STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics 5350
Master's thesis in economics
Academic year 2015–2016

Fertility and Sex Ratio in India:

Empirical evidence of gender bias among children in India

Srinidhi Srinivasan (40781)

Abstract

Research suggests that as a country develops and its economy grows, gender discrimination and the gap in the sex ratio is expected to decrease. However, in India, the sex ratio among children aged 0-6 has worsened from 945 girls per 1000 boys in 1991 to 914 girls per 1000 boys in 2011 despite huge strides in growth and development. During the same period, the total fertility in the country declined from 3.79 to 2.53 children per women. In a country with strong son preference, the concomitant trend between fertility and sex ratio suggests a potential relation. Using the data from three rounds of the National Family Health Survey (1992-93, 1998-99 and 2005-06), I empirically test the relation between the proportion of boys in each family and the total number of children (family size). Controlling for time and state fixed effects as well as for other baseline characteristics, I find that in general, smaller families have a larger proportion of boys. When the family size decreases by 1 person, the proportion of boys increases by 2.8 percentage points. The impact of a decrease family size on the proportion of boys is larger for 2005-06 data compared to 1992-93 data. Son preference has a significant influence on the realized sex ratio and also influences the size on the relation between family size and proportion of boys. Recommended policy focus areas are discussed.

Keywords: Sex Ratio, Fertility, Family Size, Son Preference, Gender Bias, India

JEL Classification: A13, N35, J13, J16

Supervisor: Martina Björkman Nyqvist
Date submitted: May 16, 2016
Date examined: May 23, 2016
Discussants: Hannes Skugghall
Examiner: Maria Perrotta Berlin

Acknowledgements

First and foremost, I would like to thank my advisor Martina Björkman Nyqvist for her valuable guidance, comments and discussions during the course of this thesis. I am grateful to Aikaterini Ploska, Marisa Basten, Sama Bombaywala and Shilaja Srinivasan for reviewing and proofreading my paper. I would also like to thank my family and friends for their moral support and motivation. All errors are my own.

Contents

1. Introduction 4
2. Background 8
2.1 Gender Bias in Asia 8
2.2 Gender Bias in India 9
3. Literature Review 12
4. Purpose and Research Questions 15
5. Data and Empirical Approach 17
5.1 Data – National Family Health Survey (NFHS) 17
5.2 Variables of Interest 18
5.3 Empirical Approach 23
5.4 Identification and threats to identification 24
6. Results 28
6.1 Estimation using first baseline specification 28
6.2 Estimation using second baseline specification 33
6.3 Estimation using interactions 35
6.4 Estimation using 2SLS 37
7. Discussion and Conclusions 40
8. References 44
9. Appendix 49
9.1 Appendix A: Age Specific Fertility Rate 49
9.2 Appendix B: Family Size and Proportion of Boys by State and Round 50
9.3 Appendix C: Regression varying family size 51

1. Introduction

The term "missing women" was first introduced by Amartya Sen, Economics Nobel Laureate, in 1990 to explain the shortfall in the number of women as seen in the suspiciously low sex ratio in Asia and North Africa.[1] He estimated that social inequality and injustice towards women has resulted in over that 100 million "missing women" worldwide (Sen, 1992). This ratio has worsened over the years especially in East and South Asia. In 2010, the estimated total number of missing women worldwide had risen to 126 million and is expected to peak at 150 million in 2035 (Bongaarts and Guilmoto, 2015). This persistence of low sex ratio despite economic development has brought the issue of gender inequality to the forefront of policy and development discussions.

A major reason for the widening gap in the sex ratio and the rise in missing girls has been attributed to the strong cultural preference for male children in Asian Countries (Park and Cho, 1995; Das Gupta and Bhat, 1997; Arnold, Choe, and Roy, 1998; Bhat and Zavier, 2003; Das Gupta, Chung and Shuzhou, 2009; Li, Yi, and Zhang, 2011; Chaudhri and Jha, 2013; Mitra, 2014). This preference has been well documented by literature in South Asia (Rahman et al, 1992; Rahman and DaVanzo, 1993; Das Gupta and Bhat, 1997), China (Wen, 1992; Wen, 1993; Ebenstein, 2010) and East Asia (Park, 1983; Park and Cho, 1995). Reasons for preference of male children is deeply rooted in patriarchal practices, religious practices, custom of dowry and dependence of parents on the son when they are older (Chaudhuri, 2012). Thus, the resulting sex ratio is an outcome of gender preference and fertility decisions of the family.

During the period of worsening sex ratio, Asian countries also saw a dramatic decline in the overall fertility rate. The total fertility rate in Asia as a whole, declined from 5.8 children per woman in 1960-65 to 2.3 children per woman in 2005-10.[2] This fertility decline has been much larger in certain countries. China and South Korea saw a fall in the fertility rate from around 6 children in 1960 to around 1.5 children in 2010, while in Pakistan the fertility rate was still high at around 3.5 children in 2010. Researchers have suggested that this fertility, coupled with the preference for male children, could explain the widening gap in the sex ratio (Park and Cho, 1995; Das Gupta, 2005; Ebenstein, 2010; Chaudhri and Jha, 2013). In China, researchers have found the introduction of the one child policy coupled with son preference to be causally linked to the "missing girls" phenomenon (Zeng et. al 1993; Das Gupta 2005; Ebenstein 2010).

[1] Sex ratio as measured in Sen (1992) is the ratio of female to male. So, a low sex ratio would indicate more number of men than women. In general, literature on sex ratio in India, by convention, uses the female to male ratio. In this remainder of this paper, I have adopted the same convention.

[2] The Total Fertility Rate (TFR) is the average number of children born to a woman if she lived to the end of her childbearing years and bear children in accordance with age-specific fertility rates of that specified year (World Bank Data).

In this paper, I aim to test the relationship between fertility rate and sex ratio in an Indian context. In particular, in a country with strong cultural preference for male children, I test the impact of changes in the family size on the proportion of male children in each family.[3]

This research is important not just from a demographic perspective, but also from social and economic perspective. The widening sex ratio causing an imbalance in the population is a hindrance for economic and human development. Studies have shown that widening sex ratio causes a squeeze in the marriage market, increases criminal behaviour, lowers labour force participation rate of women and impacts the overall rate of development of the country (Amuedo-Dorantes and Shoshana, 2007; Edlund et al, 2007; Jayachandran and Kuziemko, 2011; Wei and Zhang, 2011; Guilmoto, 2012; Agénor, Mares and Sorsa, 2015).

Demographic characteristics have been shown to influence criminal behaviour and incidence of crimes at both macro and micro levels within a country (Edlund et al, 2007; South, Trent and Bose, 2014). In China, Edlund et al (2007) found that a one percent change in the sex ratio increased violence and property crimes by 5-6%. Their research also suggests that nearly one seventh of the increase in the crime rate in China could be attributed to the rise in excess male population. Similar relation has been proven in India as well (Dreze and Khera, 2000; South, Trent and Bose, 2014; Hudson and Den Boer, 2002). In India, regions with greater imbalance in the sex ratio were associated with larger number of thefts, breaking and entering, assault, homicide rate and increased violence against women.

It has been suggested that increased number of marriages significantly reduces male criminality (Edlund et al, 2007). An imbalanced sex ratio resulting in a deficit of women particularly leads to a shortage of wives in the future. Guilmoto (2012) suggests that in India, between 2020 and 2080, there may be as many as 40 million single, unmarried men. Apart from increased criminality, unmarried men are also more likely to join military groups which could increase the incidence of domestic and regional violence (Hudson & Den Boer, 2002).

On the positive side, a shortage of girls could impact the bargaining power of women and thus raise the status of women. As the sex ratio declines, the number of females relative to male decreases causing a shortage in female supply. This shortage of females results in lower number of brides in the future. This imbalance causes more males to "fight" for brides thus improving the overall bargaining power of women. In the Chinese context, Wei and Zhang (2011) demonstrated that lower sex ratio forced parents with male children to accumulate more savings in order to compete for girls. This competition in the marriage market raises the economic status of women and could return the balance in the sex ratio in the long run.

On the economic front, poorer sex ratio has been associated with lower investment in women's education and lower female labour force participation rate. In a country with sex ratio biased towards men, women will anticipate better marriage prospects. Given the increased competition between men for brides, Wei and Zhang (2011) showed that parents invested larger amounts in

[3] Family Size refers to the total number of children per family.

male education compared to female's causing female employment prospects to be comparatively modest. Modest future earning further reduces their opportunity cost of working as many women would find it easier to be wives than work. In the United States, using individual data from Population Survey, Amuedo-Dorantes and Shoshana (2007) found that lower sex ratio reduces the likelihood of married women participating in the labour force. A lower sex ratio allowed women to perform worse than male counterparts. Furthermore, due to shortage of wives, women were more highly prized as wives and mothers. In contrast, if the sex ratio is higher, there will be a larger proportion of single women and they are more likely to work for pay than become wives. Such an effect of the sex ratio on labour force participation is significantly stronger for less educated women. Similar relation between sex ratio and female labour force participation was found in China as well (Butle, Tu and List, 2015). In general, lower female labour force participation has a huge impact on the overall economic development. It has been suggested that raising female labour force participation could boost economic growth by 2 percent over time (Agénor, Mares and Sorsa, 2015)

Son preference resulting in a skewed sex ratio and shortage of girls is detrimental to health and welfare of both women and men. It has been shown that in a country with strong son preference, daughters were less likely than sons to receive resources resulting in undernourishment, stunting, lower levels of education, lower parental time investment and excess infant mortality (Arnold, Choe and Roy, 1998; Jayaraj, 2009; Kugler and Kumar, 2011; Jayachandran and Kuziemko, 2011; Barcellos, Carvalho and Lleras-Muney, 2012). Furthermore, shortage of women leading to more single men, would also result in increased exposure of men to sex with prostitutes, thereby increasing the threat of sexually transmitted diseases, HIV and AIDS (Tucker et al, 2005). This would result in an increased expenditure in healthcare in the long run.

Given the medium and long run effect of worsening sex ratio, it is important to study the influencing factors to combat gender bias. Many of the existing literature studying the relation between family size and sex ratio focus on a visual inspection of data rather than a rigorous empirical analysis (for example Park and Cho (1995), Das Gupta Chung and Shuzhou (2009) and Mitra (2014)). The purpose of this paper is to fill this gap and quantify the impact of various factors including family size on the sex ratio. Computing this relation would allow for better forecasts and estimation of the consequences of different policy intervention.

Using women's data from the three rounds of the National Family Health Survey (NFHS) – NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06), I test whether the proportion of boys is greater in smaller families compared to larger families and how this relation has changed over time. However, it is important to note that family size is not an exogenous variable and is jointly determined by various factors including parents' education level and son preferences. The purpose of this empirical exercise is to quantifies the size and direction of the relation between the variables, thus enabling better policy focus areas to address gender bias.

My empirical analysis shows three main results. First, in line with my hypothesis, the proportion of boys significantly increases as the family size decreases. I find that a decrease in the family size by one person is associated with a 2.8 percentage point increase in the proportion of boys. This result is robust including controls for state fixed effect, round fixed effect and baseline characteristics. Second, the impact of a decrease in family size on the proportion of boys is larger for mothers with strong son preference compare to mothers without strong son preference. Finally, analysing the relation between family size and proportion of boys separately for NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06), I find that the effect is larger for NFHS III compared to NFHS I. This indicates that the gender bias has become more pronounced over time as fertility transitions to lower levels.

The rest of the paper is structured as follows. In Section 2, I present the background of the situation in Asia in general and India in particular. Section 3 presents an overview of the literature connecting fertility and gender bias. The purpose of the study, research questions and hypotheses are formulated in Section 4. Section 5 describes the National Family Health Survey (NFHS) Data used for analysis, variables of interest and the empirical approach. Results are presented in Section 6. Discussion of the results, recommended policy focus areas and concluding remarks are presented in the final section.

2. Background

In this section, I describe the trends in fertility and gender bias in Asian countries in general and India in particular.

2.1Gender Bias in Asia

The trends in the sex ratio observed in various countries is a complex consequence of cultural, demographic and economic characteristics. In the absence of any form of manipulation, the sex ratio at birth is commonly assumed to be between 105–107 male births for every 100 female births (Hesketh and Zhu, 2006). However, male infants are inherently more vulnerable and have lower resilience resulting in higher male infant mortality compared to females. These two effects would balance out over time and ensures a ratio of close to parity. However, among children in Asian countries, this ratio far exceeds the normal range. In 2005, this ratio among children in the 1-4 age group in China was peaking at 126 boys per 100 girls in rural areas with six provinces citing ratios of over 130 boys per 100 girls (Zhu, Lu and Hesketh, 2008). In South Korea, the sex ratio at birth peaked at almost 115 boys per 100 girls in the early 1990s and started to decline gradually since. This deviation from natural rates implies that the gap in the sex ratio is a result of sex selection and abortion rather than biological factors.

