10.14754/CEU.2019.02

Essays in Development Economics with Special Emphasis on Gender Inequalities

by

Suparna Das

Submitted to

Central European University

Department of Economics and Business

in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

Advisor: **Andrea Weber**

Budapest, Hungary

10.14754/CEU.2019.02

10.14754/CEU.2019.02

CENTRAL EUROPEAN UNIVERSITY
DEPARTMENT OF ECONOMICS AND BUSINESS

The undersigned hereby certify that they have read and recommend to the Department of Economics and Business for acceptance a thesis entitled **"Essays in Development Economics with Special Emphasis on Gender Inequalities"** by **Suparna Das**
Dated: March 5, 2019

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Chair of the Thesis Committee:

Julius Horvath

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Advisor:

Andrea Weber

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Internal Examiner:

Arieda Muço

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
External Examiner:

Anna Raute

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
External Member:

Agnes Szabo-Morvai

10.14754/CEU.2019.02

CENTRAL EUROPEAN UNIVERSITY

DEPARTMENT OF ECONOMICS AND BUSINESS

Author: **Suparna Das**

Title: **Essays in Development Economics with Special Emphasis on Gender Inequalities**

Degree: **Ph.D.**

Date: **March 5, 2019**

Hereby, I testify that this thesis contains no material accepted for any other degree at any other institution and that it contains no material previously written and/or published by another person except when the appropriate acknowledgement is made.

Signature of the author: Suparna Das .....

10.14754/CEU.2019.02

I dedicate my doctoral thesis to my father Pronab Kumar Das.

10.14754/CEU.2019.02

Table of Contents

1 An economic analysis of gender gap in household demand for education: Evidence from India **11**

1.1 Motivation 11

1.2 Literature Review 14

1.3 Theoretical Background 18

1.3.1 Household Demand for Education: 18

1.3.2 Household Demand for Education based on the Gender of Child . 23

1.4 Data and Estimation Methodology: 26

1.4.1 Data 26

1.4.2 Estimation Strategy 28

1.5 Results 32

1.5.1 Descriptive Statistics 32

1.5.2 Measuring difference in education demand between boys and girls: 34

1.5.3 Gender difference in enrollment across levels of education 39

1.5.4 Labor market implications on enrollment 42

1.5.5 Reciprocity and Enrollment 49

1.5.6 Decomposition of Gender gap in school enrollment and contribution of the identified factors: 54

1.5.7 Other Crucial Factors and gender gap in enrollment 57
1.5.8 Robustness Checks 65
1.6 Discussion and Policy Implications 67

2 Impact of 'Having a Son' on Women's Intra-household Status: Evidence from India 83

2.1 Motivation 83
2.2 Literature Review 84
2.3 Data 88
2.4 Empirical Strategy 91
2.5 Results 99
2.5.1 Descriptive Statistics 99
2.5.2 Probit Estimation results: 102
2.5.3 Ordered Probit Estimation: Using Categories of Responses 112
2.5.4 Impact of having son on domestic violence to mothers 115
2.5.5 Other important factors to women's bargaining power within household 117
2.5.6 Robustness Check 125
2.6 Conclusion 126

3 The Co-existence of Biased Sex Ratio and Crime against Women in India: Examining the Causality 133

3.1 Motivation 133
3.2 Theoretical Background 139
3.3 Literature Review 144
3.4 Data 147
3.5 Empirical Strategy 149
3.6 Results 153
3.6.1 Descriptive Statistics: 153

3.6.2 Results: The relationship between the sex ratios and crime against women 157
3.6.3 Comparison of relationships of crime against women, total crime and other crime with sex ratio 162
3.6.4 YSR Quintile-wise analysis of the relationship between sex ratios and Crime against women 163
3.6.5 The region wise analysis of the relationship between sex ratios and crime against women 166
3.6.6 Crime-head wise analysis: 167
3.7 Discussion and Policy Implications 172

10.14754/CEU.2019.02

10.14754/CEU.2019.02

List of Figures

1.1 Gross Enrollment Rate at different education levels (2011)- India 13
1.2 Gender and age-wise enrollment of children at school-going age 32
1.3 Distribution of siblings and male siblings in eldest children sample 33
1.4 Drop out rate at different school levels 40
1.5 Gender wage gap across education levels 43
1.6 Gender-wise standard deviation of wages across education levels 43
1.7 Gender-wise success in income quintiles across education levels 44
A1.1 Age distribution of children at school-going age 81

2.1 Parity-wise Sex ratio in Birth-order 98

3.1 Sex Ratios (1961-2011) in India 134
3.2 Growth in crime against women, other crimes and total crime (1996-2014) 135
3.3 Share of different crimes in Crime against women (1995-2013) 136
3.4 Crime-head-wise growth of different crime under CRW: 1995-2014 137
A3.1 Age group-wise percentage of persons arrested 182

10.14754/CEU.2019.02

10.14754/CEU.2019.02

List of Tables

1.1 Out-of school rate among children (in per cent), 2014 12
1.2 Descriptive Statistics of school-going age children 35
1.3 Probit Regression: Marginal effects on enrollment of children 35
1.4 Marginal effects on enrollment of single children 39
1.5 Probit Results: Enrollment of children across different school levels . . . 40
1.6 Labor market impact on enrollment: $Pr(E_{Si}|ER_{Si}, Var(W_{Si}))$ 45
1.7 Labor market impact on enrollment across education levels: 49
1.8 Impact of parents' reciprocity expectation on enrollment $Pr(E_{Si}|R_i)$ . . . 50
1.9 Gender difference in imapct of reciprocity on enrollment 53
1.10 Decomposition of Gender Gap in Enrollment: All and Eldest Children Samples 55
1.11 Marginal effects on enrollment of boys and girls 60
1.12 Ordered Probit Regression Results - Marginal effects of gender on schooling choice between No schooling, Public / Private schooling 63
1.13 Enrollment of children: Using other estimation methods and IHDS-2005 . 66
A1.1 Enrollment of children at school-going age 75
A1.2 Probit Regression: Marginal effects on enrollment of children - 76
A1.3 Marginal effects of gender on enrollment decision of Single children - . . . 77

A1.4 Missing values of labor market variables are substituted 77
A1.5 Marginal impact of having very young siblings on enrollment 78
A1.6 Descriptive statistics of the families with single child and multiple children: 78
A1.7 Labor market success on enrollment of children 78
A1.8 Impact of age on enrollment of children across different age-groups 79
A1.9 Marginal Impact of Mother's say on enrollment of children 79
A1.10 Panel Probit Regression Results with state dummies: 80
A1.11 Enrollment across age-groups and completion levels: 80

2.1 Sex ratio of Children across the order of birth 96
2.2 Abortion and Gender composition of children 96
2.3 Summary of Household Characteristics 100
2.4 Summary of Individual Characteristics 100
2.5 Women having at least a son and having say in HH decisions 101
2.6 Probit Regression Results 103
2.7 IV Probit Regression Results 104
2.8 Reduced Form Regression Results 107
2.9 IV Probit Regression Results - Out of reproductive cycle sample 108
2.10 IV Probit Regression: Women have an independent say 110
2.11 Probit Regression Results - Women Sample 111
2.12 Ordered IV Probit Regression: Couples Sample 112
2.13 Ordered Reduced Form Regression - Couple Sample 114
2.14 Son-effect on Domestic Violence: Using Women's responses in Couple Sample 116
2.15 Son-effect on Severe Domestic Violence: Across age of the last born . . . 117
2.16 Marginal Effects from IV-Probit Estimations with Interaction terms: Women Responses 119
2.17 Marginal Effects from Probit Estimations with Interaction terms: Women Responses 120
2.18 Linear Probability Model: Using Couple Sample 125
2.19 IV Probit Regression Marginal Effects: NFHS-3 (2005-2006) 126

A2.1 Parity-wise (upto Parity 3) gender composition of children 129
A2.2 Marginal Effects from Probit Estimations - Using Women Responses . . . 129
A2.3 Marginal Effects from Probit Estimations - Using Husbands Responses . 130
A2.4 Marginal Effects from IV-Probit Estimations - Using Women Responses . 130
A2.5 Marginal Effects from IV-Probit Estimations - Using Husbands Responses 131
A2.6 Marginal Effects from Reduced form Estimations: Women Responses . . 131
A2.7 Marginal Effects from Reduced form Estimations: Husbands Responses . 132

3.1 Descriptive Statistics: 154
3.2 Highest crime incidences in districts during 1995-2013 154
3.3 Change in sex ratio over decades 157
3.4 OLS results: Impact of youth sex ratio on crime against women 158
3.5 Impact of youth sex ratio on crime against women 159
3.6 Impact of child sex ratio on crime against women 161
3.7 Relationship of sex ratios with different crime categories 162
3.8 Impact of youth sex ratio on crime against women across quintiles 164
3.9 Impact of child sex ratio on crime against women across quintiles 165
3.10 Impact of sex ratios on crime against women across regions 166
3.11 Crime-head wise analysis: Impact of sex ratios 169
3.12 Some Marriage Statistics of India: 170
A3.1 Panel level Descriptive Statistics 178
A3.2 Impact of YSR on CRW after dropping Delhi and 24 Pargana 179
A3.3 Impact of CSR_{-15} on CRW after dropping Delhi and 24 Pargana 179
A3.4 Impact of sex ratios on crime against women across quintiles 180
A3.5 Region-wise States/UTs and number of districts 180
A3.6 Impact of sex ratios on crime against women across regions 181
A3.7 Descriptive Statistics: Different crime-head under CRW (1995-2013) . . . 181
A3.8 Crime-head Wise Analysis: Excluding extreme observations (outliers) . . 181

10.14754/CEU.2019.02

10.14754/CEU.2019.02

Acknowledgement

I take this opportunity to acknowledge and convey my gratitude to all those people who helped and gave me the possibility to complete my thesis.

Foremost, I am deeply indebted to my supervisor, Andrea Weber for her continuous support and guidance during my PhD. Her patience to listen my doubts, prolific guidance and confidence in me had motivated me to perform to the best of my ability. Also, I am truly grateful to Gábor Kézdi for intellectual inspirations and insightful guidance on how to proceed with my research during the earlier years of the Doctoral program.

I would also like to thank the examiners Anna Raute and Arieda Muço for their time and efforts. I believe the present improvements in the thesis would not be possible without their astute comments and suggestions.

I also thank the Economics department of CEU for giving me the space and opportunities to learn and conduct research. I am particularly indebted to Prof László Mátyás for encouragement and stopping me to give-up at a very difficult time. I would also like to express my gratitude to my professors, and acknowledge the support from Julius Horvath, Péter Mihályi and Ádám Szeidl on various occasions. I am also thankful to Katalin Szimler, Veronika Orosz and Melinda Molnar, their continuous support made the journey much easier. Studying at the Central European University has been a wonderful experience and I am thankful to the CEU community for support and love.

10.14754/CEU.2019.02

I am also sincerely grateful to Sujata Singh for all her encouragement and love throughout the years. Among my fellow colleagues and friends, I would like to mention Franklin, Julian, Valleria and Vladimir for their help and moral support.

