399	1	8
D_Girl	304722	.4782096	.4995258	0	1
ChildAge	304722	9.257287	3.875598	3	16
First_child	304722	.5672843	.495453	0	1
SecondChild	304722	.302968	.4595422	0	1
LaterChild	304722	.1297478	.3360263	0	1
D_Girl*SecondChild	304722	.1424577	.3495195	0	1
D_Girl*LaterChild	304722	.0596872	.2369068	0	1
NoElderBr	304722	.7706697	.4204028	0	1
D_Girl*NoElderBr	304722	.3696451	.4827094	0	1
D_Girl*SecondChild*NoElderBr	304722	.0731716	.2604184	0	1
D_Girl*LaterChild*NoElderBr	304722	.0204088	.1413942	0	1
D_MotherSchool	304722	.5946732	.490956	0	1
MotherAge	304722	32.95175	6.824267	17	80
FatherAge	304722	38.12555	7.673158	17	85
D_FatherSchool	304722	.7785358	.4152329	0	1
HHInfra	304722	.0028379	1.341436	-2.448127	1.358829
Dummy_schoolpvt	256761	.3350548	.4720105	0	1
Dummy_tuition	247179	.2254601	.4178858	0	1
Std_ReadLevel	245856	.486126	.4998085	0	1
Exposure_RTE	210718	2.466377	.8629279	1	4

TABLE 2 - Impact of Birth Order

Dependent variable (Math Level)	OLS (Full Sample) (1)	OLS (Full Sample) (2)	FE (Full Sample) (3)
SecondChild	-0.0461***	-0.0452***	-0.0714***
	(0.00205)	(0.00205)	(0.00355)
LaterChild	-0.108***	-0.0845***	-0.135***
	(0.00328)	(0.00329)	(0.00662)
ChildAge	0.00864***	0.0129***	0.0100***
	(0.000857)	(0.000847)	(0.00147)
SchoolClass	0.0177***	0.00883***	0.00111
	(0.000904)	(0.000880)	(0.00137)
D_MotherSchool		0.102***	
		(0.00253)	
MotherAge		0.00649***	
		(0.000308)	
FatherAge		-0.00194***	
		(0.000269)	
D_FatherSchool		0.0766***	
		(0.00269)	
HHInfra		0.0525***	
		(0.000875)	
Constant	0.227***	-0.0374***	0.309***
	(0.00553)	(0.00747)	(0.0132)
Observations	246,121	246,121	246,121
R-squared	0.040	0.097	0.048
Hh Controls	NO	YES	NO
Mother FE	NO	NO	YES
Second Child - Third Child = 0	.0617348	.0393246	.0634373
p-value	0	0	0

Robust standard errors (clustered at Hh level) in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

TABLE 3 - Impact of Birth Order and Gender on Math Level

Dependent variable (Math Level)	OLS (Full Sample) (1)	OLS (Full Sample) (2)	FE (Full Sample) (3)	FE (More Biased) (4)	FE (Less Biased) (5)
D_Girl	-0.0465***	-0.0477***	-0.0528***	-0.0912***	-0.0106*
	(0.00256)	(0.00248)	(0.00389)	(0.00537)	(0.00562)
SecondChild	-0.0459***	-0.0455***	-0.0823***	-0.0842***	-0.0825***
	(0.00295)	(0.00291)	(0.00447)	(0.00603)	(0.00663)
LaterChild	-0.107***	-0.0841***	-0.143***	-0.145***	-0.135***
	(0.00434)	(0.00430)	(0.00743)	(0.00982)	(0.0114)
D_Girl*SecondChild	-0.00237	-0.00136	0.0198***	0.0310***	0.00450
	(0.00431)	(0.00417)	(0.00536)	(0.00733)	(0.00781)
D_Girl*LaterChild	-0.00483	-0.00378	0.00746	0.0322***	-0.0141
	(0.00575)	(0.00561)	(0.00667)	(0.00864)	(0.0106)
Observations	246,121	246,121	246,121	119,342	126,779
R-squared	0.042	0.099	0.051	0.077	0.030
Hh Controls	NO	YES	NO	NO	NO
Mother FE	NO	NO	YES	YES	YES

Robust standard errors (clustered at Hh level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 4 : Big Brother Effect - Math Level