The main reason for the existence of gap in the sex ratio in Asia has been attributed to the preference for male children (Park and Cho, 1995; Ebenstein, 2010; Arnold, Kishor and Roy, 2002). Son preference is primarily due to higher perceived utility of boys. Traditionally, the males would work in agriculture, wage labour or family business to earn money while the females would be restricted to household chores. Furthermore, since the parents live with the male children as they grow old, having male children provides a form of insurance in their old age (Das Gupta, 1987). In a patriarchal society, women have very little or no property rights. Only the sons can inherit property and take over the family lineage. In Hindu religion, only the male heir is allowed to perform the last rites and light the parent's funeral pyre. In a country with over 80% of population strictly following Hinduism practices, this is an important consideration (Arnold, Choe, and Roy, 1998). Similarly, in China, for ancestral worship rituals, the sons play an indispensable role. In the case of India, cultural practice of dowry where the bride's family have to present lavish and expensive gifts to the groom's family at the time of the wedding increases the burden of having daughters. For these reasons, son preference usually takes the form of wanting at least one or two sons (Das Gupta et al, 2003; Jayachandran, 2014).

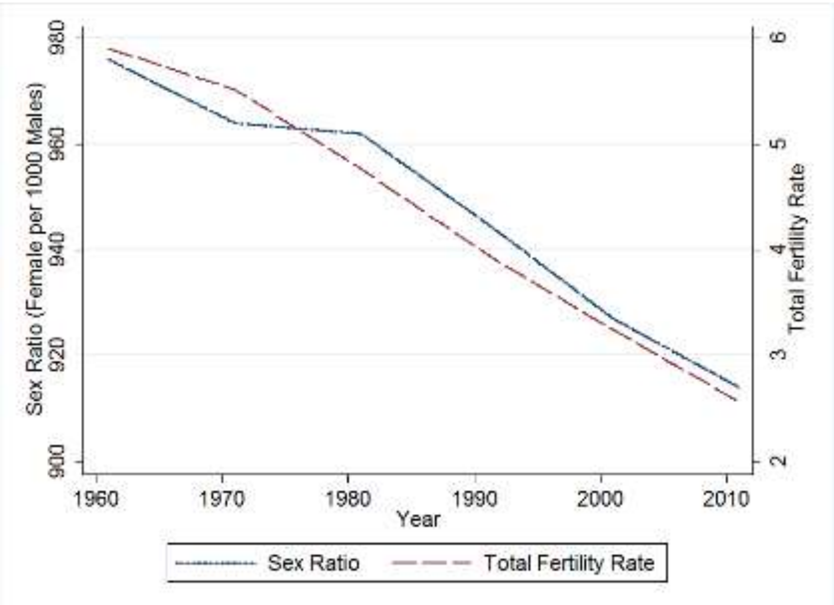
The period of worsening sex ratio also saw a dramatic decline in fertility rate. In China, following the introduction of the one child policy in 1978, fertility declined from around 4.5 in the early 1970s to 1.5 in 2000 (World Bank Data). In this same period, the sex ratio at birth changed from 0.934 girls per boy in 1975 to 0.833 girls per boys in 2000 (Guilmoto, 2009). In South Korea, the fertility rate declined from over 4.5 children in 1970 to around 1.6 in the mid-1990s while the sex ratio worsened from around 0.93 girls per boys 1970s to 0.869 girls per boys in 1995 (Guilmoto, 2009).

Previous researchers had suggested that parental preference for sex of their children would be a barrier for fertility reduction (Amin and Mariam, 1987; Clark, 2000; Basu, 2010). That is, if parents continue to bear children till they have at least one son, this sex preference would result in a higher fertility rate. However, with improvements in medical technology, it has become easier to determine the sex of child using an ultrasound. Thus, parents have been able to ensure preferred sex composition within the smaller desired family size. This has caused the gender bias to worsen as the family size decreases.

2.2 Gender Bias in India

Over the past few decades, India has undergone significant economic development with an average growth rate of over 5% over the past 3 decades. Despite huge strides in education, health and other development indicators, the sex ratio has worsened. Economic theory predicts that growth and development would close the gap and reduce gender bias in the country. However, this has not been the case in India. Similar to other Asian countries, cultural preference for male children is strong in India as well (Das, 1987, Mitra, 2014). This indicates that a declining fertility rate coupled with son preference could explain the widening gap in the sex ratio. A first look at the relation between total fertility rate and the child sex ratio in India is shown in Figure 1.1.

Figure 1.1: Child Sex Ratio and Total Fertility Rate in India, 1961-2011

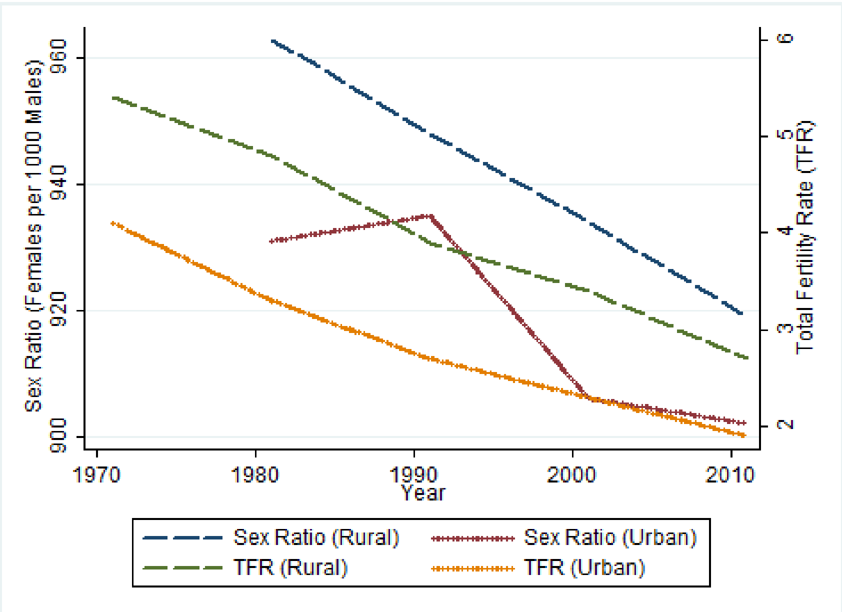


Note: The child sex ratio here refers to the sex ratio among children aged 0-6.
Data Source: India Census Data and World Bank Data

This strong co-movement between the two variables shown in the figure provides motivation for further investigation. From 1961 to 2011, the total fertility rate declined from 5.9 to 2.56 while the child sex ratio worsened from 0.976 to 0.914 in the same period. The fertility decline over this period has been attributed to the demographic transition due to improvements in economic prospects and changing economic value of children (Gubhaju, 2007).

Comparing the sex ratio and fertility by region, I find that large difference exists between rural and urban regions. Despite the differences between regions, analogous declining trend in sex ratio and fertility rate is visible in both rural and urban areas (Figure 1.2). Rural areas are those parts of the country which are not industrialized, such as agricultural areas located outside towns and cities. Urban areas on the other hand are highly urbanized and industrialized such as cities and larger towns. Similar difference between regions (rural versus urban) in the fertility and sex ratio has been observed in China as well (Hesketh, Li and Zhu 2005; Ding and Hesketh, 2006).

Figure 1.2: Child Sex Ratio and Total Fertility Rate by Region, India 1971-2011

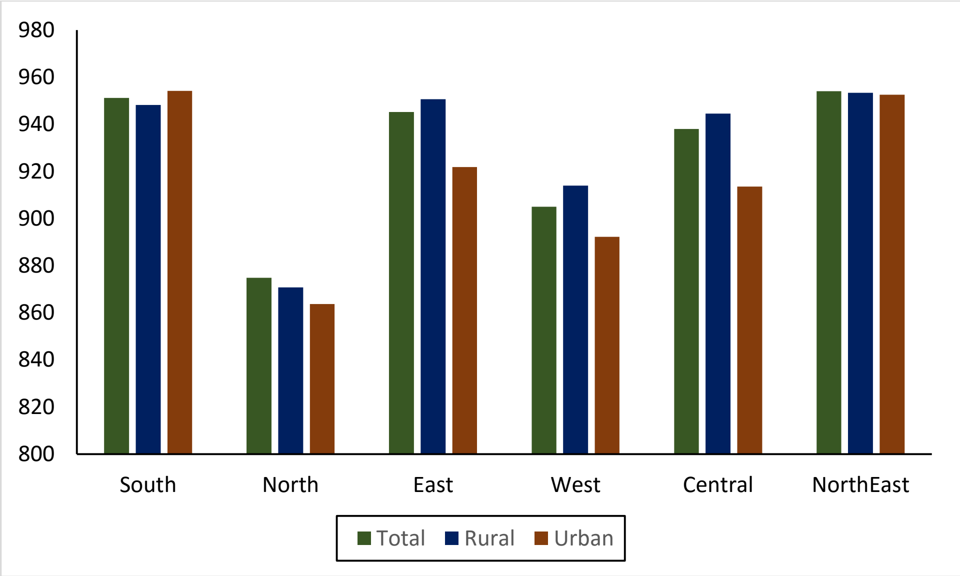


Note: Data on child sex ratio by region is available only from Census 1981; Data Source: India Census Data and Ministry of Health and Family Welfare

In Figure 1.2 The dashed line shows the trend in rural regions and the dotted lines shows the trend in urban regions. From 1971 to 2011, the total fertility rate decline from 5.4 to 2.7 children per women in rural region and from 4.1 to 1.9 in urban region. From 1981 to 2001, the sex ratio in rural areas deteriorated from 0.963 to 0.919 compared to 0.931 to 0.902 in urban areas. While both the trends shown downward movement, it is interesting to note that the sex ratio is worse in urban areas than rural areas. This might be the result of lower fertility rate and easier access to medical technology for sex selection.

There is also considerable heterogeneity across different states and regions as well. India is divided into 29 states and 7 union territories which differ vastly in terms of economics, culture and language. The child sex ratio in different regions (South, North, East, West, Central and North-East) in India based on the 2011 Census is shown in Figure 1.3.

Figure 1.3: Child Sex Ratio by Region in India, 2011



Note: Child sex ratio refers to the number of females per 1000 males among children aged 0-6; Source: Census 2011, author's own calculation

As seen in the figure, there is considerable regional difference in the child sex ratio in India. The southern, eastern and north-eastern states have a better child sex ratio compared to west and north India. North India, where son preference has been known to be the strongest, has the lowest sex ratio of 875 girls per 1000 boys. Even in North-East regions which has the highest child sex ratio of 954 girls per 1000 boys, this level is still below parity.

Apart from the regional differences in sex ratio, there is considerable diversity in terms of the extent of son preference, education and income not just between states in the same region, but also between religions groups and sects within each state. However, analysing the effects by state could be confounded by internal migration between regions and state within India. Furthermore, given that most of the policy discussions on sex ratio has been focused at a national level rather than regional level, in this paper, I focus on the national aggregate relation between family size and sex ratio controlling for state and region fixed effects.

3. Literature Review

The increasing masculinity in the population in Asian countries has raised international concern. Given the prominence of this situation and severe social and economic consequences, many literatures have focused on understanding the causes for the widening sex ratio and gender bias. However, the reasons for increasing gender bias is not as straight forward. The bias is a complex interplay of various factors including son preference, fertility, reduced infant mortality and improvements in medical technology for sex determination (Park and Cho, 1995; Das Gupta and Bhat, 1997; Jha et al, 2006). Coupled with son preference, fertility decline has been cited as a causal reason for the increased manifestation of gender bias (Bhat and Zavier, 2003; Clark, 2010; Ebenstein, 2010; Jayachandran, 2014).

Fertility decline in Asian countries have been due to two main reasons. One, demographic transition due to rapid industrialization and changing economic value of children as observed in South Korea, Hong Kong and Singapore (Gubhaju, 2007). The second reason for fertility decline is due to policy enforcement like in the case of the one child policy in China (Ebenstein, 2010). In the case of India, this decline has been due to a combination of both factors – demographic transition as well as family planning programs launched by the government. As these countries develop, infant mortality reduces, education and healthcare improve, economic prospects increase and considerable advancements in technology are made. All these factors in turn affected the fertility rate.