Especially, I am very grateful to my mother for her patience and understanding during these years. Last but not the least; I would like to give my sincere thanks to my husband Anupam, his firm belief in me, continuous encouragement to pursue the degree and love enabled me to complete the work.

10.14754/CEU.2019.02

Abstracts

An economic analysis of gender gap in household demand for education: Evidence from India

Education plays a crucial role in building tomorrow's human capital, and thus it is an important tool for economic growth and development. Followed by second Millennium Development Goal's (MDG) (2000) call for achieving universal school education by 2015, extensive government initiatives with special emphasis on girls were undertaken in India. The access to education has shown tremendous progress and become successful to bring almost all potential pupils to primary (standard I-V) school. However, starting from the elementary (standard VI-VIII) level onwards the gender gap in enrollment persists and widens with level of education in India. This paper quantifies the gender difference in enrollment decision for children and provides a theoretical structure to the underlying demand side factors that influence parents to keep girls out of the post-primary education system compared to boys. The analysis uses the 2^{nd} round dataset of India Human Development Survey (IHDS), published in 2012. This paper finds significant gender gap in enrollment, and a girl child is on average 3.6 per cent less likely to continue schooling compared to a boy. The enrollment probability of girls worsens with higher birth order, an eldest sister has significantly 5.2 per cent lower probability to continue school education compared to an eldest brother, keeping other things same. Further, it also finds

that beyond age of 14 when children are not anymore entitled to get free, compulsory education under Right to Education (RTE), girls' enrollment probability gets lower, such as a girl above age of 14 years is 7.8 per cent less likely to continue schooling and if she is the eldest sister among siblings her probability to discontinue becomes as large as 11 per cent (significantly) compared to similar boys. Labor market variables, especially returns to education and variability in wages play crucial and significant roles in schooling decision of children. Parents' reciprocity expectation, household responsibilities at adolescent years, and cost of education are also found responsible for lower enrollment of girls compared to boys.

Impact of 'Having a Son' on Women's Intra-household Status: Evidence from India

In intra-household settings, individual bargaining power is crucial for positions and control over decisions that influences resource allocation and individual well-being. Bargaining power may differ between men and women for various reasons, generally due to unequal social norms; and differential access to education, occupation and asset holding or income. Patriarchal societies often put higher values to having a son than having a daughter. The preference for son comes from the perception of higher utility gain from son(s) compared to daughter(s). If it is so, then individual's fertility outputs can play an important role in bargaining power. Also, in such societies with prevalence of patriarchy and preference for son, child-bearing and childcare are considered to be women's virtue and sole responsibility. Thus, fertility output may play more important role for women's bargaining power than men. This paper examines whether having a son has any implications in women's intra-household bargaining power and their say in different decision makings in the family. The paper primarily uses the data from sample of couples from National Family and Health Survey (NFHS) of fourth round (2015) on India. The estimations from probit regressions show that women with at least a son are significantly more likely to have a say; on average 2 - 4 per cent higher say in different decision-making in household, compared to women with no son. And women with first born son have

around 1 per cent higher say in decisions compared to women with first born daughter. However, having son does not significantly change husbands' views towards women's intra-household say, in general.

The Co-existence of Biased Sex Ratio and Crime against Women in India: Examining the Causality

Preference for son has deep historical roots in India and is evident in its highly skewed sex ratios. The country has recently been on news often for heinous rape incidences. It registers 27 per cent annual growth rate in crime against women in 2013. The economic theory indicates that the scarcity of girls should make the girls dearer to society, but the reality is opposite in India. It creates a puzzle that in spite of the scarcity India still doesn't value their women. Instead, the statistics show that crime against women increasing at higher rate compared to overall crime. Therefore, it becomes interesting to explore the puzzle and to examine whether this coincidence of pro-male biased sex ratios and higher crime on women bears any empirical relationship in case of India. The paper uses data from Census and National Crime Records Bureau for the analysis. The district-year panel analysis found that there is a significant negative relation between sex ratios and crime against women. It is found that increase in 1 female in the population of 1000 males, that is 1 unit increase in the youth sex ratio in favor of female will decrease the crime against women by 0.53 per cent, keeping other things same. Further, it is also found that the impact of sex ratios is highest on domestic violence, followed by kidnapping compared to other crime against women.

10.14754/CEU.2019.02

10.14754/CEU.2019.02

Introduction

An extraordinary surge in research interests in gender related issues, especially gender-inequalities can be observed among economists and social scientists during the last few decades. This increasing motivation among researchers comes from its relationship with economic growth and development at macro-level and also individual well-being at micro-level. A large number of studies show that growth and development of a country can be greatly improved by correcting gender imbalances. Keeping these in mind, I explore three diverse but inter-related facets of gender inequalities, and examine their causes and consequences in my doctoral thesis.

Human capital is identified as one of the crucial factors of production. Education is the most important component to build human capital, and thus, helps in economic growth. Female population constitutes, on average, half of the population globally, and their relative absence from formal education system in comparison to the male counterparts leads loss in economic output primarily by not-realizing the half of the potential human capital stock. In other words, in locations where women cannot or do not participate in different domains of economic life as equally as men, those places would not reap its full potential benefits from human resources.

The first chapter of my thesis addresses this issue of inequality in school education among the potential pupils. It provides a theoretical explanation to gender gap in parent's

decision to provide less schooling to daughters compared to sons, even if supply of (access to) education is same for both. When girls remain lower-educated, then half of the potential human capital will remain under-built and under-utilized. This motivates me to recognize the demand-side factors within household that keep the girls out-of-school with higher probability compared to their brothers. Using Indian data, this chapter quantifies the gender difference in school enrolment decision for children and also measures the contribution of the household-level demand side factors to this gap in school enrollment.

Lower participation of females in economic spheres are not only limited to formal education system, but extends to labor market, and also in political sphere. Recently, women participation in political platform has started to increase, mostly due to implementation of legislative mandatory or voluntary quotas for women. In spite of increasing visibility of women in leadership roles, women's decision making roles within household remain limited in many places in the world. In intra-household settings, individual bargaining power is crucial for positions and control over decisions that influence resource allocation and individual well-being. In many countries, especially the ones with patriarchal values, women voices remain powerless, ignored and/or stopped within the family set-up. For almost every decision in her life, a women need to seek permission from her husband, father, brother or someone elder in the family, and due to this, women's autonomy, freedom and economic independence have been greatly compromised. It is also important to mention that women's bargaining power within household matters not only for their own independence and well-being but also for the children they have, and women's lower say within household creates further inequalities in access to resources, health and education for themselves and their offspring. On this note, the second chapter of my thesis deals with unequal bargaining status of women within household. Apart from resource holding, education and income, women's fertility output plays a non-trivial role in their status within households; especially in patriarchal societies where having at least a son is important for a family. In Indian society a mother of a son earns respect and can be imagined to enjoy higher status in the family (and society) whereas mothers of daughter(s) and without having a son, are often ridiculed and even face violence within household. Therefore, the

second chapter examines the impact of women's fertility output, especially having a son on women's intra-household say in different decision-makings in Indian families.

For women, households have been paradoxical places, women who are taken/given the role of care-giver and homemaker and who may have sacrificed her career and passion in taking care of husband, children and elders, surprisingly, they themselves struggle with access to resources, face discrimination and often violence in the same household. Violence towards women within household can be a cause and/or consequences of inequalities and discriminations within household dynamics. When most of the families prefer to have son(s) instead of daughter(s), it does not come as a surprise that those locations will have a pro-male biased child and adult sex ratios. A strict and wide adherent to preference for sons in the society would result in higher number of single men due to shortages of women to pair up with these men. There are many evidences in the literature that young single men are more crime prone in general. Speaking of Indian society with a pro-male biased sex ratio, it not only records higher crime in general, but a disproportionally higher increase in crime against women. The coexistence of lower number of females and higher crime rates against them puzzles us as it contradicts the conventional economic theory which dictates that scarcity raises value.

The third chapter is a macro-study of demographic imbalances in sex ratios and its impact on women's safety and security. According to a World Health Organization report in 2013, globally, around 35 per cent women have experienced some type of violence in their lifetime. Violence against women has serious implication on their life, can widen gender imbalances further in many aspects. A society where women is not safe going out due to higher crime incidences against them, it creates further roadblocks in women's life, inhibiting them to go to schools and colleges, to go to work or even to go meet relatives and friends, and therefore, will further emphasize gender disparities in education and occupation, freedom and autonomy. The imbalances in sex ratio in favor of male have adverse consequences to society as a whole, by disbalancing the marriage market, supplying large number of single men and increasing crime rates. Preference for son has deep historical roots in India and is evident in its highly skewed sex ratios. But scarcity

of women did not automatically result in as increased value of women in Indian society, rather due to larger number of single young men, women are treated in unfair and violent ways. The third chapter explores this puzzle and examines whether this coincidence of pro-male biased sex ratios and higher crime on women bears any empirical relationship in case of India.

Gender equality plays an important positive role in economic performance of a country. Alternatively, countries that have higher, persistent and multiple gender inequalities, suffer from lower growth prospects and lower development. Not only at macro level, but gender disparities within household also have adverse outcomes. Unequal access to resources, education and voices in household decision making can be detrimental for people who suffer from lower access compared to other members and these also affect the future generation adversely. For example, a girl whose schooling is stopped early, has more likelihood to get married early and start having children early, and may remain absent from labor market, primarily because of lower opportunities due to lower education, or because she is not allowed to work and/or does not have a role in deciding whether to work or not. Girls in such situation are more vulnerable to violence and discriminations. Improving girls' participation in schooling system would prepare them for higher education, enhance their opportunities for labor market participation, and improve their value within household. India, with a very large percentage of youth population, the full potential of demographic dividend can be realized only when gender inequalities in economic and social participations are reduced. The former Secretary General of the United Nations and Nobel Prize winner, Kofi Annan said, "*Gender equality is more than a goal in itself. It is a precondition for meeting the challenge of reducing poverty, promoting sustainable development and building good governance.*"

10.14754/CEU.2019.02

CHAPTER 1

An economic analysis of gender gap in household demand for education: Evidence from India[1]

1.1 Motivation

Education plays a crucial role in building tomorrow's human capital, and thus access to education is considered as an important tool for economic growth and development. In 2000, by recognising this importance of education, the second UN Millennium Development Goal (MDG) was directed towards achieving universal school education by 2015. As from the past it was evident that girls are not going to schools as equally as boys, various supply initiatives were undertaken with additional emphasis to close the gender gap in education. More attentions were given on initiatives to address the special health needs of adolescent girls, safety concerns and ease in accessibility to schools from the neighborhood. The third MDG of promoting gender equality and empowering women includes the target of elimination of gender disparity in primary and secondary educa-

[1]I thank Andrea Weber for valuable advice.

tion, preferably by 2005, and in all levels of education no later than 2015. Despite of the widely taken global initiatives to achieve universal access to education, the gender gap in education still exists in many parts of the world, especially in developing regions. Table 1.1 shows the percentage of children at school-going age (6 - 17 years of age), who