Dependent variable (Math Level)	OLS (Full Sample) (1)	OLS (Full Sample) (2)	FE (Full Sample) (3)	FE (More Biased) (4)	FE (Less Biased) (5)
D_Girl	-0.0343***	-0.0401***	-0.0310***	-0.0459***	-0.0151**
	(0.00397)	(0.00385)	(0.00427)	(0.00557)	(0.00659)
SecondChild	-0.0422***	-0.0469***	-0.0681***	-0.0621***	-0.0789***
	(0.00255)	(0.00251)	(0.00443)	(0.00598)	(0.00657)
LaterChild	-0.103***	-0.0869***	-0.132***	-0.119***	-0.139***
	(0.00394)	(0.00390)	(0.00777)	(0.0103)	(0.0118)
NoElderBr	0.0182***	0.00357	0.0149***	0.0281***	-0.000840
	(0.00360)	(0.00351)	(0.00452)	(0.00607)	(0.00676)
D_Girl*NoElderBr	-0.0169***	-0.0107**	-0.0125**	-0.0306***	0.00771
	(0.00451)	(0.00437)	(0.00605)	(0.00805)	(0.00914)
Constant	0.237***	-0.0130	0.318***	0.276***	0.364***
	(0.00664)	(0.00824)	(0.0144)	(0.0193)	(0.0215)
Observations	246,121	246,121	246,121	119,342	126,779
R-squared	0.042	0.099	0.051	0.077	0.030
Hh Controls	NO	YES	NO	NO	NO
Mother FE	NO	NO	YES	YES	YES

Robust standard errors (clustered at Hh level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 5 - Impact of Birth Order conditional on No Elder Brother -Math Level

Dependent variable (Math Level)	FE (Full Sample) (1)	FE (More Biased) (2)	FE (Less Biased) (3)
D_Girl	-0.0580***	-0.101***	-0.00772
	(0.00476)	(0.00643)	(0.00708)
SecondChild	-0.0850***	-0.0889***	-0.0811***
	(0.00467)	(0.00625)	(0.00700)
LaterChild	-0.146***	-0.151***	-0.133***
	(0.00761)	(0.0100)	(0.0118)
D_Girl*SecondChild	0.0324***	0.0544***	-0.00173
	(0.00817)	(0.0109)	(0.0123)
D_Girl*LaterChild	0.0137	0.0484***	-0.0189
	(0.00838)	(0.0107)	(0.0135)
D_Girl*SecondChild*NoElderBr	-0.0140**	-0.0272***	0.00662
	(0.00669)	(0.00907)	(0.00989)
D_Girl*LaterChild*NoElderBr	-0.00222	-0.0174	0.00482
	(0.00948)	(0.0122)	(0.0150)
Constant	0.340***	0.317***	0.364***
	(0.0134)	(0.0181)	(0.0200)
Observations	246,121	119,342	126,779
R-squared	0.051	0.077	0.030
Hh Controls	NO	NO	NO
Mother FE	YES	YES	YES

Robust standard errors (clustered at Hh level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 6: Mechanism 1- Decision to send children for Private Tuition

Dependent variable (Dummy for Tuition)	FE (Full Sample) (1)	FE (More Biased) (2)	FE (Less Biased) (3)
SecondChild	-0.00319	0.00166	-0.0101***
	(0.00266)	(0.00368)	(0.00387)
LaterChild	-0.000314	0.00602	-0.00600
	(0.00484)	(0.00651)	(0.00727)
D_Girl	-0.0228***	-0.0319***	-0.0125***
	(0.00260)	(0.00354)	(0.00381)
NoElderBr	0.0151***	0.0259***	0.00355
	(0.00271)	(0.00374)	(0.00391)
D_Girl*NoElderBr	-0.0115***	-0.0132***	-0.0104*
	(0.00364)	(0.00497)	(0.00536)
Constant	0.169***	0.167***	0.165***
	(0.00926)	(0.0127)	(0.0136)
Observations	233,014	111,956	121,058
R-squared	0.029	0.040	0.019
Number of mother_id	138,801	62,261	76,540

Robust standard errors (clustered at Hh level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Observations are different because some points are missing
for the dependent variable