In terms of the relation between declining fertility and increased gender bias, it is not immediately clear how this relation would differ in a society with strong preference for males. Das Gupta and Bhat (1997) argued that there were two opposing effects at play - the 'parity effect' which would reduce the gender bias in the sex ratio and the 'intensification effect' which would increase it. In South Asia, it has been shown that excess mortality of girls was larger for women with more number of pregnancies (Das Gupta, 1987). Therefore, as fertility declines, the excess mortality of girls would fall as well. This was the parity effect. According to intensification effect, based on observations in South Korea and China, the fertility rate declined at a much faster rate than the preference for sons (Zeng et al, 1993; Park and Cho, 1995). Therefore, the gender bias becomes more pronounced as the family size decreases. Similar intensification effect has been observed in India as well. Das Gupta and Bhat (1997) conclude that in the case of India, the intensification effect outstripped the parity effect resulting in a net decrease in the sex ratio as fertility declines. Nevertheless, the parity effect is an important consideration since the sex ratio might have worsened even more sharply than the observed change.

The increased availability of medical technology has made it possible for couples to accommodate desired preference within smaller family sizes. Suppose the desired sex composition takes the form of wanting at least one son. The probability of having at least one son is 99% if a couple has six children. This probability falls to 76% if the couple has only two children (Jayachandran, 2014).[4] Given that the possibility of achieving the desired sex composition is lower as the family size decreases, there is a higher likelihood that the couple would indulge in sex selection through abortion or infanticide. This would also increase the gender bias as the family size decreases.

In these countries, accommodating the desire for smaller families and sex preference, distortion of the sex ratio emerges at three levels: in the overall population, between families and within families (Park, 1983). Sex selective practices at the individual family level alters the ratio not only at the within family level, but also between families and at a national level. Wen (1993) compared two provinces in China using the 1985 Fertility Survey and showed that sex ratio of siblings consistently improved with family size and abruptly changed from 130 boys per 100 girls to 108 boys per 100 girls when the family size reached four children. Even in South Korean data, there seems to be a threshold of family size below which the sex ratio is extremely low and above which it disappears. The association shows that the sex ratio is less than 0.50 females per male when the family size was one and is over 2 females per male when the family size was five (Park and Cho, 1995). These ratios at a family level aggregates to the national level and results in increased manifestation of gender bias as the fertility declines.

In China, it has been shown that the introduction of the one child policy, which caused a significant drop in fertility since its introduction 1978, could explain a large part of the worsening sex ratio. Using a difference-in-difference estimate, Li, Yi and Zhang (2011) showed that the one child policy could explain 57% of the increase in the sex ratio for the 1991-2000 birth cohort and 54% of the increase in the 2001-2005 birth cohort. In the case of India, fertility decline is shown to explain 30-50% of the declining sex ratio in India (Jayachandran, 2014). Exploiting the regional variation in the fines for having additional children in China, Ebenstein (2010) found that the areas with stricter enforcement of the one child policy and larger fines were associated with more distortion in the sex ratio compared to other regions. He showed that the fraction of male among first births rose from around 0.512 in 1975 to over 0.525 in 2000.

Similar to the case in China, many studies have observed the relation between fertility and gender bias in India as well. Das Gupta and Bhat (1997) observed this relation at all-India level. In line with other papers, they found that fertility decline has been accompanied by increased in the sex ratio. However, the rate of increase in the sex ratio differed across regions with North India exhibiting largest increase. They further showed that increased manifestation of the bias was greatest where previous bias was already large. Using data on the ideal family size and desired sex composition, Bhat and Zavier (2003) showed that the desired proportion of boys was 1.3-1.4 percentage points higher when the ideal family size decreased by one person. This effect was significant, controlling for various parental characteristics such as age, education, religion and location of residence.

The concomitant trends in fertility and sex ratio shown in Figure 1.1 coupled with various research suggest that there could be a causal link between the two variables. Although there is a vast literature on fertility and sex ratio, rigorous empirical evidence of the effect is limited

[4] The natural sex ratio is slightly skewed towards boys resulting in the probability being more than 75%.

indicating a need for further research. In the next section, I outline the purpose of the thesis and introduce my research questions which I break into three testable hypotheses.

4. Purpose and Research Questions

Many of the existing literature connecting family size and the sex ratio focus on quantitative observations rather an empirical analysis. While significant negative relation between the declining family size and widening gap in the sex ratio has been observed in various countries by researchers, the size of the impact is not estimated (for example Park and Cho (1995), Das Gupta and Bhat (1997), Das Gupta Chung and Shuzhou (2009) and Mitra (2014)). Few papers which assess the impact for India through empirical analysis (such as Bhat and Zavier (2003) and Jayachandran (2014)), focus on ideal family size and the desired sex composition rather than the realized outcome. Clark (2000) assessed the relation between the actual family size and realized sex composition using data from NFHS I (1992-93) finds significant negative relations.

Specifically, for the Indian context, there is limited literature analysing factors affecting the realized sex ratio and in particular the relation between family size and the sex ratio. My research differs from existing papers in two ways. First, I focus on the actual family size and the realized sex composition of children rather than ideal composition. Second, I use data from all three rounds of the NFHS (1992-93, 1998-99 and 2005-06) to test the effect within and across time periods. While Clark (2000) assessed the relation using realized family size and sex composition, her research focused primarily on data from NFHS I (1992-93). On the other hand, while Jayachandran (2014) used data from all three rounds of the NFHS, she quantified the relation between ideal family size and desired proportion of boys rather than the realized outcomes.

Focusing on the relation between actual family size and realized proportion of boys does pose an endogeneity threat. Given the interrelation of the key variables with various socio-economic factors, the relation established in this paper is not completely causal. However, this empirical exercise quantifying the size and direction of this relation, allows for a better understanding of the various factors influencing the sex ratio thus providing clear policy focus areas to address gender bias.

In this paper, I intend to study the relation between the family size and the proportion of boys between families and across time. More specifically, I intend to study the following questions:

1. *In India, what happens to the proportion of boys in each family as the family size decreases?*

2. *In India, can the decrease in family size explain the worsening sex ratio over time?*

It should be noted that these two questions are inter-related. The first question compares the relation between families and the second question compares this relation across time. The two research questions can be broken down into the following hypothesis:

Hypothesis 1: Parents' son preference plays a significant role in influencing the total proportion of boys

This hypothesis forms the basis for the analysis since the main reason for the existence of gender bias in the sex ratio in India is due to strong son preference (Clark, 2000; Arnold, Choe and Roy, 1998; Das Gupta et al, 2003). Thus, I expect a strong positive relation between parents' son preference and the proportion of boys in each family. Furthermore, I expect that the effect of family size on proportion of boys would vary depending on the level of son preference.

Hypothesis 2: As the family size decreases, the proportion of boys is expected to increase

When family size decreases and couples have fewer children, this puts pressure to have the desired number of male children within the smaller families. This causes families to resort to sex selection in order to achieve the desired sex composition within their desired family size. Therefore, I expect a negative relation between the family size and the proportion of boys.

Hypothesis 3: Given that the average family size has decreased over time, the relation between family size and proportion of boys is expected to be stronger for NFHS III (2005-06) than NFHS I (1992-93)

Due to demographic transition, the average family size in India has declined over time (as shown in Figure 1.2). Furthermore, the composition of families by sizes differs between NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06). That is, the proportion of families with 1, 2, 3... children in 1992-93 differs from proportion of families in 2005-06. In line with the declining average family sizes, I expect that the proportion of families with 1 or 2 children to be larger in 2005-06 than in 1992-93. Assuming that hypothesis 2 holds, I expect to see a larger negative relation in 2005-06 compared to 1998-99 and 1992-93.

5. Data and Empirical Approach

In this section, I explain the data and the regression model used to test the hypothesis mentioned in Section 4. First, the National Family Health Survey data is presented followed by an explanation of the variables used for my analysis. Second, I explain the Ordinary Least Square (OLS) model used to test the relation. Finally, I introduce the identification assumption necessary to establish causal relation and discuss potential threats to identification.

5.1 Data – National Family Health Survey (NFHS)

The National Family Health Survey (NFHS) is a multi-round, large scale survey conducted to obtain information on fertility, mortality, health, nutrition, reproductive health, family planning and other demographic indications. The survey is conducted throughout India in a representative sample of households. Three rounds of the survey have been conducted till date. The first round NFHS I was conducted in 1992-93; the second round NFHS II in 1998-99; the third round NFHS III in 2004-05. The fourth round of the NFHS survey is currently being conducted. The main objective of the NFHS has been (i) to provide essential data to the Ministry of Health and Family Welfare and other related agencies on health and family welfare for programs and policy purposes, and (ii) to provide information on emerging issues on family welfare and health (IIPS). The three rounds of NFHS data was acquired from the DHS Program Data on India.

All the rounds of the NFHS have been completed under the stewardship of the Ministry of Health and Family Welfare, India. The survey uses uniform questionnaires (translated into different Indian languages) and field procedures throughout the country in order to facilitate comparability across the states and ensure highest possible data quality. This survey is an important data source on population, health, and nutrition for India and its states.

Apart from survey data on household and child indicators, the NFHS Data also provides information about ever-married women aged 15-49 on various health and demographic dimensions. The survey also solicits information on fertility preferences such as ideal family size and sex composition of children in their ideal family size. Data was available for over 89,000 women across 25 states in NFHS I (1992-93); 33,206 women across 26 states in NFHS II (1998-99); and over 39,000 women across all 29 states in NFHS III (2005-06).

Compared to NFHS I (1992-93) and NFHS II (1998-99), NFHS III (2005-06) additionally provides information on several new and emerging issues such as male involvement in the family, family life education, welfare and health condition of slum dwellers etc. For my analysis, data on father's fertility preferences available in NFHS III (2005-06) provides additional insight into the relationship between son preference, family size and proportion of boys. Data on wealth level of the respondent is incorporated with the main data only for NFHS II (1998-99) and NFHS III (2005-06).

There are three main data requirements from the survey in order to test my hypotheses. First, data on the family size (total number of children) and sex composition of children are needed. Secondly, I need information on parent's ideal family size and sex composition in order to determine the son preference. Lastly, I need information on basic demographic characteristics of parents in order to control for baseline features.

In order to conduct my analysis, I trim the data in the following ways. To start with, I restrict our data sample only to women aged 35-49. This is because younger women are more likely to continue bearing children since they would not have reached their desired fertility level. Following Park and Cho (1995), I assume that women aged 35-49 have completed child bearing. The age specific fertility rate (ASFR) in 5 year intervals for women in NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06) is shown in Appendix 9.1.[5] Analyzing the ASFR in Appendix 9.1, I find that the births per 1000 women considerable decline for women over the age of 35 indicating that my assumption would hold.

Secondly, I only consider women who have at least one child. This is an obvious restriction since I am analyzing the proportion of boys in each family. Finally, I omit data with missing variables. The final sample consists of 11357 respondents from NFHS I (1992-93), 29091 respondents from NFHS II (1998-99) and 14274 respondents from NFHS III (2005-06).

It is important to note that since the data is restricted to women aged 35-49, this refers to women born as early as 1942. The youngest women in the survey would have been born in 1971. Thus, it is likely that the proportion of boys, family size and son preference trend captured in this restricted data would lag slightly behind current trends. However, since the trend in fertility and sex ratio has been declining for a few decades now, the analysis of this paper can also be used to explain current trends.

5.2 Variables of Interest

Proportion of boys

The main dependent variable for my analysis is the actual proportion of boys (*PropofBoys*) in each family. Similar to Bhat and Zavier (2003) and Jayachandran (2014), the proportion of boys is simply the proportion of males in the total number of children.

$$PropofBoys = \frac{\text{number of male children}}{\text{total number of children}}$$

[5] The Age Specific Fertility Rate (ASFR) is simply the total number of births in a given year per every 1,000 women of reproductive age classified in single or five-year age groups.

The proportion of boys is an alternative way for measuring the sex ratio. Instead of comparing the number of female to male children, I simply measure the proportion of boys in the total number of children. The proportion of boys is used instead of the female to male ratio because in a within-family comparison, it is likely that some households may have zero male children. In this case the female-to-male ratio would tend to infinity. Similarly, in households with zero girls, this ratio would be zero. While at a national level, the sex ratio might be better indicator than the proportion of male, at an individual family level, the proportion of boys is preferred.

Family Size

Family size (*FamSize*) is simply the total number of children per woman. In order to construct the family size, I use the total number of children rather than the total number of births. The main difference is that the total number of children refers to the number of children born to a mother that are alive compared to the total number of births which also includes children that were born but died due to various reasons. The total number of children used in my analysis refers to the number of children alive at the time of the survey.

As mentioned before, male children are inherently more vulnerable to infant mortality compared to females. However, due to preference for son, incidents of female infanticide are greater. Furthermore, the death of a child, especially if it is a male child, is likely to induce parents to have more children in the future in order to achieve their desired sex composition. This might cause a potential bias in measuring family size by the total number of births. Since the exact direction of bias is uncertain, it would be difficult to control for it. For this reason, I focus on the total number of children which are alive rather than the total number of births.