Table 1.1: Out-of school rate among children (in per cent), 2014

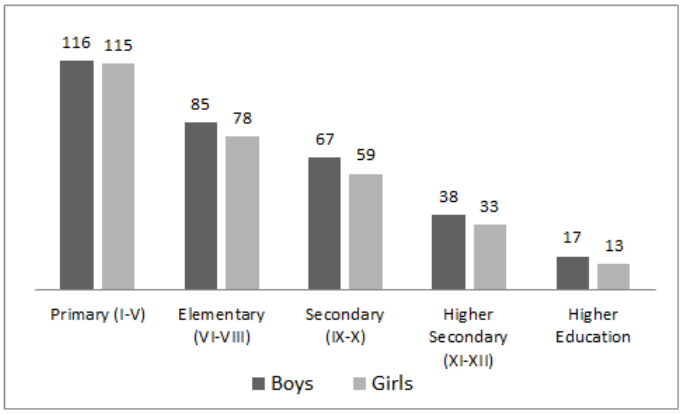
Regions	Primary 6-11 years		Lower Secondary 12-14 years		Upper Secondary 15-17 years	
	Male	Female	Male	Female	Male	Female
Caucasus and Central Asia	5.4	6.1	3.7	4.6	16.9	15.5
Developed Regions	3.4	2.9	2.2	2.0	8.5	7.3
Eastern Asia	3.0	3.0	6.6	6.6	21.2	12.4
Latin America and the Caribbean	6.2	5.8	8.0	7.3	24.9	24.9
Northern Africa	1.0	0.7	1.0	0.7	23.8	26.6
Oceania	8.9	14.3				
South-Eastern Asia	5.3	5.1	15.2	13.1	36.8	36.5
Southern Asia	5.8	6.8	21.4	17.7	48.7	51.2
Sub-Saharan Africa	19.2	23.3	31.8	36.5	54.6	60.8
Western Asia	7.8	13.7	13.1	20.0	29.8	35.5
World	8.1	9.7	16.0	16.0	36.9	37.5

Source: UNESCO Institute for Statistics database (UNESCO, 2014).

remain out of school in different regions of the world. The UNESCO data tell us that in Sub-Saharan Africa and Southern and Western Asia higher percentage of girls are out-of school compared to boys and this gap increases as the level of education rises.

Followed by MDGs in 2000, extensive government initiatives with special emphasis on girls were undertaken in India and recognition towards importance of universal education was reflected in its plans, programs and policies. These initiatives brought impressive progress to bring almost all potential pupils at the age-group of primary level (6-10 years of age) to school. In 2009, the Parliament of India enacted the Right of Children to Free and Compulsory Education Act or Right to Education Act (RTE). The act was implemented in 2010, and incorporated *free and compulsory* education to all children up to age 14, that is the age of completing elementary education ideally. Under RTE, *'free education'* means that no child (other than a child who has been admitted by his or her parents to a private school) shall be liable to pay any kind of fee/charges/expenses which may prevent him or her from pursuing and completing elementary education. *'Compulsory education'* indicates an obligation to the Government and local authorities to provide and ensure admission, attendance and completion of elementary education by all children in the 6-14 years of age group.

Figure 1.1: Gross Enrollment Rate at different education levels (2011)- India



Source: Ministry of Human Resource Development, Government of India (MHRD, 2012).

In 2011, the gross enrollment rate (GER) at primary level (I-V) is 116 for boys and 115 for girls[2] and at elementary level (VI-VIII), it is 85 for boys and 78 for girls (Ministry of Human Resource Development, Government of India, 2011). Starting from elementary (VI-VIII), secondary (IX - X) to higher secondary (XI - XII) levels, the gender gap in enrollment persists and widens with the level of education (Figure 1.1). Beyond enrollment in schools, attendance rates and learning outcomes even at the primary level are still questionable. Overall, the girls in India still lag behind boys in terms of literacy, enrollment, attendance, retention and learning at different education levels. Therefore it remains a concern that despite the enhanced infrastructure and policies to improve supply in Indian education system, girls still do not continue schooling beyond primary level as equally as boys. If supply side initiatives are adequately adding up to the access to education universally for all children, we need to look at the demand side factors within household that may have a gender discriminated demand for schooling.

The Nobel laureate economist, Amartya Sen once mentioned that policy making towards Indian education system requires "...the analysis of the characteristics of the economic and social forces operating in India, and response of public policy to these forces" (Sen, 1972). The goal of ensuring access to education for all does not automatically mean use of

[2]GER can exceed 100 per cent as it includes students, who are early and late entrants and also students who are in grade repetition.

education system equally by all. It is crucial to identify the constraints within household that inhibit full enrollment beyond primary level and to analyze the link between economic and social fabrics which creates these constraints that are responsible for the low rate of usage among girls compared to boys.

This paper attempts to identify the underlying demand side factors that keep girls out of the post-primary education system and/or attaining schools beyond the compulsory elementary levels. This paper addresses the following questions:

- Does the demand for school education differ between boys and girls? If so, how big is the gap?

  The household demand for schooling of children primarily depends on parents' preferences and decisions.

  - Do parents prioritise son's education over daughter's education?
  - Which factors are responsible for gender gap in parents' demand for children's education?

The paper provides a detail analysis of the household level factors that are responsible for the gender gap in enrollment at school education and also explains the policy implications of the findings that can help improving usage of school education system, especially beyond 'free and compulsory level' of schooling universally for girls and boys.

The paper is organised as follows: Section 1.2 reviews relevant theoretical and empirical literature on the topic; Section 1.3 gives theoretical background for the empirical model that will be used for analysis; Section 1.4 details of estimation methodology and data used; Section 1.5 presents the results as well as its robustness tests; and Section 1.6 concludes with discussion on results and relevant policy implications.

1.2 Literature Review

Research on human capital and labor market productivity have identified schooling years as important determinants of wage/earning. From household perspective, per-

ceived/expected returns from education motivate parents to spend on children's education. A paper by Jensen (2010) examines the importance of returns to education in schooling decisions. He used survey information on perceived knowledge about the returns from education from eighth-grade boys in the Dominican Republic and found that when randomly selected school students were made aware of the higher actual measured returns, it leads to 0.20–0.35 more years of schooling on average for the aware students over the next four years than those who were not aware. O. P. Attanasio and Kaufmann (2014) and O. Attanasio and Kaufmann (2009) investigated the role of expected returns to schooling and related risks as determinants of schooling decisions in Mexico and found that mothers' and youth's subjective expectations play crucial role in decision to enter college and continue high school.

The returns to education and parents' demand for child's education are actually linked due to parent's expectation that the child will grow up to an earning individual and then will reciprocate by providing old-age care when parents will retire from job market. Parish and Willis (1993) highlight that parents' altruistic behavior leads to investment in child's education in Taiwan. Alderman and King (1998) discuss the possible sources of gender disparity in parental investment on children and claimed that such disparities in investment can come through differences in returns realised by parents that is the expectation of future transfers from children to parents even when market returns to children themselves do not differ.

Greenhalgh (1985) discussed that patriarchal norms and parents' preference for sons in Taiwan are responsible for different treatment towards girls' education compared to boys. The author also mentioned that parents often send their girls to work due to resource constraints within the household and also to generate resource for brother's higher studies.

Using 1985-86 Peru Living Standards Survey, Gertler and Glewwe (1992) showed that parents perceive lower net returns to education for girls which leads to lack of parental desire to invest in daughter's education compared to son's education. Similarly, Gandhi Kingdon (2002) mentioned that parent's gender preference and thus differential treatments to sons and daughters lead to gender gap in education in developing countries like India.

A large number of literature have highlighted that higher birth order, sibling composition and large family size are responsible for lower usage of education in developing countries (Gomes (1984) - Africa; Knodel, Havanon, and Sittitrai (1990) - Thailand; Pong (1997) - Malaysia; Shreeniwas (1993) - Malaysia; Greenhalgh (1985) - Taiwan; Lillard and Willis (1994) - Malayisa; Parish and Willis (1993) - Taiwan; Black, Devereux, and Salvanes (2005) - Norway; Knodel and Wongsith (1991) - Thailand). Black et al. (2005) examine the effects of family size and birth order on the educational attainment of children using a dataset on the entire population of Norway and find a negative correlation between family size and children's education, but instrumenting for birth order or twin births the family size effects become negligible. Additionally, they showed that higher birth order has a significant and large negative effect on children's education.

The study by Knodel and Wongsith (1991) shows that family size has a significant negative impact on the probability of secondary school enrollment among children in Thailand as family resources per-child decrease when number of children increases.

Many literature have also shown marriage and related age is responsible for girl's dropout from formal educational institutions. Hill and King (1995) discuss about the barriers to female education. Marriage prospects can encourage or discourage girl's education depending on the relationship they hold between them. Also, social customs like patrilocality, and seeing women as primary care-giver and/or home-maker discourage parents to invest in girl's education as equally as in boy's education. Parish and Willis (1993) also show that elder or eldest daughters are taken out from schools earlier and married away earlier in Taiwan. Cochrane, Mehra, and Osheba (1986) show that parents' education has stronger influence on children's education in Egypt and educated parents attach higher value to education and more likely to educate their girls as similarly as boys.

Bommier and Lambert (2000) found that in Tanzania particularly, boys and girls follow fundamentally different patterns of schooling due to different returns from pre-school training in the family's economic activities or marriage prospects of girls. Their model predicts that when school quality decreases or schooling costs increase, parents send their children to school at a later age and for a shorter duration. Specifically for girls, despite

they attain lower level of education than boys, girls enroll earlier possibly due to relatively lower returns to pre-school experience for girls than for boys. It could also be the sign of interactions between education and marriage decisions that parents are eager to make their daughters ready for marriage prospects as soon as possible.

Studies also found that children's schooling varies with household wealth and location type (i.e, rural or urban). Mauldin, Mimura, and Lino (2001) explore the factors and amount related to parents' allocation of money for children's primary and secondary education and found after-tax income, parent's education, region, age and race are important determinants to decide the allocation of parents' money on children's schooling.

Using probit models, Glick and Sahn (2000) investigate gender gap in schooling indicators, such as grade attainment, enrollment, and drop-out from school in urban West Africa and found that an increase in household income leads to greater investments in girls' schooling but have no significant impact on boys. Education of father improves schooling of children of both gender, however, mother's education has significant impact only on daughters' schooling. Opportunity cost of schooling and increasing domestic responsibilities such as taking care of very young siblings have strong negative impact on girls' education but not on boys' schooling in India (Pal, 2004).

A large volume of literature in human capital, labor and education has identified either a factor or factors in combination that are responsible for gender gap in education. However, so far as best of my knowledge, no existing literature provides a holistic structure to household demand for school education. This paper contributes to this gap in the existing literature by identifying the fundamental factors that generate parents' demand for schooling of a child and combines the demographic and economic factors of the household that may influence the household schooling demand. The paper provides a simple theoretical framework to household demand for education and further derives comparative statics on various demand side factors. Using Indian data, the paper gives a general and also gender-disaggregated measurement of the contribution of all these factors in demand for schooling. After separating out the impact of the identified determinants, it also measures the inherent gender gap in parent's decisions of schooling. In addition,

it investigates the cultural factors that are prevalent in Indian society that may have contributed to gender gap in school education.