TABLE 7: Mechanism 2- Decision to send children to Private School

Dependent variable (Dummy for Private School)	FE (Full Sample) (1)	FE (More Biased) (2)	FE (Less Biased) (3)
D_Girl	-0.0470***	-0.0631***	-0.0263***
	(0.00287)	(0.00392)	(0.00420)
SecondChild	-0.0368***	-0.0238***	-0.0486***
	(0.00302)	(0.00424)	(0.00427)
LaterChild	-0.0648***	-0.0505***	-0.0814***
	(0.00549)	(0.00753)	(0.00804)
NoElderBr	-0.000703	0.00594	-0.00368
	(0.00299)	(0.00412)	(0.00434)
D_Girl*NoElderBr	-0.0149***	-0.0160***	-0.0154**
	(0.00414)	(0.00567)	(0.00603)
Constant	0.368***	0.401***	0.356***
	(0.0107)	(0.0151)	(0.0150)
Observations	245,577	119,105	126,472
R-squared	0.019	0.033	0.010
Number of mother_id	145,112	65,770	79,342

Robust standard errors (clustered at Hh level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Observations are different because some points are missing
for the dependent variable

TABLE 8: Mechanism 3- Regression at Households level: Girls tend to be in Bigger Households

Dependent variable (Number of Children)	OLS (All States) (All Hh) (1)	OLS (More Biased) (All Hh) (2)	OLS (Less Biased) (All Hh) (3)	OLS (All States) (Hh>2children) (4)	OLS (More Biased) (Hh>2children) (5)	OLS (Less Biased) (Hh>2children) (6)
Girl_Proportion	0.131***	0.188***	0.0988***			
	(0.00477)	(0.00815)	(0.00566)			
First&Second_BothGirls				0.271***	0.321***	0.233***
				(0.00605)	(0.00950)	(0.00757)
MotherAge	0.00670***	0.000648	0.00427***	0.0109***	0.00548***	0.00767***
	(0.000609)	(0.00108)	(0.000727)	(0.000704)	(0.00122)	(0.000851)
FatherAge	-0.00323***	0.00264***	-0.00105	-0.00227***	0.00335***	0.000290
	(0.000540)	(0.000984)	(0.000642)	(0.000624)	(0.00109)	(0.000761)
D_MotherSchool	-0.146***	-0.145***	-0.0888***	-0.146***	-0.126***	-0.112***
	(0.00586)	(0.00862)	(0.00791)	(0.00584)	(0.00848)	(0.00797)
D_FatherSchool	0.0590***	0.0519***	0.0343***	0.0112*	0.00586	-0.00843
	(0.00668)	(0.0101)	(0.00866)	(0.00677)	(0.00995)	(0.00891)
HHInfra	-0.0746***	-0.0788***	-0.0477***	-0.0662***	-0.0628***	-0.0485***
	(0.00224)	(0.00310)	(0.00318)	(0.00216)	(0.00294)	(0.00309)
Constant	1.656***	1.702***	1.566***	2.169***	2.196***	2.102***
	(0.0136)	(0.0210)	(0.0173)	(0.0150)	(0.0222)	(0.0196)
Observations	172,864	78,285	94,579	92,321	45,040	47,281
R-squared	0.029	0.032	0.013	0.073	0.068	0.058

Robust standard errors (clustered at village level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 9 - Does Right to Education drive the Results?

Dependent variable (Math Level)	FE (Full Sample) (1)	FE (More Biased) (2)	FE (Less Biased) (3)
D_Girl	-0.0374***	-0.0568***	-0.0182**
	(0.00486)	(0.00647)	(0.00729)
SecondChild	-0.0459***	-0.0432***	-0.0549***
	(0.00515)	(0.00703)	(0.00752)
LaterChild	-0.110***	-0.101***	-0.113***
	(0.00906)	(0.0122)	(0.0135)
NoElderBr	0.0151***	0.0277***	0.000433
	(0.00506)	(0.00687)	(0.00744)
D_Girl*NoElderBr	-0.0128*	-0.0281***	0.00353
	(0.00686)	(0.00926)	(0.0102)
Exposure_to_RTE	-0.0601***	-0.0371***	-0.0823***
	(0.00176)	(0.00237)	(0.00260)
Constant	0.359***	0.298***	0.413***
	(0.0184)	(0.0250)	(0.0269)
Observations	210,718	99,853	110,865
R-squared	0.067	0.089	0.056
Hh Controls	NO	NO	NO
Mother FE	YES	YES	YES

Robust standard errors(clustered at Hh level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Number of observations different because RTE not implemented in Jammu & Kashmir and some children not exposed to RTE