Son Preference

Son preference is a key variable since the practices of sex selection in Asian countries is primarily due to the preference for male children over females. The data from the NFHS solicits information on mother's fertility preference. They ask the respondent to state their ideal family size and the desired sex composition of children. To measure the desired sex composition, the respondents are asked to state desired sex as male, female or either/neutral preference.

In line with the index created by Bhat and Zavier (2003), I state that son preference is strong if the ideal proportion of boys is strictly greater than the ideal proportion of girls. In other words, if in their ideal family size, the women report preference for more boys than girls, then son preference exists. For example, if a respondent states their ideal family size to be 2 and they prefer one child to be a son and is indifferent about the gender of the other child, then I have classified this as son preference. In this case, the ideal proportion of boys is 0.5 and the ideal proportion of girls is 0. However, if the parents had stated that they would prefer the gender of the second child to be a girl, then I have not classified that as son preference. Simple put:

$$SonPreference = \begin{cases} 1, & if\ ideal\ proportion\ of\ boys > ideal\ proportion\ of\ girls \\ 0, & otherwise \end{cases}$$

The data on mother's fertility preference is available for NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06). This enables us to construct mother's son preference (*mSonpreference*) for all three rounds. However, father's preference (*fSonpreference*) is available only for NFHS III (2005-06).

Other explanatory variables

To start with, I include the baseline characteristics of the mothers – mother's age (*mAge*), mother's education (*mEduc*) and mother's working status (*mWork*). By introducing mother's age as an explanatory variable, I intend to capture the cohort effect for son preference. That is, I expect that younger women will have a weaker preference for sons than older women. Mother's education is a crucial explanatory variable. However, the exact direction of impact of mother's education on preference for sons is uncertain. On one hand, improved access to education would decrease the preference for sons because of increased exposure and better self-earning prospects (Bhat and Zavier, 2003). On the other hand, highly educated individuals are more likely to be able to act on their desires for sons through increased access of medical technology for sex selection thus worsening the sex ratio (Bhat and Zavier, 2003; Clark, 2000). Overall, I expect mother's education to have a negative effect on the proportion of sons. I also include a binary variable to capture the working status of women categorized as 1 if the individual is working and 0 otherwise.

Religion also plays a role in son preference, women's status and the resulting sex composition in families. The practice of dowry, common among Hindus has often been cited as a key factor in parents' preference for sons (Arnold et al., 1998). Among Muslims cultures, the status of women is generally considered to be lower. Given the difference in preference due to religious beliefs, I include religion as a variable to control for the heterogeneity across religions.

Geography plays a key role in shaping the social and cultural environment. Extensive literatures have found large regional variance in fertility rate, sex ratio and son preference across regions in India (Das Gupta and Bhat, 1997; Clark, 2000). Table 9.2 in the Appendix shows the proportion of boys and family size in different states by round. From 1992-93 to 2005-06, the average family size for women aged 35-49 decreased for all states. However, the extent of decline varies. Manipur, a state in the North East, saw a decline in the average family size from 5.73 to 3.59 over the three rounds while Andhra Pradesh declined from 4.2 to 3 children per family. In terms of sex ratio over the three rounds, some states such as Punjab saw a dramatic increase in the average proportion of boys from 0.50 to 0.58; some state such as Manipur saw a moderate increase in the proportion of boys from 0.512 to 0.528; and one state, Meghalaya actually saw a slight decline from 0.531 to 0.529. Given the considerable heterogeneity across states, I introduce state fixed effects to control for this variation.

Economic factors also play an important role in determining the proportion of boys in a family. Since a crucial reason for son preference in India is to provide security to the parents in their old age, I introduce the wealth level (*wealthlevel*) of the family as an explanatory variable. Furthermore, with cultural practices such as dowry, the burden of a female child is heavier on poorer families. However, increase in wealth also improves access to medical technology for sex selection which might result in higher proportion of boys in wealthier families. The net effect of wealth level on the proportion of boys is uncertain. The data on wealth level is incorporated in the main data only for NFHS II (1998-99) and NFHS III (2005-06). The wealth level measured is categorized into 5 levels – poorest, poor, middle, rich, richest.

Other measures included in the multivariate analyses are father's age (*fAge*), father's education (*fEduc*) and a dummy variable for residence (which is equal to 1 if the individual resides in a rural area).

Summary

The table below presenting the mean and standard deviation of the all the variables used in the analysis.

Table 5.1: Descriptive Statistics of Variable

	NFHS I (1992-93)		NFHS II (1998-99)		NFHS III (2005-06)	
	Mean	SD	Mean	SD	Mean	SD
Proportion of Boys	0.49199	-0.22061	0.54172	-0.27808	0.53968	-0.29542
Family Size	4.78216	-1.74933	3.74635	-1.72439	3.35680	-1.64918
Mother's Age	42.69437	-4.04686	40.83803	-4.20534	40.23897	-3.98463
Mother's education level (in years)	1.74562	-3.21700	3.60215	-4.71295	4.81435	-5.27716
Proportion of Mother's working	0.35837	-0.47954	0.38166	-0.48580	0.43996	-0.49640
Proportion of Mother's with strong son preference	0.39209	-0.48824	0.32904	-0.46987	0.23168	-0.42192
Father's age	49.58449	-6.54476	46.98807	-7.13309	45.29599	-5.69137
Father's education level (in years)	4.53773	-5.08703	6.35372	-5.19715	7.70583	-8.75160
Proportion of Father's with strong son preference					0.24471	-0.42993
Rural Population Proportion	0.74588	-0.43538	0.64986	-0.47702	0.51289	-0.49985
Religion by Proportion						
Hindu	0.80444	-0.39665	0.77845	-0.41529	0.75284	-0.43138
Muslim	0.10998	-0.31287	0.10924	-0.31195	0.10397	-0.30523
Christian	0.04182	-0.20020	0.06143	-0.24012	0.09220	-0.28931
N	11357		29091		14274	

Note: The summary statistics are calculated based on the data on women aged 35-49 in each round of the NFHS

The proportion of boys increased dramatically from 0.491 in 1992-93 to 0.541 in 1998-99 and marginally reduced to 0.539 in 2005-06. As expected, the number of children per women aged 35-49 decreased over time from 4.78 children in 1992-93 to 3.35 children in 2005-06. The decrease in family size is a result of demographic transition observed when countries transition from pre-industrial to an industrialized system. In line with the transition and economic development of the country, the average years of education of mothers and fathers have improved. The data indicates that mother's education has more than doubled over the three rounds from an average of 1.7 years in 1992-93 to 4.81 years in 2005-06. Along with improvements in education, I also see a greater proportion of mothers who are working. Nearly 44% of the surveyed mothers were working in 2005-06 compared to 35.8% in 1992-93. The improvements in education coupled with higher proportion of working mothers could also lead to the decline in the average family size.

It is also interesting to note that over the three rounds, the proportion of mothers with strong son preference has nearly halved from nearly 40% of them preferring sons in 1992-93 to just 23% with strong son preference in 2005-06. This could be a result of improvements in education and job prospects of women. Data on father's preference available only for NFHS III (2005-06) indicates that over 24% of the fathers have a strong son preference. Since it is possible that the stated preference might differ from the actual preference, caution needs to be exercised when interpreting the results.

The data also shows the proportion of population which are Hindus, Muslim and Christian. However, the proportion indicated in the sample differs from the religious composition in the actual population. According to 2001 Census, of the total population, 80.5% were Hindus, 13.4% were Muslims, 2.4% were Christians and the remaining belonged to other religions such as Sikhs, Jains and Buddhists. This difference between the actual population and surveyed data could be a result of sampling and data trimming.

5.3 Empirical Approach

In order to empirically test the hypotheses formulated in section 4, two OLS models are constructed. The two models are identical in set up but differ whether the regression is run by pooling all three rounds of the NFHS (equation 5.1) or by running each round separately (equation 5.2).

The first baseline specification is:

$$PropofBoys_{isr} = \alpha + \beta_1 * FamSize_{isr} + \beta_2 X_{isr} + v_s + u_r + \varepsilon_{isr} \quad (5.1)$$

Where $PropofBoys_{isr}$ is the proportion of boys for family i, in state s, at round r. The main independent variable of interest is $FamSize_{isr}$ which is the total number of children in each family. As mentioned before, the total number of children refers to the number of children that are alive rather than the total number of births. X_{isr} is a vector of baseline covariates which includes mother's age (*mAge*), mother's education level (*mEduc*), mother's working status (*mWork*), father's age (*fAge*), father's education level (*fEduc*), region (rural dummy variable) and religion. I control for state fixed effects (v_s) in order to account for unobservable state-level factors which does not change over time. I also control for survey round fixed effects (u_r) in order to account for the unobservable panel-invariant effects for each round of the NFHS.

I run the first baseline specification (equation 5.1) in three ways. First, is a simple regression slowly including additional controls to quantify the effect of family size on the proportion of boys. Second, I run the regression for different family sizes in order to test for effect at different levels of family size. In both these specification, I include a separate control for son preference (*mSonpreference*) as well. Finally, I run this regression separately for mothers with and without strong son preference.

The second baseline specification is:

$$PropofBoys_{is} = \alpha + \beta_1 * FamSize_{is} + \beta_2 X_{is} + v_s + \varepsilon_{is} \quad (5.2)$$

The second baseline specification, which is run individually for NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06) is identical to equation 5.1 except that the round effects are removed. Similar to the first baseline specification, the coefficient of interest is β_1 which measures the effect of *FamSize* on *PropofBoys*. The list of baseline characteristics included slightly differs across the three rounds since, additional data is available for NFHS II (1998-99) and NFHS III (2005-06). The explanatory variables included in all three rounds are mother's age (*mAge*), mother's education level (*mEduc*), mother's working status (*mWork*), mother's son preference (*mSonpreference*), father's age (*fAge*), father's education level (*fEduc*), region (rural dummy), religion and state fixed effects. Additionally, I include the explanatory variable wealth level (*wealthlevel*) for NFHS II (1998-99) and NFHS III (2005-06) and father's son preference (*fSonpreference*) only for NFHS III (2005-06).

5.4 Identification and Threats to Identification

In order to establish a causal relationship between the family size and the proportion of boys in equation (5.1) and (5.2), the identification assumption is that the total number of children are conditionally exogenous to the sex ratio and other determinants of the sex composition of children. In other words, the family size is determined exogenously and is not affected by the dependent variable and other explanatory variables.

There are two main threats to this assumption. The first is that the family size might be endogenous and thus the proportion of boys and family size might be jointly determined by another explanatory variable. The second threat is the possibility of reverse causality due to differential stopping behaviour. In this subsection, I discuss these two threats and propose strategy to deal with them.

Endogeneity and joint determination

There are multiple social and economic factors influencing both the family size and the proportion of boys. Furthermore, the outcomes of both the variables are jointly determined by the parents. While the sex composition of children is less under the parents' control, possibility to sex selection due to availability of medical technology, this raises parent's influence in controlling the gender outcome of the children.

One of the key parental characteristics which influences both the fertility decision and sex composition is mother's education level. Education, especially female education has been cited as key factor in reducing gender inequality (Clark, 2000; Duflo, 2011; Jayachandran, 2014). As mentioned in Section 5.2, improvements in mother's education has two opposing effects on the sex ratio. First, improved access to education would decrease the preference for sons because of increased exposure and better self-earning prospects (Bhat and Zavier, 2003). Second, highly education individuals are more likely to be able to act on their desires for sons through increased access of medical technology for sex selection thus worsening the sex ratio (Bhat and Zavier, 2003; Clark, 2000).

In terms of impact on fertility rate, it has been shown that mother's education decreases the fertility rate. This is because, improved education raises employment prospects of women. This in turn affects the opportunity cost of raising children. This causes parents to make a trade-off between the quantity and quality of children as economic prospects improve (Becker and Lewis, 1973). Thus in general, increased levels of education lead to lower fertility rates. For this reason, I introduce an interaction variable between mother's education and family size (*FamSize*mEduc*).