1.3 Theoretical Background

1.3.1 Household Demand for Education:

The household demand for schooling of children primarily depends on parents' preferences and choices. However, beyond schooling, tertiary (college) education is a combined decision taken by both parents and children. Without government's education subsidy, parents are solely responsible to bear the cost of school education and to decide whether to enroll, how long to keep the children in school, or discontinue a child's schooling. After completion of school education, students often take up part-time jobs to finance (fully or partially) their own education and also play a crucial role in decision of whether continuing education further or not and in which specialization. Beyond school education, perhaps parents and children together bear the expenses. As this paper focuses on parent's decision making on children's education, I would consider only school level education as at such levels parents are the primary decision makers. The crucial responsibility of educating the offspring relies upon parents choices about sending whom to which school and till when. Decision on children's schooling have both consumption and investment purposes.

When value of education is positive, parents would like to provide schooling to children as it feels good to have educated successful (from labor market and earning perspectives) children. The consumption motive behind schooling of children depends on preference for other goods and services that is how much parents value child's education compared to other goods and services. Parents' schooling decision for children can also be considered as an investment component as it requires to bear the cost (both direct and indirect) of schooling currently and gets return in the future in terms of old-age care from grown-up children. Parents' personal monetary benefits from the investment in child's education come from transfer of funds as financial support from grown-up child when he/she starts

earning in the future and parents retire from job.

The schooling decision is constrained by the household income. I can write this as utility from enrollment of child i at level S:

$$U(E_{Si}) \text{ is constrained by } I = C + \sum_{i=1}^{K} T_i$$

where U stands for utility; E_{Si} implies enrollment of child i at level S; I represents household disposable income; C denotes consumption of any other goods and services; K is the total number of children at school-going age in the household and T_i denotes total expenditures on schooling of child i. $\sum_{i=1}^{K} T_i$ denotes total household expenses on K number of children's school education. The household decision towards children's education depends on the current expenses (T) required to send children to school, that is the cost of schooling, both direct (tuition fees, transport cost to school, uniforms, books and stationary) and indirect (opportunity cost of children's schooling hours) costs. As this research considers only school education, any household expenditure on college education or higher education of children can be considered as part of consumption (C). The decision of parents towards a child's schooling depends on the utility (U) gains from choosing one option (i.e, to continue the child's enrollment) over another (i.e, to discontinue his/her schooling). Rational parents will keep sending their child to school if and only if the utility gain from sending him/her to school is higher than the utility gain from not sending him/her to school. And parents will not send a child to school when utility from sending to school is lower than utility from not sending. That is parents will,

- Continue child's schooling if $U(E_{Si} = 1) - U(E_{Si} = 0) \geqslant 0$
- Stop child's schooling if $U(E_{Si} = 1) - U(E_{Si} = 0) < 0$

where E_{Si} takes value 1 if child i is enrolled at S and 0 otherwise.

Based on parents' incentives to educate a child, the utility from providing school education depends on the expected remuneration of working in future and on the probability that the grown-up earning child will take care of retired parents.

Therefore, after controlling for economic and demographic characteristics of households, the incentives to send a child to school depend on parents':

- perception towards returns from education in future when the child will start earning; and
- expectation that the child will reciprocate in terms of providing old-age (economic and social) care to parents.

Parents are more likely to keep a child in school for longer years when the return has a positive relationship with years and levels of education. However, the future returns from education cannot be observed at current times and parents' perceptions towards future earning from a level of education are formed from the information on current actual wages in the known circle (family members, relatives and people in the neighborhood) with that level of education. The information set includes not only the distribution of wage rates for different education levels but also the associated risks in earning and access to opportunities. Therefore, the perceived returns to education are estimated from the wage distribution and also take into account variability in the distribution. So I will assume that the expectation of returns to education is formed by the entire distribution of current actual wage returns for different education level in the neighborhood. The effect of this distribution can be summarized by its moments.

The expected average return from an education level S is defined as discounted difference between average (expected value) wages at education levels (S) and $(S-1)$. I categorize schooling years into different levels, such as *below primary* (I-IV), *primary* (V), *below secondary or elementary* (VI-VIII), *secondary* (IX-X) and *higher secondary* level (XI-XII). The returns to education across these levels are then defined as:

$$ER_{Si} = \frac{W_{Si} - W_{(S-1)i}}{(1+r)^t} \tag{1.1}$$

where,

ER stands for expected returns from education;

W is the average wage of respective level of education;

S = below primary, primary, elementary, secondary, higher secondary;

$S-1$ = no schooling, below primary, primary, elementary, secondary;[3]

r is discount rate and

t is time in future when i will earn.

In addition, the returns to education are also attached with uncertainties related to matching and other labor market imperfections and can be measured in terms of variance (standard deviation) of the wage distribution of a particular level of education, $Var(W_{Si})$. Parents also recognize that the higher the earning of the grown-up children, the larger will be their capacity to provide old-age care to parents. If years of education positively influence its returns then it will also positively impact parents' reciprocity expectation. The expectation of old-age care that parents have from a child i (R_i), depends on social customs, feasibility and capacity of the child to provide economic and social support to parents at their old ages.

If continuing child's education is a component of parent's utility function, then this utility (U) from child's schooling can be explained as,

$$U(E_{Si}) = U(ER_{Si},\ Var(W_{Si}),\ R_i) \tag{1.2}$$

The relationship of the components in the right hand side of (1.2) that would generate utility for parents by taking enrollment decision are expected to be as follows:

- If expected return from education level S compared to level $(S-1)$ is positive then parents will be interested to continue the child's education into level S and won't

[3]

$$ER_{below\ primary} = \frac{E(W_{below\ primary}) - E(W_{no\ schooling})}{(1+r)^t}$$

$$ER_{primary} = \frac{E(W_{primary}) - E(W_{below\ primary})}{(1+r)^t}$$

$$ER_{elementary} = \frac{E(W_{elementary}) - E(W_{primary})}{(1+r)^t}$$

$$ER_{secondary} = \frac{E(W_{secondary}) - E(W_{elementary})}{(1+r)^t}$$

$$ER_{higher\ secondary} = \frac{E(W_{higher\ secondary}) - E(W_{secondary})}{(1+r)^t}$$

stop his(her) schooling after completion of level $(S-1)$ that is, $\frac{\delta ER_S}{\delta S} \geqslant 0$ leads to $\frac{\delta U(E_{Si})}{\delta S} \geqslant 0$.

- The variability of wages may have different impacts on enrollment decision. If the wage distribution of level S has higher variance that is higher uncertainties to get the returns (in terms of remunerations) or in getting opportunities, then parents will be discouraged to continue child's education in level S. But if the variability decreases with increase in education level that is if variability is lower in S compared to $(S-1)$, then parents will encourage the child to continue education for more years to have a more secured future, vice versa. It is also likely that parents are willing to take the risk as the returns are much higher for level S compared to $(S-1)$.

- And if parents expect reciprocal behavior from a child, whom they want to stay with and/or to get financial help from; then they will continue the child's education for longer years, given the positive relationship of returns with levels.

Apart from these fundamental components that contribute to parents' utility from a child's schooling, there are other factors that can influence the decision of schooling.

Parent's income and cost of schooling: The expenditure on children's schooling is constrained by parents' disposable income and plays an important role in schooling decision. If the schooling costs, direct or indirect are higher for higher levels, then parents will be less likely to continue children's education at higher levels given other factors and returns to education remain same. Poor parents with lower income level have to take out a child from school due to fund constraint even if they want to continue all children's education equally. It is also more likely for poor families to send children to schools where education costs are low if not free. The incentives to save can influence the decision of educating children and vice versa. An educated child when grown-up can earn and will be capable to provide old-age care to parents. If parents perceive so, then the motivation of savings for old age will be lower. To avoid this complexity in decision making I have ignored saving possibilities in this paper.

Family size and Sibling composition: Parents' decision to a child's schooling also depends on the family size and composition of children the parents have. Larger family and large number of children lead to division of household resources between more persons and the per capita resources available will be lower compared to smaller families, holding income constant. Children with higher birth order and higher number of siblings are more likely to drop out from school.

1.3.2 Household Demand for Education based on the Gender of Child

In many developing countries, presence of school-going age children in educational institutions differs vastly between boys and girls. In spite of extensive government initiatives; recognition towards importance of girls' education as equally as boys, and the promotion of universal access to education irrespective of gender; girls still lag behind boys in terms of usage of the education system. This paper intends to explore the source and dynamics of gender gap in school education.

If parents are biased towards a gender among children, such as if parents prefer son over daughter, then investment in education differs between boys and girls. However, even when parents are gender neutral, their demand for girls' education may differ from boys' education if any of the factors, such as the expected future returns from education, its variability and expected reciprocity differ between gender.

Labor market discrimination: Labor market opportunities differ between girls and boys. There is a considerable gender wage gap across occupation globally, male workers earn more compared to female workers with same level of education, experience and location. $W_{Sb} > W_{Sg}$, where b represents boys and g represents girls.

Also, I assume that the female wages are less elastic to schooling years compared to male wages, that is, $W_{Sb} - W_{Sg} \geqslant W_{(S-1)b} - W_{(S-1)g}$ or ,

$$\frac{\delta W_b}{\delta S} \geqslant \frac{\delta W_g}{\delta S} \quad \longrightarrow \quad \frac{\delta ER_b}{\delta S} \geqslant \frac{\delta ER_g}{\delta S}$$

Therefore, son's higher education is more beneficial compared to daughter's higher education and it is more likely that parents will discontinue daughter's schooling earlier than son's schooling. The prevalence of gender wage gap in almost any occupation leads to different investment (demand) functions for girl's education than boy's education. Then, it implies that,

$$\frac{\delta U_b}{\delta S} \gtrless \frac{\delta U_g}{\delta S} \quad \longrightarrow \quad \frac{\delta Pr(E_{Sb})}{\delta S} \gtrless \frac{\delta Pr(E_{Sg})}{\delta S}$$

Further, if the riskiness in earning opportunities is higher for girls compared to boys with same level of education, parents will prefer to continue boys education longer compared to girls. But, it also can be the case that women with lowest education level is less likely to or can never reach the higher income level due to lower opportunities compared to men with lowest education level. Similarly, women with highest education level also may not be as successful as men with highest education level due to labor market discrimination in hiring, remuneration differences and societal gender stereotype role of men and women. It is not always true that $Var(W_{Sb}) \leqslant Var(W_{Sg})$.