TABLE 10: Robustness 1- Birth order and Big Brother effect with English level

Dependent variable (English Level)	FE (More Biased) (1)	FE (Less Biased) (2)	FE (More Biased) (3)	FE (Less Biased) (4)
D_Girl	-0.0291***	0.0246***	-0.00733	0.0445***
	(0.00484)	(0.00505)	(0.00511)	(0.00590)
SecondChild	-0.0429***	-0.0605***	-0.0335***	-0.0538***
	(0.00544)	(0.00594)	(0.00544)	(0.00599)
LaterChild	-0.0854***	-0.101***	-0.0787***	-0.102***
	(0.00900)	(0.0103)	(0.00950)	(0.0107)
D_Girl*SecondChild	0.0145**	0.0139*		
	(0.00671)	(0.00711)		
D_Girl*LaterChild	0.00651	-0.000913		
	(0.00795)	(0.00961)		
NoElderBr			0.0139**	0.00953
			(0.00557)	(0.00612)
D_Girl*NoElderBr			-0.0188**	-0.0198**
			(0.00739)	(0.00835)
Constant	-0.0272*	-0.0154	-0.0425**	-0.0253
	(0.0163)	(0.0178)	(0.0175)	(0.0194)
Observations	119,220	126,636	119,220	126,636
R-squared	0.326	0.317	0.326	0.317
Mother FE	YES	YES	YES	YES

Robust standard errors (clustered at Hh level) in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Observations Number is different because dependent
variable is missing for some

TABLE 11: Robustness 2- Birth order and Big Brother effect for unbiased states

Dependent variable (Math Level)	FE (Not Biased) (1)	FE (Not Biased) (2)
D_Girl	-0.0193	-0.0450**
	(0.0156)	(0.0178)
SecondChild	-0.0601***	-0.0872***
	(0.0186)	(0.0176)
LaterChild	-0.0891***	-0.133***
	(0.0312)	(0.0311)
D_Girl*SecondChild	-0.0208	
	(0.0214)	
D_Girl*LaterChild	-0.0362	
	(0.0271)	
NoElderBr		-0.0303*
		(0.0182)
D_Girl*NoElderBr		-0.00384
		(0.0245)
Constant	0.358***	0.406***
	(0.0553)	(0.0587)
Observations	16,287	16,287
R-squared	0.032	0.032
Mother FE	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Observations Number is different because sample is of states with sex ratio > 1000

Table A1 - Comparing Coefficients between More & Less Biased states

(Birth Order and Gender)

Dependent variable (Math Level)	FE (Full Sample) (1)
D_Girl	-0.0106*
	(0.00562)
SecondChild	-0.0825***
	(0.00663)
LaterChild	-0.135***
	(0.0114)
D_Girl*SecondChild	0.00450
	(0.00781)
D_Girl*LaterChild	-0.0141
	(0.0106)
D_BiasedState*D_Girl	-0.0806***
	(0.00777)
D_BiasedState*SecondChild	-0.00163
	(0.00896)
D_BiasedState*LaterChild	-0.0101
	(0.0151)
D_BiasedState*D_Girl*SecondChild	0.0265**
	(0.0107)
D_BiasedState*D_Girl*LaterChild	0.0463***
	(0.0136)
Constant	0.340***
	(0.0135)
Observations	246,121
R-squared	0.055
Hh Controls	NO
Mother FE	YES

Robust standard errors (clustered at Hh level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2 - Comparing Coefficients between More & Less Biased states

(Big Brother Effect)

Dependent variable (Math Level)	FE (Full Sample) (1)
D_Girl	-0.0151**
	(0.00659)
SecondChild	-0.0789***
	(0.00657)
LaterChild	-0.139***
	(0.0118)
NoElderBr	-0.000840
	(0.00676)
D_Girl*NoElderBr	0.00771
	(0.00914)
D_BiasedState*D_Girl	-0.0308***
	(0.00863)
D_BiasedState*SecondChild	0.0168*
	(0.00889)
D_BiasedState*LaterChild	0.0201
	(0.0157)
D_BiasedState*NoElderBr	0.0289***
	(0.00908)
D_BiasedState*D_Girl*NoElderBr	-0.0383***
	(0.0122)
Constant	0.322***
	(0.0145)
Observations	246,121
R-squared	0.055
Hh Controls	NO
Mother FE	YES

Robust standard errors (clustered at Hh level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1