Another important factor parental characteristic which affects both the number of children as well as the sex composition is son preference. Strong son preference increases the likelihood of parents manipulating the sex of their children in smaller families (Park and Cho, 1995). Furthermore, in families not indulging in sex selection due to moral or financial reasons, if the first born is female and the desire for a son is strong, then couples are likely to continue child bearing till they have at least one son. As a result, girls will tend to grow up in larger families with more siblings and boys in smaller families with fewer siblings (Basu and de Jong, 2010). For this reason, I introduce another interaction variable between mothers' son preference and family size (*FamSize*mSonPreference*)

The modified regression model is presented below:

$$PropofBoys_{isr} = \alpha + \beta_1 * FamSize_{isr} + \beta_2 X_{isr} + \beta_3 * FamSize_{isr} * mEduc_{isr} + \beta_4 * FamSize_{isr} * mSonpreference_{isr} + v_s + u_r + \varepsilon_{isr} \quad (5.3)$$

The above model is similar to the model presented in equation 5.1 with the addition of interaction variables. Similar to equation 5.1, I control for state fixed effects, round fixed effects and baseline characteristics. In equation 5.3, despite the joint determination of family size and the proportion of boys by various factors, I include family size as an explanatory variable since other socioeconomic variables and interactions are unlikely to capture all the factors responsible for fertility decisions. That is, the difference in family size could be the result of decreased mortality, changing preference, demographic transition or other factors which are independent of mother's education and son preference.

Reverse Causality

Another important link between family size and sex ratio is the birth order of the child. If the first born is female and the desire for a son is strong, then couples are likely to continue child bearing till they have at least one son. If they reach their desired composition within the first child, they are more likely to stop bearing more children. The impact of this differential stopping behaviour (DSB) on the family size is well documented in literature in South Asia and Korea (Park, 1983; Clark, 2000, Chaudhuri, 2012). Using data on Indian women from NFHS III (2005-06), Chaudhuri (2012) showed that at any given family size, the last-born child of women who had stopped bearing children was more likely to be a son than a daughter.

This differential stopping behaviour alters the causal relation between proportion of sons and family size. That is, it is possible that the proportion of boys could causally affect the family size rather than the other way around. In essence, the differential stopping behaviour would result in a high proportion of boys in families which have male children early on and a lower proportion in families which have male children later (Chaudhuri, 2012). In this case, the estimate for causal relations using OLS might be biased.

For this reason, I use the two stage least square (2SLS) method. The instrument used to control for the endogeneity is the sex of the first child. A necessary condition for the sex of the first child to be a valid instrument is that the family size must be highly correlated with the sex of the first born. In countries with son preference, research has shown that the sex of the first child is an important source of exogenous variation in the family size (Clark, 2000; Basu, 2010). It is possible that the use of medical technology for sex determination and sex selection would undermine the validity of the instrument. However, ultra sound and sex determination has been illegal in India since 1994, and access to this medical technology is not as widespread across the country. Furthermore, Retherford and Roy (2003) found little or no evidence of sex selection on first birth in India using NFHS I (1992-93) and NFHS II (1998-99). This suggests that the sex of the first child can be considered as random and therefore exogenous.

The sex of the first child as an instrument for family size has been used in other studies as well (such as Lee (2008) and Kumar and Kugler (2010)). The 2SLS model is presented below:

First Stage:

$$FamSize_{isr} = \delta_0 + \delta_1 * Male1_{isr} + \delta_2 X_{isr} + v_s + u_r + \epsilon_{isr} \quad (5.4)$$

Second Stage:

$$PropofBoys_{isr} = \alpha + \beta_1 * \widehat{FamSize}_{isr} + \beta_2 X_{isr} + v_s + u_r + \varepsilon_{isr} \quad (5.5)$$

In the first stage, the family size is regressed on all exogenous variables including the instrumental variable controlling for state and round fixed effects. The instrument, which is the sex of the first child (*Male1*) is a dummy variable which equals 1 if the first child is male. In the second stage, I regress the proportion of boys on the estimated family size derived from the first stage as well as baseline characteristics controlling for fixed effects.

Using interactions model and 2SLS, I try to control for the endogeneity to establish a causal relation. However, given that the interconnection and interdependence of the key variables with various socio-economic factors, it is difficult to completely eliminate endogeneity to establish causality.

6. Results

In this section, I present the results of empirical strategy outlined in the previous section. As mentioned before, it is important to keep in mind that this data is restricted to women aged 35-49 and therefore, the results capture the trend for women born between 1942 and 1971. To start with, Table 6.1 shows the proportion of boys and the corresponding sex ratio by family size for each round of the NFHS.

Table 6.1: Proportion of Boys by Family size and round

	Family Size	1	2	3	4	5	6	7+
NFHS I	PropofBoys	-	0.517024	0.52579	0.5026	0.475862	0.469715	0.469335
1992-93	Sex Ratio	-	0.934146	0.901901	0.989653	1.101449	1.128951	1.130674
NFHS II	PropofBoys	0.575139	0.604532	0.569851	0.520008	0.489415	0.47617	0.478466
1998-99	Sex Ratio	0.73871	0.654173	0.754846	0.923047	1.043257	1.10009	1.090014
NFHS III	PropofBoys	0.583333	0.590317	0.547224	0.505926	0.475352	0.467702	0.478468
2005-06	Sex Ratio	0.714286	0.694006	0.827404	0.976575	1.103706	1.138114	1.090004

Note: Sex ratio is represented as the ratio of females to male. A ratio of 1 implies parity.

There is a clear negative relation between the family size and the proportion of boys for NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06). As the family size decreases from over seven children to two children, the proportion of boys jump from 0.478 to 0.604 for NFHS II (1998-99). The corresponding sex ratio transition from having slightly more girls (109 girls per 100 boys) to significantly greater number of boys (65 girls per 100 boys). It is interesting to note that the proportion of boys in families with 2 or 3 children is much lower in 1992-93 compared to 1998-99 and 2005-06. The proportion of boys is greater than parity and quite similar across all three periods for families with 5 and more children indicating that the gender bias is concentrated at lower family sizes. Another interesting observation is that when the family size increases from 1 to 2 children, there is a slight jump in the proportion of boys from 0.57 to 0.60 in NFHS II (1998-99) and from 0.58 to 0.59 in NFHS III (2005-06). This could indicate that the son preference for the sample in this data set takes the form of wanting at least two sons.

6.1 Estimation using First Baseline Specification

Table 6.2 presents the OLS regression combining all three rounds of the NFHS data in which results in column (6) corresponds to the first baseline equation 5.1.

The dependent variable is the *PropofBoys* in all specifications. The key variable of interest is *FamSize*. Through column (1) to (3), explanatory variables are slowly added without controlling for state fixed effects or round fixed effects. Compared to column (2), column (3) includes mother's son preference (*mSonpreference*) as well. Column (4) to (6) repeat this same exercise, but controlling for state fixed effects and round fixed effects.

Table 6.2: Proportion of Boys and Family Size

Dep. var Propofboys	(1)	(2)	(3)	(4)	(5)	(6)
famsize	-0.0223***	-0.0267***	-0.0282***	-0.0239***	-0.0281***	-0.0286***
	(0.00262)	(0.00297)	(0.00293)	(0.00251)	(0.00285)	(0.00288)
mEduc		-0.00312***	-0.00220***		-0.00259***	-0.00189***
		(0.000491)	(0.000449)		(0.000421)	(0.000424)
mAge		0.00210***	0.00198***		0.00207***	0.00209***
		(0.000542)	(0.000526)		(0.000517)	(0.000511)
mWorks		-0.0205***	-0.0190***		-0.0193***	-0.0193***
		(0.00350)	(0.00311)		(0.00291)	(0.00292)
fEduc		0.0000474	0.0000206		-0.000385	-0.000334
		(0.000195)	(0.000192)		(0.000192)	(0.000190)
fAge		-0.000935***	-0.000786**		-0.000215	-0.000193
		(0.000251)	(0.000230)		(0.000249)	(0.000239)
mSonpreference			0.0682***			0.0645***
			(0.00488)			(0.00490)
rural		0.00839*	0.00438		0.0108***	0.00807**
		(0.00355)	(0.00339)		(0.00266)	(0.00263)
Religion (Others=0)						
Hindu		-0.0142	-0.0148		-0.00441	-0.00669
		(0.00980)	(0.00881)		(0.00685)	(0.00701)
Muslim		0.000326	0.00320		0.0165*	0.0168*
		(0.0109)	(0.0100)		(0.00791)	(0.00788)
Christian		-0.0213*	-0.0170		-0.00258	-0.00380
		(0.00972)	(0.00939)		(0.00637)	(0.00654)
Round Fixed Effect	No	No	No	Yes	Yes	Yes
State Fixed Effect	No	No	No	Yes	Yes	Yes
N	54722	54722	54722	54722	54722	54722
R-sq	0.021	0.026	0.039	0.030	0.035	0.046

Notes: In column 1-3, explanatory variables are added without controlling for round fixed effect or state fixed effect. In column 4-6, round fixed effect and state fixed effects are included. Column 6 represents the baseline specification (equation 5.1). Standard errors clustered at the state level (29 clusters) are in parentheses. *p<0.05, **p<0.01, ***p<0.001

To start with, there is a clear negative association between family size and the proportion of boys. The estimates for family size are significant at the 0.1 percentage level through all six specifications. The coefficient of family size in column (6) estimates that the proportion of boys increases by 2.86 percentage points when the family size decreases by 1 person. The estimate of family size is quite robust to the inclusion of control variables and fixed effects.

In terms of other control variables, I find that mothers with strong son preference (*mSonpreference*) have a much larger proportion of boys compared to mothers who do not. For mothers with strong preference, the proportion of boys jumps by 6.82 percentage points without state and round fixed effects and by 6.45 percentage points controlling for fixed effects.

The proportion of boys seems to be negatively associated with mother's education and mother's working status, as expected. While mother's education decreases proportion of boys by a mere 0.2 percentage points, working status has a larger impact of almost 2 percentage points.

There is also a clear cohort effect on the proportion of boys. Older mothers have a larger proportion of boys compared to younger mothers. The effect shows and increment of 0.2 percentage point for per addition year increase in mother's age. Thus, when comparing mothers aged 35 and mothers aged 49, the proportion of boys could differ by over 3 percentage points. Controlling for fixed effects, I observe that mother's characteristics have a more significant impact on the proportion of boys compared to father's characteristics.

The regression model suggests that the proportion of boys is slightly higher in rural areas compared to urban. This is quite contrary to evidence presented in the Figure 1.2 where rural areas were characterized by higher fertility and lower sex ratio. I suspect that this contrary evidence might be the result of sampling and data trimming bias.

In order to test the fluctuation in the relation between family size and proportion of boys, I re-run the regression equation 5.1 varying family size. That is, I test the relation between *PropofBoys* and *Famsize* when the family size is 1&2, 2&3, 3&4 and so on. The results are presented below in Table 6.3:

Table 6.3: Proportion of boys and Family Size - varying Family Size

	FamSize=1&2	FamSize=2&3	FamSize=3&4	FamSize=4&5	FamSize=5&6
Dep. var. PropofBoys	(1)	(2)	(3)	(4)	(5)
famsize	0.0190	-0.0613***	-0.0499***	-0.0440***	-0.0187***
	(0.0107)	(0.00619)	(0.00651)	(0.00427)	(0.00422)
mEduc	-0.00103	-0.00284***	-0.00435***	-0.00625***	-0.00737***
	(0.000887)	(0.000692)	(0.000647)	(0.000582)	(0.000796)
mAge	0.00210*	0.00114	0.00251***	0.00265***	0.00244*
	(0.000990)	(0.000652)	(0.000481)	(0.000619)	(0.000953)
mWorks	-0.0288***	-0.0210***	-0.0181**	-0.0157***	-0.0165***
	(0.00612)	(0.00468)	(0.00494)	(0.00345)	(0.00399)
fEduc	0.000746	0.000370	-0.000207	-0.000662	-0.000909*
	(0.000770)	(0.000340)	(0.000273)	(0.000383)	(0.000346)
fAge	-0.00117	-0.000286	-0.000403	0.000208	0.000519
	(0.000736)	(0.000407)	(0.000288)	(0.000222)	(0.000340)
mSonpreference	0.145***	0.0929***	0.0749***	0.0535***	0.0315***
	(0.0138)	(0.00653)	(0.00430)	(0.00410)	(0.00499)
rural	0.0223*	0.0120*	0.00285	0.00276	0.00133
	(0.00822)	(0.00555)	(0.00338)	(0.00380)	(0.00497)
Religion (Others=0)					
Hindu	-0.0171	0.00601	0.00792	-0.0108	-0.000802
	(0.0113)	(0.00946)	(0.00790)	(0.0117)	(0.0104)
Muslim	-0.0354	-0.00699	0.0194	0.0153	0.0284*
	(0.0187)	(0.0100)	(0.0109)	(0.0153)	(0.0135)
Christian	-0.0322*	-0.00781	0.0109	0.0104	0.0280
	(0.0142)	(0.0120)	(0.0107)	(0.00973)	(0.0146)
Round Fixed Effect	Yes	Yes	Yes	Yes	Yes
State Fixed Effect	Yes	Yes	Yes	Yes	Yes
N	13041	23799	24596	18756	12425
R-sq	0.044	0.049	0.052	0.044	0.039

Note: Each column, I run the regression with different family sizes controlling for round fixed effect and state fixed effects. Standard errors clustered at the state (29 clusters) level are in parentheses. *p<0.05, **p<0.01, ***p<0.001

The impact of family size on the proportion of boys significantly differs between the five columns. To start with, the estimate for family size is largest when the family size transitions from 3 to 2. That is, when the family size decreases from 3 to 2 children, the proportion of boys increases by over 6 percentage points. This estimate is still significantly larger (4.4 percentage points and 5 percentage points) when the family size transitions from 5 to 4 and 4 to 3 children.