Traditionally, due to gender-stereotype views of society, women are seen as home-makers and caregivers whereas men are seen as primary bread-earners. This perception is much more evident in patriarchal society. In such societies, the value of women within household is generally measured by their efficiency in managing home and taking care of children and/or elders; and to perform these roles of women, parents may consider education as irrelevant. In recent times, though larger numbers of women are joining the labor force, the labor market structure still contributes to and reinforces gender stereotype in the society. The hiring, remuneration and promotion strategies often favor the male workers over the females and during recession time companies lay off female workers first, as it is considered that job-loss of a male worker will be more harmful to family than job-loss of a female worker. The situation is worse in the informal sectors where workers are primarily school educated or low educated. Due to the society assigned gender roles, women often take breaks from labor market participation due to marital, reproductive and nurturing responsibilities. Thus employers consider women as less loyal and reliable, and this leads to preference towards male candidates over females in responsible positions and offering

of remunerations.

Patri-locality and providing old-age care to parents: One of the primary incentives to provide schooling to a child can come from parents' expectation that the child will reciprocate by providing old-age care to parents in future. However, there are uncertainties attached to this reciprocal behavior. The probability of providing old-age support to parents is low among girls than boys, especially in patriarchal and patrilocal societies. In such societies, daughters are married away to live with in-laws family, whereas married sons stay with parents. Therefore, married daughters will get lower opportunities to take care of own parents compared to married sons. Due to increasing migration towards cities and even foreign countries for getting better job opportunities, it becomes common that retired parents stay away from sons and receive only remittances for their monetary needs. Physical presence of grown-up children whether sons and/or daughters with parents has become less likely than earlier. It remains difficult for married daughter to provide even monetary support if husband and in-laws don't allow her. Therefore, on average parents bear less expectation from daughters in regard to physical and monetary support at their retired age compared to sons. This perception provides lower incentives for parents to continue daughter's education longer as similarly as son's. Therefore, as

$$R_{Sg} \leqslant R_{Sb} \quad \text{then} \quad Pr(E_{Sg}) < Pr(E_{Sb})$$

Household income (I) and cost of schooling (T): Household or parents' income influences the decision of schooling, such as poor families either send their children to free school or schools with lower cost, or choose between children for schooling. Due to fund constraint, if parents have to choose between children's schooling as they can't afford everyone's schooling, it is more likely that parents stop girls' schooling and continue boys' schooling. Parents' decision for a child's schooling may also differ if the cost of schooling is different for boys and girls. Controlling for economic and demographic factors, the direct cost of education for girls and boys in a household is likely to be same, but the indirect cost of education may differ between boys and girls. After a certain age, especially adolescent girls are expected to take up some of the household responsibilities, such as helping

mothers at household chores, taking care of younger siblings etc. The adolescent boys are not asked to take such household responsibilities in general. If,

$$T_{Sg} \geqslant T_{Sb} \quad \text{then} \quad Pr(E_{Sg}) < Pr(E_{Sb})$$

that is if it is more costly for parents to send the daughters to school compared to sons, there will be higher probability that girls education will discontinue compared to boys.

Parents preference for son: If parents have a preference towards boy child over girl child then parents will perceive $U(E_{Sb}) > U(E_{Sg})$, and this leads to lower probability to continue girls' schooling compared to boys that is parents' preference for boys will emphasize the inequality as $Pr(E_{Sg}) < Pr(E_{Sb})$.

1.4 Data and Estimation Methodology:

1.4.1 Data

For the empirical analysis, I primarily use the second round dataset of India Human Development Survey (IHDS), published in 2012. The first round of IHDS data was published in 2005. IHDS 2012 is a nationally representative, multi-topic survey of 42,152 households and 204,565 individuals in 1503 villages and 971 cities across India. The survey has both household and individual level information on income and employment; consumption and standard of living; household and family structure; education; marriage and gender relations; fertility and health; and social and cultural capital.

Enrollment: The data has information on the current enrollment status as in terms of whether an individual goes to school currently or does not. The binary enrollment variable is constructed as if a child is enrolled in school currently, gets value 1; if a child went to school earlier but is not enrolled in school currently, gets value 0; and also if a child has never been enrolled in school also gets value 0.

Sibling composition: The household and individual data has information about the birth history of the children in the household, and using these information the sibling compo-

sition for a child is generated. The analysis uses the information of number of siblings and the number of male siblings a child has.

Gender of the child: Gender of the child is main explanatory variable in this paper, as the main purpose is to measure gender gap in school enrollment. The gender variable is constructed as if the child is female, she gets value 1, and otherwise gets value 0.

Labor market variables: Children's education decision by parents depends on their perceived monetary returns (ER_i) from spending on child i's education for another year/level. Many research papers have used survey information on perceived future returns from education and modeled current education decisions conditional on such subjective information (Reuben, Wiswall, and Zafar (2017); O. P. Attanasio and Kaufmann (2014); O. Attanasio, Kugler, and Meghir (2011); Keane and Wolpin (1997)).

The IHDS data do not have information on expected returns from education. However, we can assume that parent's perception about the returns to education is formed by the information they have about current wage returns from different levels of education. The information set is a function of actual wage returns across education levels among the known-circle of people in the family, relatives and neighborhood. Returns to education from a level, say level S, is the difference between average wage at level S and average wage at level (S-1). For this, gender-wise average wages are calculated using the gender-disaggregated actual wage distributions in the locality for different education levels. Thus, conditional on earning money as wage, the expected returns to education are calculated based on education levels, locations and gender. Here, locations mean the primary sampling units (PSUs) in the survey and each PSUs were formed with randomly selected 150-200 households in villages and urban blocks.

Exploiting the distribution of actual wages, not only the returns but also the riskiness attached with the wage opportunities is also captured. Variability of labor market is calculated from the standard deviation of gender-disaggregated local wages for an education level, conditional on earning money as wage.

Parents' reciprocity expectation: The survey has asked questions on preference towards gender and expectation of old-age care. The mothers were asked questions as: *Who do*

you expect to live with when you get old? Would you consider living with your daughter when you get old? Who do you expect will support you financially when you get older? and *Would you consider being financially supported by your daughter?* I use these four questions and construct four reciprocity expectation variables as if a mother has expectation of reciprocity from any of her child then reciprocity variable (R) takes value 1, otherwise 0. Also, I create another reciprocity variable (termed as *recipro4* in the analysis) taking into account all four reciprocity information together, such as *recipro4* gets value 0 if none of the 4 responses takes value 1, *recipro4* gets value 1 if at least one of the responses takes value 1, *recipro4* gets value 2 if at least two of the responses take value 1 and so on.

Other covariates: Apart from these, the analysis takes into account other demographic and household level characteristics that may have influences on children's enrollment. The location variable is constructed if the household is situated in urban locality then gets value 1, otherwise 0. Household income and consumption data have very high range in scales and thus these variables are scaled down by dividing the income and consumption data and also calculated returns by 100,000. The logarithmic scaling is not used as in some cases calculated returns have negative values. In addition, child's age, parents' income and education (schooling years), religion and caste are used as other variables in the analysis.

1.4.2 Estimation Strategy

I use probability models to estimate the schooling decisions for children. Parents' decision to keep a child enrolled in school depends on their perceived difference in utilities from two alternative choices: utility from keeping the child in school minus the utility from taking-out the child from school. This difference in utilities can't be observed, instead we only observe the current enrollment status of a child. So I assume that rational parents have made the decision comparing the two alternative choices. Let Y_i^* represents the

unobserved latent variable and can be defined as,

$$Y_i^* = U(E_{Si} = 1) - U(E_{Si} = 0) \tag{1.3}$$

where,

E_{Si} is a binary variable that is whether child i is enrolled ($= 1$) currently in S or is taken out from school (enrolled $= 0$).

Based on this difference in utilities, parents keep their child i enrolled in school if $Y_i^* \geqslant 0$ or decide to take out child i from school if $Y_i^* < 0$. Therefore I can write this as,

$$\begin{aligned} E_{Si} = 1 \ \ \text{(enrolled) if} \quad Y_i^* \geqslant 0 \\ = 0 \ \ \text{(not enrolled) if} \quad Y_i^* < 0 \end{aligned} \tag{1.4}$$

The equation for estimation can be formulated as:

$$\begin{aligned} Pr(E_{Si} = 1|Z_i) &= Pr(Y_i^* \geqslant 0|Z_i) \\ &= Pr(\beta_0 + \sum_{n=1}^{N} \beta_n.Z_{ni} + \epsilon_i \geqslant 0) \quad \text{where} \quad n \in [1, N] \end{aligned} \tag{1.5}$$

where, Z_{ni} is the vectors of all regressors, ϵ is the error term. and,

$$\sum_{n=1}^{N} \beta_n.Z_{ni} \equiv \beta_1.G_i + \beta_2.Sb_i + \beta_3.ER_{Si} + \beta_4.Var(W_{Si}) + \beta_5.R_i + \beta_6.X_i \tag{1.6}$$

where,

G_i: If child i is a girl then takes value 1 and if a boy then gets value 0.

Sb_i: The composition of siblings includes two variables, the number of siblings child i has and the male siblings i has.

ER_{Si}: The expected returns from education level S that child i has completed and dropped out or the level child i currently studying.

$Var(W_{Si})$: Standard deviation of (neighborhood) wages of the corresponding education level S that child i has completed and dropped out or the level child i currently studying.

R_i: Parent's expectation from child i that he/she will take care of parents when they

retire.

X_i: Control variables such as age, urban or rural location, parents education and income/consumption of the household, religion and caste dummies.

As I want to measure the gender gap in education demand within household, the main explanatory variable (G_i) is gender of the child, that is whether the child i is a girl ($= 1$) or a boy ($= 0$). The main coefficient of interest for measuring gender gap in enrollment is β_1. The probability model with binary choice dependent variable (enrollment) can be estimated by probit model assuming that the unobserved determinants of enrollment after controlling for observed factors and the stochastic errors provide a normally distributed random disturbance.

As discussed in the theoretical part (section 1.3) of the paper, crucial factors that influence the household decision-making of schooling are returns from education, labor market variability captured in terms of standard deviation of wages across education level and parents' expectation of reciprocal behavior from children in terms of old-age care. Therefore, in the estimation I use these factors as predictors of enrollment. In addition, composition of children that the parents have also may influence the schooling decision of children. Parents with fewer children can provide better access to education to children compared to parents who have many children. Other demographic variables that may have impacts on the schooling decision of children can be age of the child, completed schooling years, mother's and father's education, location type (urban or rural), income or consumption of the household, and religion and caste/tribe dummies. After considering the main explanatory factors and demographic control variables in the estimation, β_1 measures how being a boy or a girl can make differences in parents' decision on child's education. Therefore, β_1 gives a measure of parents inherent gender preference in schooling decision.

Before including all these variables together in a regression, I examine the presence of multicollinearity by computing the correlations between the variables and using Variance

Inflation Factor (VIF)[4]. The presence of multicollinearity between the right hand side variables will reduce the precision of estimated results. The correlations between *age of the child* and *completed schooling years* is 0.85 and between *hindu* and *muslim* is 0.81, therefore these give evidence for the presence of multicollinearity. The VIF and $1/VIF$ values also confirm multicollinearity[5]. It indicates that neither age and completed schooling years and nor hindu and muslim should be used together in the same regression. To decide which variables are better to explain the model, Bayesian Information Criteria (BIC) and Akaike Information Criteria (AIC) are used. Based on the BIC and AIC, I decide to keep age of the child instead of completed years of schooling and keep muslim instead of dummy for hindu religion. Earning of parents creates a spending constraint to investment in education of children. The information of earning of the parents from different sources (remuneration from job, earning of business etc) can be used as income of the household. A proxy for earning can be household consumption. Using BIC and AIC, I decide to keep consumption rather than income (parents' combined earnings) of the household.

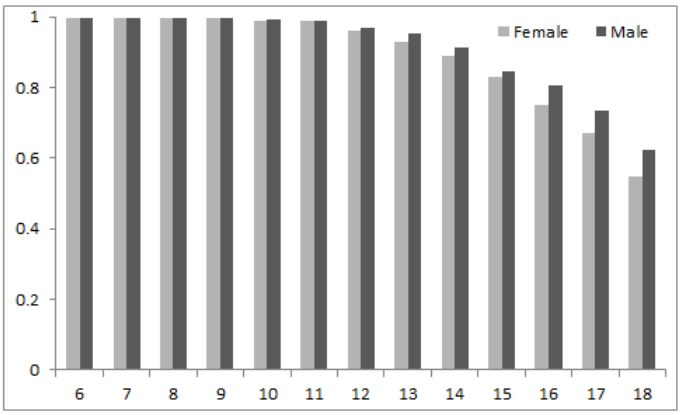
The IHDS survey includes information on all the household members. In India living as joint-family[6] is common till date, such as brothers and even cousins live in the same household with their own families and children. Then there would be cases where many children come from the same household, even in eldest and single children cases (where a parent is brother/sister to another parents within the same household). Therefore, the observations of children within the same household would be correlated and for variables, such as religion, location etc the observations would have the same values. The standard errors of all the estimations in this paper are clustered at household level.

[4]VIF value exceeds $1/(1 - Rsq)$ in presence of multicollinearity.

[5]VIF values are higher than 2.5 and $1/VIF$ values are less than 0.40 between *age* and *completed schooling years* and also between *hindu* and *muslim*.

[6]A joint family or undivided family is an extended family arrangement, which is common in Indian culture, and in this arrangement many generations, brothers and cousins even after marriage live in the same household, all bound by the common relationship.

Figure 1.2: Gender and age-wise enrollment of children at school-going age



Source: Author's calculation from IHDS 2012.

1.5 Results

1.5.1 Descriptive Statistics

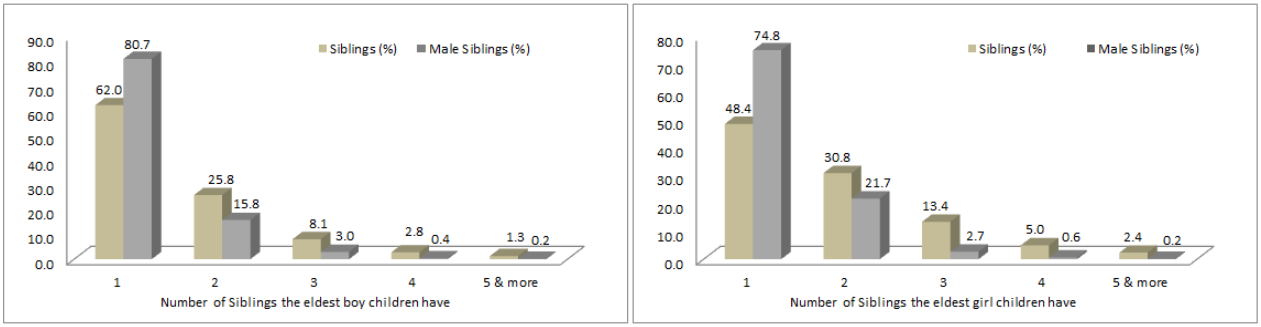
In the total sample of 2,04,568 individuals, 51,399 (25 per cent) are of school-going age that is between age 6 to 18 years. Among these school-going age children, 52 per cent are boy child and 48 per cent are girls. In these children, highest percentage (around 10 per cent) of children are of age 12 and the lowest percentage (6.5 per cent) children are at age 11 (Figure A1.1 in appendix).

Figure 1.2 shows the age-wise school enrollment rates among boys and girls. The enrollment of children at the age-group of 6 to 11 years is almost full (100 per cent), with marginally lower rates for girls at age 10. Starting from age 12 and onward, the enrollment rates start to diverge from the full enrollment, with higher difference for girls.

Ideally, at the age of 6 a child should start schooling at grade I, and complete grade I by age 7. Accordingly the ideal grade completion ages are as follows: 8 for II, 9 for III and so on. Therefore, children at age 11 should finish primary, at 14 should finish elementary (VIII), age 16 should finish secondary (X) and at 18 should finish higher secondary (XII). The data reveals that few children have finished the levels early than the ideal level-completion age[7]. There are large percentages of children who finished the

[7]1.8 per cent children completed primary level early; 1 per cent completed elementary level early;

Figure 1.3: Distribution of siblings and male siblings in eldest children sample



Source: Author's calculation from IHDS 2012.

level later than ideal age of completion.

Among 51,399 school-going age children, around 12 per cent (6,018) of the school-going age children are single children, who do not have any siblings. The number of eldest (first born children with siblings) children is 15,486 (30 per cent). Around 31 per cent children are without any male siblings, and 69 per cent with at least a male sibling. Figure 1.3 presents the distribution of siblings and male siblings that eldest children have. The percentage of eldest boys with one male sibling (80.7 per cent) is more than the percentage of eldest girls with a male sibling (74.8 per cent). It indicates that more parents on average want at least two sons. Similarly, the data on eldest child with two siblings indicates that eldest sisters have more male siblings (21.7 per cent) than eldest brothers (15.8 per cent) and so on.

The descriptive statistics of the children at school going age are given in Table 1.2. The mean age of school-going age children is 12 years. On average, the children have 2 siblings. Mother's mean year of education is 4 and father's education years is around 5 years on average. The average distance of schools is 2.7 km from home, and average annual schooling cost is 2112 INR (Indian Rupees). The negative values of cost mean that these children receive stipend from government or any other sources.

0.3 per cent and 0.1 per cent completed secondary and higher secondary levels early respectively.

1.5.2 Measuring difference in education demand between boys and girls:

To measure the difference in parents' decision in education, I use different subsamples of children, such as children in general at school-going age (is termed as *All Children Sample*); sample of first born children (is termed as *Eldest Children Sample*); and sample of children without any siblings (is termed as *Single Children Sample*). In table 1.3 for all estimations, the dependent variable is the current enrollment status of children that is whether a child i is currently enrolled in school ($E_i = 1$) or not ($E_i = 0$). To measure the difference in parents' decision-making based on the gender of the children, the main explanatory variable in all estimations is '*girl child*', the binary variable that is if the child is girl takes value 1 and if boy takes 0. The results in row 1 measure the marginal effects of being a girl child on probability of enrollment compared to being a boy[8].

Keeping the dependent and main explanatory variables same, different models include different sets of control variables: Model I includes number of siblings and male siblings; Model II adds return to education and standard deviation of wages; and Model III includes mother's reciprocity expectation, age of the child, urban/rural location, household consumption, both parents' education, muslim dummy and schedule caste/tribe dummy. Among 51,399 total children of school-going age, the observation included in the estimations of all children and eldest children are only 12049 and 3841 respectively, as the observations with missing values for variables are dropped to equalize the total number of observations in estimations with different set of controls and to make different estimations comparable. The exclusion of missing value observations also excludes children who have never been in school, and only keeps children who have some schooling and either continue schooling or have dropped out. The results without excluding the missing values are given in table A1.1 in appendix. The estimations in Table A1.1 also include

[8]Marginal effects represent percentage change in probability of enrollment due to discrete change of binary explanatory variables from 0 (being a boy) to 1 (for being a girl).
In case of the continuous explanatory variable, marginal effects indicate percentage change in probability of enrollment due to 1 unit change in continuous variables.

Table 1.2: Descriptive Statistics of school-going age children

Variables	Obs	Mean	Std. Dev.	Min	Max
Descriptive:					
Age	51399	11.58	3.48	6	18
Siblings	51399	2.06	1.37	0	9
Male Siblings	51399	1.07	0.96	0	8
Female Siblings	51399	0.99	1.03	0	7
Mother's Education (years)	48202	4.30	4.67	0	16
Father's Education (years)	43913	5.23	5.00	0	16
Household (HH) Demographic:					
HH Members	51399	6.29	2.71	1	33
Male HH Members	51399	3.08	1.61	0	17
Female HH Members	51399	3.21	1.71	0	17
Urban HH	51399	0.315	0.46	0	1
HH Income (INR)	51399	118136.6	230840.0	0	2439999
HH Consumption (INR)	51379	121089.2	105294.3	6000	4028836
School Demographic:					
Cost of education (INR)	41685	2112.5	9053.0	-13200	470900
School Distance (Km)	43825	2.75	5.41	1	99

Table 1.3: Probit Regression: Marginal effects on enrollment of children

Dependent Variable: Enrollment	All Children Sample			Eldest Children Sample		
Explanatory Variables:	Model I	Model II	Model III	Model I	Model II	Model III
Girl Child	-0.0218***	-0.0128*	-0.0364***	-0.0156	-0.0042	-0.0521***
	(0.0071)	(0.0073)	(0.0063)	(0.0120)	(0.0126)	(0.0108)
No. of Siblings	-0.0037	-0.0031	-0.0038	-0.0471***	-0.0462***	-0.0140***
	(0.0031)	(0.0031)	(0.0027)	(0.0056)	(0.0055)	(0.0047)
No. of Male Siblings	-0.0328***	-0.0326***	-0.0081**	-0.0199***	-0.0195***	-0.0014
	(0.0040)	(0.0040)	(0.0034)	(0.0076)	(0.0076)	(0.0062)
Return from Education		0.0168*	0.0147*		0.0062	0.0043
		(0.0086)	(0.0076)		(0.0143)	(0.0147)
Std. Dev. of Wage		0.0608***	0.0850***		0.0803***	0.0780**
		(0.0144)	(0.0155)		(0.0287)	(0.0337)
Reciprocity			0.0175			0.0010
			(0.0183)			(0.0260)
Age			-0.0436***			-0.0454***
			(0.0011)			(0.0020)
Urban			-0.0284***			-0.0279**
			(0.0072)			(0.0116)
HH consumption			0.0189***			0.0135*
			(0.0043)			(0.0076)
Mother's Education			0.0110***			0.0126***
			(0.0009)			(0.0015)
Father's Education			0.0069***			0.0077***
			(0.0007)			(0.0011)
Muslim			-0.0632***			-0.0741***
			(0.0078)			(0.0128)
Scheduled caste and tribe			-0.0062			-0.0123
			(0.0071)			(0.0119)
N	12049	12049	12049	3841	3841	3841

Note: Standard errors clustered at HH level are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
This table is re-estimated taking into account survey weights that is used in the survey sample design: See table A1.2

children who have no schooling with children who have some schooling.