In all five columns, mothers' preference and working status has significant impact. It is interesting to note that as I increase the family size, levels explanatory power of mother's son preference and working status decreases. When comparing families with 1 and 2 children, son preference and mother's working status explains most of the variation in the proportion of boys compared to the family size which has no significant impact. In all five iterations, father's age and education has a non-significant impact on the dependent variable.

The results from an alternative method of varying family size is presented in Appendix 9.3. I re-run the regression equation 5.1 separately when the family size is 2 or less, 3 or less, and so on. This provides an additional perspective to assess the fluctuation in the relation between family size and proportion of boys.

Given that preference for male children is the primary reason for the existence of gender bias in India, I test the relation between family size and proportion of boys separately for mothers without and with son preference. The results are presented below in Table 6.4:

Table 6.4: Proportion of Boys and Family Size - by existence of strong son preference

	mSonPreference=0	mSonPreference=1
Dep. Var. *PropofBoys*	(1)	(2)
famsize	-0.0232***	-0.0396***
	(0.00284)	(0.00407)
mEduc	-0.00246***	0.00162
	(0.000515)	(0.000818)
mAge	0.00206**	0.00199***
	(0.000612)	(0.000518)
mWorks	-0.0196***	-0.0194***
	(0.00383)	(0.00303)
fEduc	-0.000458	-0.000194
	(0.000243)	(0.000372)
fAge	-0.000332	0.0000594
	(0.000321)	(0.000255)
rural	0.00934**	0.00279
	(0.00316)	(0.00385)
Religion (Others=0)		
Hindu	-0.00383	-0.0127
	(0.00703)	(0.0102)
Muslim	0.0172*	0.00936
	(0.00829)	(0.0117)
Christian	-0.00193	-0.00134
	(0.00704)	(0.0144)
Round Fixed Effects	Yes	Yes
State Fixed Effects	Yes	Yes
N	37390	17332
R-sq	0.026	0.081

Note: Column 1 shows results for mothers without strong son preference and column 2, for mothers with strong son preference. Both columns control for round fixed effects and state fixed effects. Standard errors clustered at the state (29 clusters) level are in parentheses. *p<0.05, **p<0.01, ***p<0.001

It is interesting to see that a decrease in family size by one person, increases the proportion of boys by almost 4 percentage points for mothers with strong son preference compared to 2.3 percentage points change for mothers without strong son preference. Both these results are significant at a 0.1 percentage significance level. The negative relation between family size and the proportion of boys observed for mothers without strong son preference could indicates that the possibility of weak son preference.

Another interesting observation from Table 6.4 is the effect of mother's education on the proportion of boys. Mother's education level has no impact on the proportion of boys for mother with strong son preference. For mothers without strong son preference, a one-year increase in education reduces the proportion of sons by 0.2 percentage points, significant at 0.1 percentage significance level. Mother's work status has a significant impact in reducing the proportion of boys irrespective of son preference.

6.2 Estimation using Second Baseline Specification

In order to test hypothesis 3, I regress the proportion of boys separately for NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-09). These results corresponding to equation 5.2 are presented below in Table 6.5.

Table 6.5: Proportion of Boys and Family Size: by Round

	NFHS I (1992-93)	NFHS II (1998-99)	NFHS III (2005-06)
Dep. var			
PropofBoys	(1)	(2)	(3)
famsize	-0.0136***	-0.0330***	-0.0370***
	(0.00270)	(0.00293)	(0.00549)
mEduc	-0.000810	-0.00255***	-0.00198**
	(0.00130)	(0.000652)	(0.000610)
mAge	0.00253***	0.00220***	0.00148
	(0.000668)	(0.000482)	(0.000996)
mWorks	-0.0169**	-0.0194***	-0.0163**
	(0.00456)	(0.00450)	(0.00545)
fEduc	0.0000597	-0.00181***	-0.00000801
	(0.000460)	(0.000423)	(0.000209)
fAge	0.000293	-0.000657*	0.000366
	(0.000328)	(0.000237)	(0.000727)
mSonpreference	0.0505***	0.0675***	0.0741***
	(0.00594)	(0.00592)	(0.0106)
fSonpreference			0.0557***
			(0.00831)
rural	0.0132*	0.0103*	0.0110
	(0.00547)	(0.00416)	(0.00735)
wealthlevel		0.00661***	0.00346
		(0.00167)	(0.00257)
Religion (Others=0)			
Hindu	-0.0104	-0.0106	0.00356
	(0.00981)	(0.00607)	(0.0163)
Muslim	0.00357	0.0211*	0.0167
	(0.0123)	(0.00856)	(0.0168)
Christian	-0.0295	0.0000242	0.0118
	(0.0183)	(0.00846)	(0.0197)
Round Fixed Effect	Yes	Yes	Yes
State Fixed Effect	Yes	Yes	Yes
N	11357	29091	14274
R-sq	0.030	0.049	0.050

Note: Each column, I run the regression for different round controlling for state fixed effects. Standard errors clustered at the state (29 clusters) level are in parentheses. *p<0.05, **p<0.01, ***p<0.001

In all three rounds, family size has a significant negative impact on the proportion of boys. However, the impact differs across each round. As hypothesized, the impact is much larger in 2005-06 compared to 1992-93. As shown in the summary Table 5.1, the average family size for women aged 35-49 decreased from 4.7 in 1992-93 to 3.35 in 2005-06. Greater negative relation between family size and proportion of boys in 2005-06, indicates that as the family size decrease, gender bias manifestation intensifies. When the family size decreases by 1 person, the proportion of boys increases by 3.7 percentage points for NFHS III (2005-06) compared to 3.3 percentage points increase in NFHS II (1998-99) and 1.36 percentage point increase in NFHS I (1992-93).

In all three rounds, mothers' son preference plays a large significant role. However, for mothers with strong son preference, the proportion of boys increased by 7.4 percentage points in NFHS III (2005-06) compared to 5.05 percentage points increase in NFHS I (1992-93). In summary Table 5.1, it was shown that son preference has decreased over time. However, studying its impact on the proportion of boys, it is likely that mothers with son preference are more likely to resort to sex selection in 2005-06 compared to 1992-93.

Over all three rounds, working mothers have slightly lower proportion of boys in their families compared to non-working mothers. Considering data on father's son preference, which is available only for NFHS III (2005-06), the proportion of boys increases by 5.57 percentage points for father's with stronger son preference.

6.3 Estimation using Interactions

The first threat to identification is the joint determination of family size and the proportion of boys. Equation 5.3 introduced a model with two key interactions: *FamSize*mEduc* and *FamSize*mSonpreference*. Table 6.6 (column (2)) presents the results corresponding to equation 5.3.

Table 6.6: Proportion of Boys and Family Size - Including Interactions

Dep. Var. PropofBoys	(1)	(2)
famsize	-0.0222***	-0.0168***
	(0.00282)	(0.00228)
mEduc	-0.00154***	0.00474***
	(0.000417)	(0.00107)
mAge	0.00202***	0.00207***
	(0.000521)	(0.000515)
mWorks	-0.0196***	-0.0216***
	(0.00291)	(0.00287)
fEduc	-0.000346	-0.000300
	(0.000188)	(0.000194)
fAge	-0.000194	-0.000176
	(0.000247)	(0.000248)
mSonpreference	0.143***	0.154***
	(0.0159)	(0.0146)
FamSize*mSonpreference	-0.0192***	-0.0215***
	(0.00343)	(0.00324)
FamSize*mEduc		-0.00217***
		(0.000361)
rural	0.00760**	0.00856**
	(0.00264)	(0.00258)
Religion (Others=0)		
Hindu	-0.00576	-0.00530
	(0.00689)	(0.00684)
Muslim	0.0159*	0.0161
	(0.00775)	(0.00795)
Christian	-0.00285	-0.00116
	(0.00646)	(0.00658)
Round Fixed Effects	Yes	Yes
State Fixed Effects	Yes	Yes
N	54722	54722
R-sq	0.049	0.051

Note: Column (1) includes interaction for son preference only and Column (2) include interaction with mother's education as well. Standard errors clustered at the state level (29 clusters) are in parentheses. *p<0.05, **p<0.01, ***p<0.001

In Column (1), the effect of change in family size on the proportion of boys for mothers without son preference, is 2.2 percentage points. For mothers with strong son preference, the effect of change in family size on the proportion of boys is stronger. In Column (2), both interaction terms (*FamSize*mSonpreference* and *FamSize*mEduc*) are included. Similar to column (1), the partial effect of family size on proportion of boys is negative and this relation intensifies for mothers with strong son preference. It is interesting to note that with the inclusion of interaction term with mother's education, the partial effect of mother's education on the proportion of boys is positive. That is, increases in mother's education level would increase the proportion of boys. The negative sign for the interaction term *FamSize*mEduc* indicates that the effect of a decrease family size by one person would increase the proportion of boys at higher education levels. This provides evidence that highly educated individuals are more likely to act on their desires for sons through increased use of medical technology for sex selection thus worsening the sex ratio.

6.4 Estimation using 2SLS

The second limitation discussed in order to conclude causation between family size and the proportion of boys is the threat of reverse causality. As explained in section 5.4, I use a 2SLS to address this simultaneity bias. Using sex of the first child as instrument, the results from the first stage and second state (equation 5.4 and 5.5) along with OLS results is shown below:

Table 6.7: Proportion of Boys and Family Size: OLS and 2SLS

	OLS	2SLS	
		Second Stage	First Stage
	(1)	(2)	(3)
Dep. Var.	propofboys	propofboys	famsize
famsize	-0.0286***	-0.831***	
	(0.00288)	(0.0413)	
mEduc	-0.00189***	-0.0769***	-0.0939***
	(0.000424)	(0.00401)	(0.00486)
mAge	0.00209***	0.0307***	0.0368***
	(0.000511)	(0.00165)	(0.00406)
mWorks	-0.0193***	-0.0503***	-0.0457
	(0.00292)	(0.00288)	(0.0225)
fEduc	-0.000334	-0.00914***	-0.0110***
	(0.000190)	(0.000478)	(0.00209)
fAge	-0.000193	0.00784***	0.0101***
	(0.000239)	(0.000381)	(0.00140)
mSonpreference	0.0645***	0.135***	0.111***
	(0.00490)	(0.00613)	(0.0214)
rural	0.00807**	0.109***	0.126**
	(0.00263)	(0.00606)	(0.0403)
Religion			
Hindu	-0.00669	0.0340***	0.0453
	(0.00701)	(0.00840)	(0.0784)
Muslim	0.0168*	0.970***	1.187***
	(0.00788)	(0.0490)	(0.108)
Christian	-0.00380	0.279***	0.349
	(0.00654)	(0.0171)	(0.172)
male1			-0.309***
			(0.0277)
Round Fixed Effect	Yes	Yes	Yes
State Fixed Effect	Yes	Yes	Yes
N	54722	54722	54722
R-sq	0.046	0.240	0.297
F-stat			418.45
Wald Chi-sq		328.67	

Note: In the 2SLS regression, sex of the first child is used as an instrument for family size. Standard errors clustered at the state level (29 clusters) are in parentheses. *p<0.05, **p<0.01, ***p<0.001

The first stage results (column (3)) shows that sex of the first child is strongly correlated to the family size and is significant at a 0.1% significance level. It shows that family size decreases by 0.309 children when the first born child is male. Furthermore, the first stage F-statistic is highly significant suggesting that the analysis does not suffer from weak identification. The IV results shown in column (2) indicates a large negative relation between the family size and the proportion of boys significant at a 0.1% significance level. This relation confirms that a decrease in family size considerably increases the proportion of boys in each family. Compared to the effect using OLS model, the IV method shows a decrease in proportion of boys by 83 percentage points.

A potential problem in interpreting the IV result is that the sex of the first child and the proportion of boys are also highly correlated. The correlation between the two variables is 0.479. Part of this correlation might be independent of the effect of *male1* on *famsize*. I suspect that this correlation might be the result of the way the variable proportion of boys is constructed. Proportion of boys is simply the fraction of boy children within a family. Given that, if the sex of the first child is male, the proportion of boys is likely to be higher. In this case, the instrument *male1* affects the key dependent variable *propofboys* not only through *famsize* but also independently causing an overestimation of the IV estimate.

7. Discussion and Conclusions

In this section, I will discuss the results presented in the previous section follow by recommended policy focus areas and conclusion.

Discussion of Results

Based on literatures comparing the effects of fertility on the sex ratio, I hypothesized a negative effect of family size on the proportion of boys in the family. Empirical analysis presented in Section 6 provides compelling evidence of the increased manifestation of gender bias as fertility decreases. In particular, the proportion of boys considerable increases when the family size decreases.

On average, the proportion of boys increases by 2.86 percentage points when the family size decreases by one person. Given that on between NFHS I (1992-93) and NFHS III (2005-06), the average family size for women aged 35-49 decreased by 1.35 children, this implies that the proportion of boys would have increased by a total of 3.86 percentage points. Assuming that the sex ratio was at parity in 1992-93, a 3.86 percentage point increase in the proportion of boys would imply a change in the sex ratio from 1 to 0.856 females per male.

The effects found in this paper using actual family size and realized proportion of boys is similar to the results found by Jayachandran (2014) using desired proportion of sons in ideal family size. She showed that on average, desired percent of sons increases by 2.3 percentage points when the desired family size decreases by one child.

However, Jayachandran (2014) states that due to the positive correlation between family size and son preference, the negative effect of family size on the desired sex might be underestimated. In the first stage regression (shown in Table 6.7, Column (3)), I find that a positive correlation between son preference and family size could indeed be found in the data. Furthermore, in Table 6.4 where the regression results presented separately for mothers with and without strong son preference, I find that a decrease in family size by one person increases the proportion of boys by almost 4 percentage points compared to 2.32 percentage points for mothers without strong son preference. This indicates that son preference acts as a moderator and affects the strength of the relation between family size and the proportion of boys.

Comparing across time, I find that the effect of family size on the proportion of boys is larger in NFHS III (2005-06) compared to NFHS I (1992-93). However, these results are not directly comparable since I include additional controls for NFHS II and NFHS III (*wealthlevel* and *fSonpreference*). Despite the inclusion of additional controls, there is a dramatic increase in the effect of family size on the proportion of boys from -0.136 (for NFHS I (1992-93)) to -0.033 (NFHS II (1998-99)) and -0.037 (for NFHS III (2005-06)).

Another interesting observation comparing across time is the increased explanatory power of *mSonpreference*, despite decreases in the proportion of mothers with strong son preference (as shown in summary Table 5.1). A potential reason for this increase might be due to the increased availability of medical technology for sex determination couples with desire for smaller family sizes. Access to this technology makes it possible for mothers with strong son preference to indulge in sex selection in order to achieve desired sex composition within smaller family sizes rather than have more children in order to satisfy the need to have at least one son. That is, while the son preference might have decline over time, there might be more couples indulging in sex selection in NFHS III (2005-06) compared to NFHS I (1992-93).

Given the interrelation between the key variables and various socio-economic factors, it has been difficult to establish a causal relation between family size and the proportion of boys. However, the results from the interaction model and 2SLS provides a better understanding of the relation as well as the mechanism between variables. Using the results from Table 6.2 to 6.7, below I recommend some policy focus areas to address gender bias in India.

Policy Recommendations

In China, the unanticipated increase in the sex ratio over the years as a result of the one child policy has now led to the government relaxing this restriction. This policy relaxation could bring the sex ratio back to biologically normal level over time by allowing couples to have a son without resorting to sex selective abortion.

However, in the case of India, given the concerns over population explosion, it would not be ideal to increase fertility in order to combat the widening gap in sex ratio. The negative relation between family size and sex ratio demonstrated in this paper suggests that if the fundamental issue of son preference is not addressed, there would be an exacerbated manifestation of gender bias in the coming years. That is, as the country's fertility rate declines further, the national child sex ratio could fall below 900 girls per 1000 boys compared to the current ratio of 914 girls per 100 boys. This worsening sex ratio could result in increased incidences of crime, violence against women and lower female labour force participation thus hindering the growth and development of the country.

Using the results presented in Section 6, I recommend two interrelated policy focus areas. First is to improve female labour force participation rate. Second, is to reduce son preference.

Through all model specification, mother's work status (*mWork*) has been found to have a significant negative impact on the proportion of boys. Even for mothers with strong son preference, *mWork* is associated with a decrease in the proportion of boys by almost 2 percentage points. In China, Qian (2008) found improvements in the sex ratio when the demand for female labourers increased. Using an exogenous shock to the price of tea, which primarily employs female laborers, she found the increase in the demand for female labour and increases in the female agricultural income to be causally linked to the improvements in the survival rate of girls.

A potential policy to increase employment opportunity for women could be affirmative action. That is, encouraging employment equality by creating a specific employment quota for women. Increased female representation in leadership roles has been found to help reshape attitude towards women as well as raise aspirations and increase investment in girls (Beaman et al, 2009). Furthermore, increased participation of females in politics has also been associated with increased implementation of policies reflecting women's preference (Chattopadhyay and Duflo, 2004). While such policies already exist in India, further encouragement and stricter enforcement of the policy is required to ensure desired impact.

While increasing female labour force participation rate also reducing the overall son preference by raising the economic utility of girls, specific policies is also needed to change the cultural mindset and attitude towards women in order to reduce son preference. One potential way to influence a shift in cultural mindset is through education and increased exposure. While the results in section 6 show a significant negative relation between mother's education and proportion of boys in the baseline models, the results from the interaction model shown in Table 6.6 (Column (2)) indicates that education could actually worsen the sex ratio. Furthermore, results from the first stage of the 2SLS indicates that education decreases the fertility rate which could further worsen the sex ratio.

An alternative way which has proved quite effective in improving gender perception is through increased exposure to media and TV (Jensen and Oster, 2009). Increased exposure provides viewers with information about the outside world and different ways of life which may in turn influence their behaviour and attitude. In India, introduction of cable TV has been found to significantly decrease violence against women and reduce overall son preference (Jensen and Oster, 2009). Encouraging more TV series to portray women in stronger positions could result in an overall decrease in son preference and promote gender equality throughout the country.

Conclusion

In this thesis, I study the various factors influencing the sex ratio in general and more specifically the relation between family size and the proportion of boys. The empirical analysis shows three main results. First, I find that in general, smaller families have a larger proportion of boys. Second, son preference has a significant impact on the realized sex ratio and the effect of family size on the proportion of boys is larger for mothers with strong son preference compared to mothers without. Finally, the impact of a decrease in family size on the proportion of boys is larger in 2005-06 compared to 1992-93. These results explain the increased manifestation of gender bias observed in the Census Data. Empirical analysis presented in this thesis adds to the existing limited body of literature quantifying the relation between actual family size on the realized sex ratio.

As India transitions to lower fertility levels, this thesis suggests that the ratio is expected to worsen if the fundamental issue of sex preference is not addressed. Policies specifically focused on improving economic prospects of women needs to initiated. Recently, there has been some evidence of turnaround of this phenomenon of worsening sex ratio as fertility declines in South Korea. Chung and Das Gupta (2007) showed that the sex ratio bottomed reaching its lowest in the mid-1990s and has started to return to biologically normal. Similar turnaround could be expected from India as well in the future, even before reaching similar level of economic development as Korea if the government actively launches programs aimed at promoting equality and bridging the gap.

8. References

Agénor, P., Mares, J. and Sorsa, P. (2015): "Gender Equality and Economic Growth in India: A Quantitative Framework", OECD Economics Department Working Papers, No. 1263, OECD Publishing, Paris

Amin, R. and Mariam, A. G. (1987): "Son Preference in Bangladesh: An Emerging Barrier to Fertility Regulation", Journal of Biosocial Science, 19, 221-228

Amuedo-Dorantes, C., and Shoshana, G. (2007): "Cohort-Level Sex Ratio Effects on Women's Labor Force Participation", Review of Economics of the Household, 5 (3), 249-78

Arnold, F., Choe, M.K., Roy, T.K. (1998): "Son Preference, the Family-Building Process and Child Mortality in India", Population Studies, 52, 301-315

Arnold, F., Kishor, S., and Roy, T.K. (2002): "Sex Selective Abortions in India", Population and Development Review, 28 (4), 759-785

Barcellos, S. H., L. Carvalho, and A. Lleras-Muney (2012): "Child Gender and Parental Investments in India: Are Boys and Girls Treated Differently?", National Bureau of Economic Research, Working Paper 17781

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

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

Becker, G. & Lewis, G. H. (1973): "On the Interaction between the Quantity and Quality of Children", Journal of Political Economy, 81 (2), 279-288

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2009): "Powerful Women: Does Exposure Reduce Bias?", Quarterly Journal of Economics, 124, 1497-1540.

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

Bongaarts, J., and Guilmoto. C. Z. (2015): "How Many More Missing Women? Excess Female Mortality and Prenatal Sex Selection, 1970-2050", Population and Development Review, 41 (2), 241-269

Butle, E., Tu, Q., List, J. (2015): "Battle of the Sexes: How Sex Ratios Affect Female Bargaining Power", Economic Development and Cultural Change, 64 (1), 143-161

Chattopadhyay, R. and Duflo, E. (2004): "Women as Policy Makers: Evidence from a Nationwide Randomized Experiment in India", Econometrica, 72, 1409-1443.

Chaudhuri, S. (2012): "The Desire for Sons and Excess Fertility: A Household-level Analysis of Parity Progression in India", International Perspectives on Sexual and Reproductive Health, 38 (4), 178-186

Chaudhri, D.P. and Jha, R. (2013): "India's Gender Bias in Child Population, Female Education and Growing Prosperity", International Review of Applied Economics, 27 (1), 23-43

Chung, W. and Das Gupta, M. (2007): "The Decline of Son Preference in South Korea: The Roles of Development and Public Policy", Population and Development Review, 33 (4), 757-783

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

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

Das Gupta, M. (2005): "Explaining Asia's 'Missing Women': A New Look at the Data", Population and Development Review, 31 (3), 529-35

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

Das Gupta, M., Zhenghua, J., Bohua, Li, Xie, Z., Woojin, C. and Bae, H. (2003): "Why is Son Sreference so Persistent in East and South Asia? A Cross-country Study of China, India and the Republic of Korea", Journal of Development Studies, 40 (2), 153-187

Das Gupta, M., Chung, W., and Shuzhuo, L. (2009): "Evidence for an Incipient Decline in Numbers of Missing Girls in China and India", Population and Development Review, 35 (2), 401-416

Dreze, J., and Khera, R. (2000): "Crime, Gender, and Society in India: Insights from Homicide Data", Population and Development Review, 26 (2), 335-352

Ding, Q. J., and Hesketh, T. (2006): "Family Size, Fertility Preferences and Sex Ratio in China in the Era of the One Child Policy: Results from National Planning and Reproductive Health Survey", BMJ 333, 371-373

Duflo, E. (2012): "Women's Empowerment and Economic Development", Journal of Economic Literature, 50 (4), 1051-1079

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

Edlund, L., Li, H., Yi, J., & Zhang, J. (2007): "Sex ratios and crime: Evidence from China's One-Child Policy", IZA Discussion Paper No. 3214, Bonn, Germany: Institute for the Study of Labor

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

Guilmoto, C. Z. (2012): "Skewed Sex Ratios at Birth and Future Marriage Squeeze in China and India, 2005–2100", Demography 49 (1), 77-100

Graham, E. (2007): "Son preference, Female Deficit and Singapore's Fertility Transition" In: Attané, I. and Guilmoto, C.Z. (eds.). Watering the neighbour's garden: The growing demographic female deficit in Asia. Paris: CICRED: 73-88

Gubhaju, B. (2007): "Fertility Decline in Asia: Opportunities and Challenges", Japanese Journal of Population, 5 (1), 19-42

Hesketh, T., Li, L., Zhu, W.X. (2005): "The Effect of China's One-child Family Policy after 25 Years", New England Journal of Medicine, 353, 1171-1176

Hudson, V., Den Boer, A.M. (2002): "A Surplus of Men, a Deficit of Peace: Security and Sex Ratios in Asia's Largest States", International Security, 26, 5-38

Jayachandran, S. (2014): "Fertility Decline and Missing Women", National Bureau of Economic Research, Working Paper 20272

Jayachandran, S. and Kuziemko, I. (2009): "Why Do Mothers Breastfeed Girls Less than Boys: Evidence and Implications for Child Health in India", National Bureau of Economic Research, Working Paper 15041