Further, observing the characteristics of the missing values, it is found that from 51,399 children, 39,350 (around 75 per cent) children were dropped from estimations, and the highest number of observations (around 65 per cent) are dropped due to missing values in returns to education and standard deviation of wages. When returns to education and standard deviation of wages are constructed, a large number of missing values are generated as when no one (can be absent for a gender) in a location earn money with a certain level of education then missing values are created for such location or gender for an education level. As robustness check, I re-estimate the probabilities of enrollment by substituting the missing values of labor market variables with district and/or state average and the results are given in table A1.4 in appendix. The results for gender of the child remain robust in sign and significance, however, the size of the impact on enrollment declines in both all and eldest children samples. And, the impact of variability in wages increases when missing values of standard deviation of wages are substituted with district average and then further by state average. In addition, around 10 per cent observations are dropped due to missing values in other covariates.

Table 1.3 presents the marginal effects of all the right hand side variables with different set of control variables. The standard errors of the estimated effects of the variables are given in parentheses. The columns 1, 2 and 3 show estimated results from all children sample and columns 4, 5 and 6 present eldest children sample estimations.

All Children Sample: In Model I, including only sibling compositions and gender of the child, the results imply that girls are (significantly) 2.2 percentage points less likely to remain enrolled compared to boys, given the numbers of siblings and male siblings remain same. In Model II, the probability of enrollment of a girl child compared to a boy child remains negative but the size and significance of the impact decline. It may also mean that part of the gender gap in enrollment is explained by labor market variables, and thus Model I suffers from omitted variable bias. The labor market variables, return to education have positive and standard deviation of wages have positive and significant impacts on enrollment of a child. After controlling for the full set of variables, Model

III results indicate that a girl in general has 3.6 per cent significantly (at 1 per cent level) lower probability of enrollment compared to similar boys. Comparison of third model with the second model indicates that the size of the negative impact of 'being a girl' on enrollment increases in Model III. Examining the effect of each variables on enrollment separately, it is found that the major increase in the size of the gender impact is contributed due to the addition of child's age in the estimation. This can be examined in detail by using decomposition method later in section 1.5.6 and using interaction terms of different variables with girl child in section 1.5.7. Part of the increase in gender impact on enrollment in Model III can be due to the presence of unobserved heterogeneity in the data.

Eldest Children Sample: The results in table 1.3, columns 4, 5 and 6 present the results using eldest children only. In Model I, the results indicate that eldest sisters have 1.6 per cent lower (but statistically insignificant) probability of enrollment compared to similar eldest brothers. Addition of labor market variables (Model II) also keeps the enrollment probability of eldest sisters statistically indifferent from the probability of eldest brothers. However, with full set of control variables (Model III), it is found that an eldest sister on average has 5.2 per cent lower chances of enrollment compared to a similar first born boy. Comparing the all children sample with the eldest children sample results, I find that the gender (i.e, being a girl) impact on enrollment is higher in size for eldest children than children in general. Girls who are eldest among the siblings have higher probability of being withdrawn from schools compared to eldest boys than the probability of the same among general girls compared to similar boys. The marginal impact of 'being a girl' on enrollment status shows that being eldest daughter is worse than being a daughter in general in terms of decision that parents take for children's schooling.

One possible reason to eldest sisters' lower attendance in school can be their household responsibility and especially younger sibling-care in the household. Among 51,399 school-going age children, 13 per cent children have very young siblings; whereas around 16 per cent girls and 11 per cent boys have very young siblings. To examine the impact of 'having

very young sibling(s)' on schooling of children, *young sibling* variable is constructed as binary variable, that is children who have very young sibling(s), whose age is 3 years or below, get value 1, otherwise 0. The estimated impact of having *young sibling* on enrollment of children in general and also on enrollment of girls and boys are shown in table A1.5 in appendix. Having young sibling(s) can in general effect the enrollment of children in negative manner. The results from interaction terms indicate that a girl child with very young sibling has significantly 4.3 per cent lower probability of continue schooling and a first born girl has 3.8 per cent (insignificant) lower chances to continue schooling if she has very young sibling(s). Having young sibling does not influence boys' enrollment decision in a significant way. These results do not provide significant evidence for eldest sisters but the impact of having young siblings is higher and significant for girls in general.

The results in table 1.3 show the cases in general and when parents have to decide on schooling among children. It would be interesting to examine the gender gap among the children who have no siblings that is sample of single children.

Single Children Sample: However, the subsample of single children may have potential endogeneity issues, as the decision of number of children is not random. On the one hand, parents who have preference for a particular gender may have shorter family if they get the preferred gender composition of children earlier and on the other hand, couples may continue having more children until they get preferred child(ren). Therefore, parents' decision to stop having more children after having the first is not a random decision. The stopping decision can also be influenced by couples' economic and social life. Therefore, the single children sample may have potential endogeneity problem. I compare the characteristic of parents with single child and multiple children and find that single child parents are higher educated, earn higher income and have lower reciprocity expectation on average compared to parents who have multiple children (see table A1.6 in appendix). In this paper as the decision of child bearing do not have a direct importance to the main results, and as it is difficult to solve the endogeneity issue in decision of having child, I would estimate the single children sample results considering that the

Table 1.4: Marginal effects on enrollment of single children

Dependent Variable: Enrollment	Model I	Model II	Model III
Explanatory Variables:			
Girl Child	-0.00867	0.00414	-0.00572
	(0.0285)	(0.0289)	(0.0249)
Return from Education		0.0135	0.0048
		(0.0237)	(0.0245)
Std. Dev. of Wage		0.0818	0.0482
		(0.0497)	(0.0429)
Reciprocity			0.0703**
			(0.0358)
Age			-0.0419***
			(0.0049)
Urban			0.0158
			(0.0256)
HH consumption			0.0277
			(0.0198)
Mother's Education			0.0100***
			(0.0026)
Father's Education			0.0079***
			(0.0018)
Muslim			-0.0708*
			(0.0376)
Scheduled caste and tribe			-0.0109
			(0.0231)
N	875	875	875
Controls	Set 1	Set 2	Set 3

Standard errors clustered at HH level are in parentheses.
* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$.
This table is re-estimated taking into account survey weights that is used in the survey sample design: See table A1.3

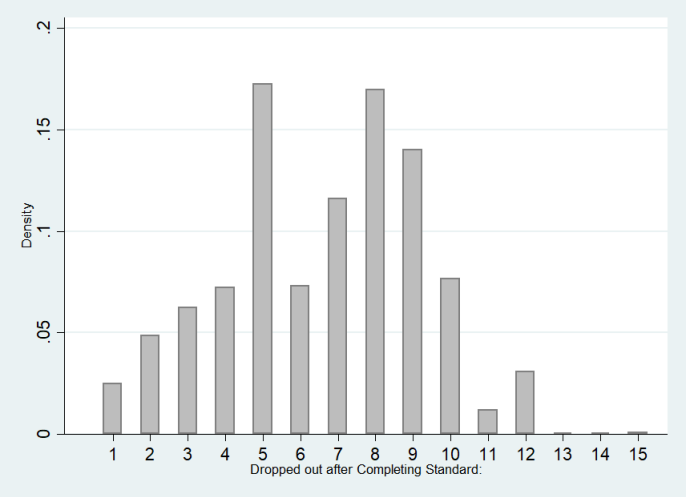
results are not with best precision in presence of the potential endogeneity.

Using single children sample, table 1.4 shows estimations of three similar models as in table 1.3. Among single children, the effects of being a girl child on enrollment are negative but statistically insignificant in the three models with different set of controls. It indicates that when parents have only one child, there is no significant difference in decision on schooling between parents who have girl child and parents with a boy child.

1.5.3 Gender difference in enrollment across levels of education

The examination of school enrollment across different schooling standards reveals that the drop-out rates are higher at the transitions from one level to another. Figure 1.4 shows that the drop-out rate is higher after completion of standard V or primary level compared

Figure 1.4: Drop out rate at different school levels



Source: Author's calculation from IHDS 2012.

to immediate pre and post standards, such as standards IV and VI. Higher percentage of children also drop-out after completion of elementary school (standard VIII) compared to standards VII and X. Though, the drop-out rate at secondary level is also higher in comparison to drop-out upon completion of standards XI and XII, but it is lower than drop-out after standard IX. The gender-wise drop-out rates are also examined, both have the similar patterns as figure 1.4, with higher drop-out among girls.

Table 1.5: Probit Results: Enrollment of children across different school levels

Probit Estimation -	Dependent variable: Enrollement		
Explanatory Variables: Girl Child	All	Eldest	Single
Panel A			
Completed Primary level (Std V)			
dy/dx	-0.0288	-0.0067	0.0308*
	(0.0192)	(0.0343)	(0.0166)
N	861	245	59
Completed Below Secondary level (Std VIII)			
dy/dx	-0.0375***	-0.0477***	0.0136
	(0.0101)	(0.0163)	(0.0463)
N	5675	1878	419
Completed Secondary level (Std X)			
dy/dx	-0.0667***	-0.0787***	-0.1144
	(0.0194)	(0.0305)	(0.0698)
N	1171	453	65
Panel B			
Children above age 14			
dy/dx	-0.0785***	-0.1127***	-0.0021
	(0.0169)	(0.0260)	(0.0627)
N	3962	1447	328

Note: Controls used in estimations are number of siblings and male siblings (excluded in single children sample), returns to educations, std dev of wages mother's reciprocity expectation, child's age, urban/rural, HH consumption, parents' education, religion (muslim) and scheduled caste/tribes dummies. Standard errors clustered at HH level are given in parentheses.
* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$.

With the evidence of higher drop-out rates after completion of levels, I decide to examine the influence of gender of the child on enrollment decision in the transition period from one level to another. For the estimations in Table 1.5 Panel A, I use three subsamples; children who have completed primary level or standard V, children who have completed elementary level or standard VIII and children who have completed secondary level or standard X; under all, eldest and single children samples. These estimations include the full set of control variables.

For pupils who have completed primary schooling, gender of the child do not significantly influence their enrollment decision in all children and eldest children samples. However, for single children the girl child has significantly (at 10 per cent level) higher probability by 3 per cent to remain enrolled compared to the single boys after primary level. After completion of elementary level, girls in general and first born girls are significantly less likely by 3.7 per cent and 4.8 per cent to continue schoolings compared to boys in general and eldest boys respectively. After elementary level, as schooling does not remain free and parents have to bear the costs, and then girls' education becomes less preferred to parents compared to boys' education. However, for single child the marginal effects of gender on enrollment after elementary level are insignificant and positive, that is the single girls do not differ significantly from single boys in access to schooling after completion of standard VIII.

At completion of secondary level, girls on average have 6.7 per cent lower probability to continue schooling compared to boys in general, keeping other things same. And, eldest girls have significantly 7.9 per cent lower chances to continue schooling compared to eldest boys after completing secondary school, *ceteris paribus*. The gender difference in enrollment decisions of single children remains insignificant even after completion of secondary schooling. From the level-wise analysis it is evident that the probability of girls' and eldest sisters' drop-out increases as the level of education increases, given all other things remain same. Parents will more likely discontinue daughter's education compared to son's education when schooling incurs costs.