Jayaraj, D. (2009): "Exploring the Importance of Excess Female Mortality and Discrimination in 'Natality' in Explaining the 'Lowness' of the Sex Ratio in India", Developing Economies, 47 (2), 177-201

Jensen, R. and Oster, E. (2009): "The Power of TV: Cable Television and Women's Status in India", Quarterly Journal of Economics, 124, 1057-1094

Jha P, Kumar R, Vasa P, Dhingra N, Thiruchelvam D, Moineddin R. (2006): "Low Female-to-male Sex Ratio of Children Born in India: National Survey of 1.1 Million Households", Lancet 367, 211-218

Kumar, S and Kugler. A (2011): "Testing the Children Quantity-Quality Trade-Off in India", MPRA Paper 42487 University Library of Munich, Germany

Lee, J. (2008): "Sibling Size and Investment in Children's Education: An Asian Instrument", Journal of Population Economics, 21 (4), 855-875

Li, H., J. Yi, and J. Zhang (2011): "Estimating the Effect of the One-Child Policy on the Sex Ratio Imbalance in China: Identification Based on the Difference-in-Differences", Demography, 48, 1535-1557

Meslé, F., Vallin, J., and Badurashvili, I. (2007): "A Sharp Increase in Sex Ratio at Birth in the Caucasus. Why? How?", In: Attané, I. and Guilmoto, C.Z. (eds.). Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia. Paris: CICRED: 73-88

Mitra, A. (2014): "Son Preference in India: Implications for Gender Development", Journal of Economic Issues, 48 (4), 1021-37

Park, C. B., and Cho, N.H. (1995): "Consequences of Son Preference in a Low-Fertility Society: Imbalance of the Sex Ratio at Birth in Korea", Population and Development Review, 21 (1), 59-84

Park, C. B. (1983): "Preference for Sons, Family Size, and Sex Ratio: An Empirical Study in Korea", Demography, 20, 333-352

Qian, N. (2008): "Missing Women and the Price of Tea in China: The Effect of Sex Specific Earnings on Sex Imbalance", Quarterly Journal of Economics, 123 (3), 1251-1285

Rahman, M. and DaVanzo, J. (1993): "Gender Preference and Birthspacing in Matlab, Bangladesh", Demography, 30, 315-332

Rahman, M., Jalaluddin, A., Phillips, J. F., and Becker, S. (1992): "Contraceptive use in Matlab, Bangladesh: The role of gender preference", Studies in Family Planning, 23, 229-242

Retherford, R. D. and Roy, T.K. (2003): "Factors Affecting Sex-Selective Abortion in India and 17 Major States", Technical Report, Mumbai, India

Sen, A. (1990): "More Than 100 Million Women Are Missing", The New York Review of Books, 37 (20)

Sen, A. (1992), "Missing Women", British Medical Journal, 304, 587-588

South, S. J., Trent, K., & Bose, S. (2014): "Skewed Sex Ratios and Criminal Victimization in India", Demography, 51, 1019-1040

Subramanian, S. V., and Selvaraj, S. (2009): "Social Analysis of Sex Imbalance in India: Before and After the Implementation of the Pre-Natal Diagnostic Techniques (PNDT) Act.", Journal of Epidemiology and Community Health, 63, 245-252

Tucker J.D., Henderson G.E., Wang T.F., et al.(2005): "Surplus Men, Sex Work, and the Spread of HIV in China", AIDS 19, 539-547

Wei, S., and Zhang, X. (2009): "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China", National Bureau of Economic Research, Working Paper No. 15093

Wen, X. (1992): "The Effect of Sex Preference on Subsequent Fertility in Two Provinces of China", Asia-Pacific Population Journal, 7 (4), 25-40

Wen, X. (1993): "Effect of Son Preference and Population Policy on Sex Ratios at Birth in Two Provinces of China", Journal of Biosocial Science, 25, 509-521

Zeng, Y., Tu, P., Gu, B., Xu, Y. Li, B. and Li, Y. (1993): "Causes and Implications of the Recent Increase in the Reported Sex Ratio at Birth in China", Population and Development Review, 19, 283-302

Zhu, W. X., Li,L. and Hesketh, T. (2009): "China's Excess Males, Sex Selective Abortion, and One Child Policy: Analysis of Data from 2005 National Intercensus Survey", British Medical Journal, 338, b1211

9. Appendix

9.1 Appendix A: Age Specific Fertility Rate

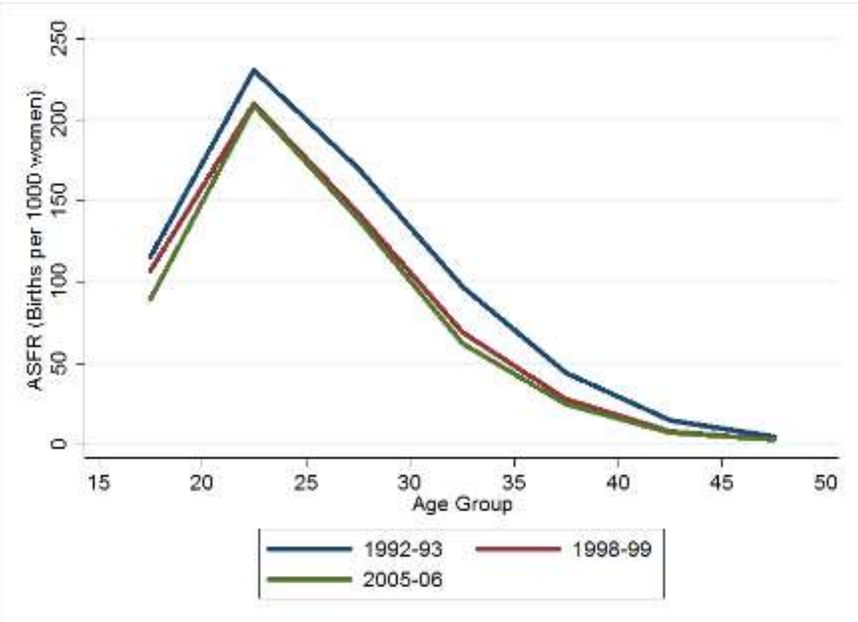
The Age Specific Fertility Rate (ASFR) is defined as the total number of births occurring during a given year or reference period for every 1,000 women of reproductive age classified in single or five-year age groups. It is calculated as:

$$ASFR_a = (\frac{B_a}{N_a}) \times 1000$$

Where B_a is the total number of births to women in age group 'a' in a given year or reference period and N_a is the number of women in age group 'a' during the specified reference period.

The ASFR according to NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06) is shown below:

Figure 9.1: Age Specific Fertility Rate by Round



Note: The ASFR is classified in 5-year groups. Data Source: Health Education to Village (HETV), NFHS 3 India Reports

It is clearly visible that the ASFR peaks in the 20-25 age group and considerable drops to under 50 births per 1000 women in the 35-39 age range and declines further in the 40-44 and 45-59 age range for all three rounds of the NFHS. This suggests that the women in the above the age 35 have far fewer births and are more likely to have completed child bearing.

9.2 Appendix B: Family Size and Proportion of Boys by State and Round

The Table below summarizes the proportion of boys and family size in different state according to NFHS I (1992-93), NFHS II (1998-99) and NFHS III (2005-06).

Table 9.1: Summary - Proportion of boys and Family size: By State and Round

	NFHS I (1992-93)		NFHS II (1998-99)		NFHS III (2005-06)	
District	Propofboys	Famsize	Propofboys	Famsize	Propofboys	Famsize
Andhra Pradesh	0.4763	4.20	0.5317	3.47	0.5113	3.00
Arunachal Pradesh	0.5067	4.97	0.5286	3.87	0.5074	4.00
Assam	0.4729	5.44	0.5332	3.91	0.5334	3.23
Bihar	0.4921	5.02	0.5474	4.47	0.5304	4.55
Chhattisgarh					0.5077	3.75
Goa	0.4885	4.50	0.5190	2.89	0.5078	2.47
Gujarat	0.4982	4.42	0.5492	3.44	0.5684	3.18
Haryana	0.5104	4.71	0.5822	3.76	0.5781	3.48
Himachal Pradesh	0.5011	4.25	0.5538	3.33	0.5422	3.05
Jammu and Kashmir	0.4956	5.10	0.5188	4.12	0.5364	3.85
Jharkhand					0.5719	3.74
Karnataka	0.4704	4.83	0.5299	3.53	0.5365	2.87
Kerala	0.4615	4.39	0.5064	2.88	0.4841	2.35
Madhya Pradesh	0.5043	4.91	0.5559	4.11	0.5743	3.44
Maharashtra	0.4945	4.50	0.5406	3.26	0.5527	2.92
Manipur	0.5125	5.73	0.5258	4.31	0.5284	3.59
Meghalaya	0.5318	5.69	0.5152	4.82	0.5290	4.34
Mizoram	0.4742	5.07	0.5135	3.96	0.5044	3.80
Nagaland	0.5477	6.48	0.5115	4.83	0.5159	4.26
Odisha	0.4761	4.54	0.5469	3.48	0.5412	3.06
Punjab	0.5007	4.48	0.5867	3.29	0.5818	3.12
Rajasthan	0.4992	5.10	0.5485	4.31	0.5617	4.20
Sikkim			0.5133	4.08	0.5171	3.11
Tamil Nadu	0.4740	4.22	0.5200	2.95	0.5305	2.57
Tripura	0.5244	5.33	0.5179	3.43	0.5201	3.02
Uttar Pradesh	0.5036	5.13	0.5541	4.45	0.5553	4.25
Uttarakhand					0.5602	3.45
West Bengal	0.4860	4.87	0.5307	3.26	0.5334	2.93
New Delhi	0.4941	4.68	0.5570	3.31	0.5774	3.08
N	11357		29091		14274	

Notes: The number of states surveyed differs in each round. The proportion of boys and family size is calculated on the trimmed data as explained in Section 5.

9.3 Appendix C: Regression varying Family Size

The table below shows an alternative method for fluctuating family size to assess the relation between proportion of boys and family size.

Table 9.2: Proportion of boys and Family Size: Varying Family Size

	Famsize<=2	Famsize<=3	Famsize<=4	Famsize<=5	Famsize<=6
Dep. var Propofboys	(1)	(2)	(3)	(4)	(5)
famsize	0.019	-0.0351***	-0.0421***	-0.0433***	-0.0394***
	(0.0107)	(0.00552)	(0.00486)	(0.00417)	(0.00383)
mEduc	-0.00103	-0.00181**	-0.00237***	-0.00271***	-0.00266***
	(0.000887)	(0.000621)	(0.000622)	(0.000510)	(0.000483)
mAge	0.00210*	0.00167*	0.00216***	0.00199**	0.00210**
	(0.000990)	(0.000681)	(0.000542)	(0.000590)	(0.000573)
mWorks	-0.0288***	-0.0220***	-0.0204***	-0.0187***	-0.0193***
	(0.00612)	(0.00437)	(0.00370)	(0.00319)	(0.00309)
fAge	-0.00117	-0.00077	-0.000654*	-0.00036	-0.00029
	(0.000736)	(0.000409)	(0.000293)	(0.000290)	(0.000269)
fEduc	0.000746	0.000402	0.000128	-5.1E-05	-0.00018
	(0.000770)	(0.000382)	(0.000276)	(0.000223)	(0.000192)
mSonpreference	0.145***	0.103***	0.0899***	0.0786***	0.0705***
	(0.0138)	(0.00767)	(0.00505)	(0.00481)	(0.00498)
rural	0.0223*	0.0133*	0.00998*	0.00935*	0.00817*
	(0.00822)	(0.00553)	(0.00418)	(0.00364)	(0.00308)
Religion (Others=0)					
Hindu	-0.0171	-0.00194	-0.00523	-0.00612	-0.00615
	(0.0113)	(0.0104)	(0.00845)	(0.00713)	(0.00677)
Muslim	-0.0354	-0.0111	-0.00264	0.00406	0.00894
	(0.0187)	(0.0103)	(0.0103)	(0.00941)	(0.00855)
Christian	-0.0322*	-0.0149	-0.0107	-0.00819	-0.0049
	(0.0142)	(0.0120)	(0.0101)	(0.00747)	(0.00646)
Round Fixed Effect	Yes	Yes	Yes	Yes	Yes
State Fixed Effect	Yes	Yes	Yes	Yes	Yes
N	13041	26484	37637	45240	50062
R-sq	0.044	0.041	0.045	0.049	0.05

Note: Each column, I run the regression with different family sizes controlling for round fixed effect and state fixed effects. Standard errors clustered at the state (29 clusters) level are in parentheses. *p<0.05, **p<0.01, ***p<0.001