As the implementation of RTE Act (2009-2010) ensures that every child upto age of 14

years has right to full-time free and compulsory education, parents become motivated to send children upto age of 14 years to free schools irrespective of gender of the children, but beyond this age the schooling demand decreases as schooling is not free anymore. To evidence this, I estimate the impact of gender on enrollment at this age threshold by using children above age 14 (Table 1.5 Panel B). It is found that among children above age 14 years, the gender impact increases and girls in general have 7.8 per cent lower probability to remain enrolled in school compared to similar boys, keeping other things same. Among the eldest children, girls have 11 per cent lower chance to continue schooling compared to similar eldest boys, given other things remain same. Both the impact on sample of all children and eldest children are statistically significant at 1 per cent level. The enrollment of girls, who do not have any sibling and above the age of 14, does not differ significantly compared to similar single boys.

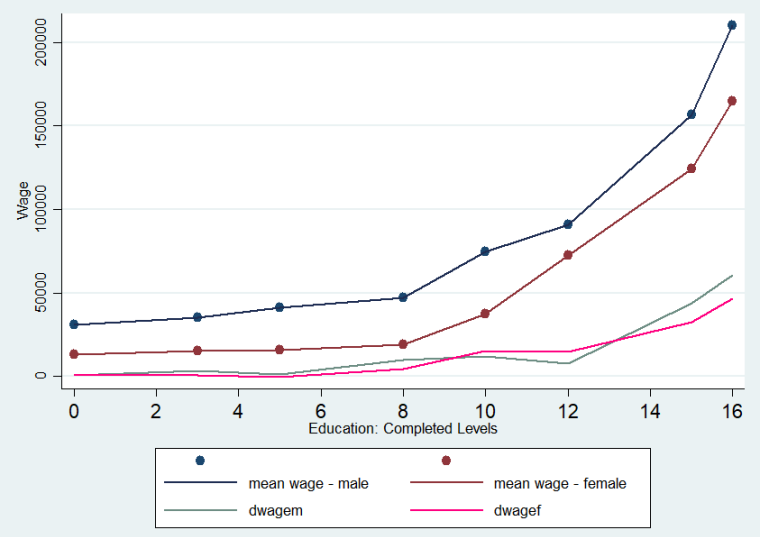
The results in panel B of table 1.5 thus imply that the probability of continuing schooling worsens among girls beyond age 14, and becomes worst for first born girls whereas parents' decision on schooling does not differ significantly between single boys and girls.

1.5.4 Labor market implications on enrollment

Jensen's (2012) paper shows that labor market variables play important roles for girls and improvement in employment and economic opportunities influence investment in girls' education positively. I want to examine empirically how labor market variables influence enrollment decision of children and the difference of these impacts on girls' and boys' school enrollment. Using the IHDS 2012 data, figure 1.5 shows that gender gap in wages persists at all levels of education, and $W_{Sb} > W_{Sg}$ is true in the data. However, it cannot be said that the return to education is always lower for females than males though in most parts it is so. At the secondary and higher secondary levels the growth in female wages expedites compared to males, then for tertiary education level female education return again falls below the male returns.

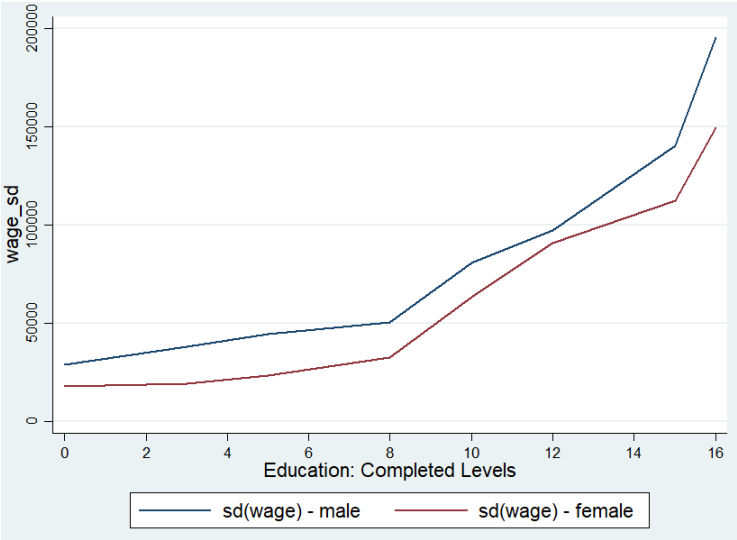
Figure 1.6 shows the standard deviation of wages across education levels and that it is lower for females than males. It can be explained as the females are less likely to even

Figure 1.5: Gender wage gap across education levels



Source: Author's calculation from IHDS 2012.

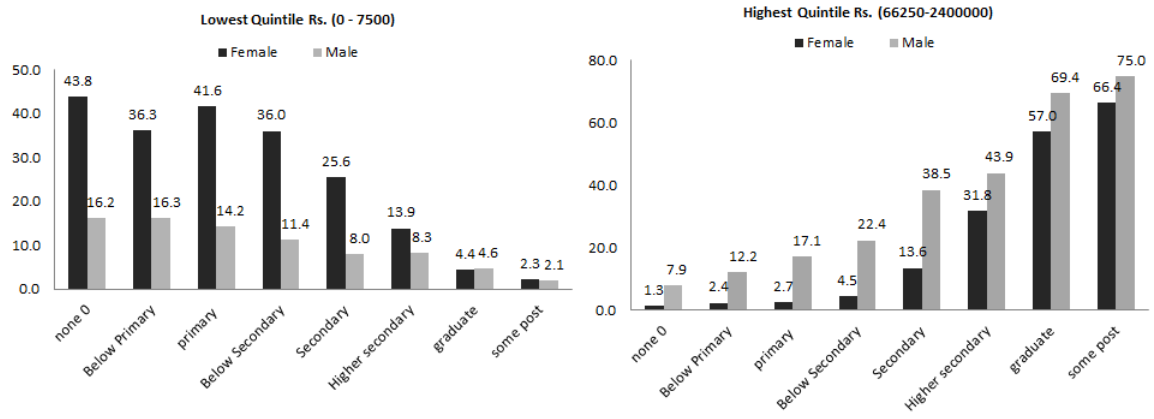
Figure 1.6: Gender-wise standard deviation of wages across education levels



Source: Author's Calculation from IHDS 2012.

apply for the highly paid jobs when they have lower education levels whereas their male counterpart with similar qualification are more likely to try their luck even in the highest paid jobs. Therefore, the variability in the male wages are higher not only because the inherent riskiness in opportunities in the labor market but also because males actually try and get highly paid jobs with comparatively lower educational qualification than females. Figure 1.7 also reinforces this argument as it shows that only 2.4 per cent females with below primary education reach the highest quintile of income distribution compared to

Figure 1.7: Gender-wise success in income quintiles across education levels



Source: Author's Calculation from IHDS 2012.

12.2 per cent of males with below primary education and only 2.7 per cent females with primary education reach the highest quintile of income compared to 17.1 per cent males with similar qualification. Similarly, among the lowest educated females around 36 per cent with below primary and 41.6 per cent primary educated females work with income in the lowest quintile whereas only 16.3 per cent below primary and 14.2 per cent primary educated males work with similar payments.

Table 1.6 shows the impact of labor market variables on enrollment of children and also how it impacts the enrollment of boys and girls differently. Using both return to education and standard deviation of wages as the explanatory variables and enrollment as the dependent variable, I perform probit regression in all, eldest and single children samples. Panel A results in the first row indicate that returns to education have positive impact on the school enrollment of children, and especially the impact is statistically significant at 10 per cent level in case of all children sample, keeping other variables constant. By using interaction terms between gender-disaggregated average returns with gender of the child, Panel A results in second and third rows represent the impact of male and female returns to education on enrollment of boys and girls respectively. I find that male returns to education have no statistically significant effect in boys' enrollment.

The marginal effects of female wage returns on enrollment of girls are positive and higher in size compared to the impact of male returns on boys' enrollment. For example, in all children sample estimations with full set of control variables (column 2), the results

indicate that 1 unit[9] increase in female returns to education improves the chances of girls' enrollment significantly by 9.2 per cent, whereas 1 unit increase in male returns can improve boys' enrollment probability only by 1.1 per cent (insignificant). For eldest and single children samples estimations, the results remain statistically insignificant though the size of the impact of female returns on girls' enrollment is much higher compared to impact of male returns on boys' enrollment.

In table 1.6, the marginal effects of standard deviation of wages on enrollment of children are shown in the first row of Panel B. Standard deviation of wages have significant and positive impact on the enrollment of all and eldest children, but statistically insignificant in single children sample when estimated with full set of control variables. Panel B's result

Table 1.6: Labor market impact on enrollment: $Pr(E_{Si}|ER_{Si}, Var(W_{Si}))$

Dependent Variable: Enrollment	All Children		Eldest Children		Single Children	
Panel A						
Return to Education	0.0138	0.0144*	0.0044	0.0041	0.0138	0.0029
	(0.0086)	(0.0076)	(0.0151)	(0.0146)	(0.0237)	(0.0248)
Using interaction term between Return to Education and girl child						
Return to education	0.0114	0.0112	0.0034	0.0033	0.0120	0.0030
	(0.0086)	(0.0077)	(0.0150)	(0.0149)	(0.0240)	(0.0248)
Return to education X Girl Child	0.1186*	0.0924*	0.0618	0.0204	0.0837	0.0886
	(0.0640)	(0.0488)	(0.1091)	(0.0778)	(0.1435)	(0.1815)
Panel B						
Std Deviation of wages	0.0788***	0.0858***	0.1092***	0.0782**	0.0800*	0.0531
	(0.0147)	(0.0155)	(0.0292)	(0.0337)	(0.0485)	(0.0437)
Using interaction term between Std Deviation of wages and girl child						
Std Deviation of wages	0.0709***	0.0848***	0.1021***	0.0794**	0.0750	0.0459
	(0.0153)	(0.0161)	(0.0313)	(0.0358)	(0.0500)	(0.0428)
Std Deviation of wages X Girl Child	-0.0139	0.0045	-0.0420	-0.0197	0.1939	0.1278
	(0.0589)	(0.0539)	(0.0933)	(0.0853)	(0.2925)	(0.2823)
N	12049	12049	3841	3841	875	875
Controls	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2

Note: No control variables are included in Set 1 except return to education and Std deviation of return. Set 2 further adds gender of the child and number of siblings (excluded in single children sample), child's age, urban/rural, HH consumption, both parent's education, religion and caste dummies. Clustered standard errors at HH level are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

rows 2 and 3 show the difference in impacts of variabilities in male and female wages on enrollment decision of boys and girls respectively. With full set of controls, probability of

[9]Return values are scaled down by dividing with 100,000, Therefore, 1 unit increase in return means increase of INR 100,000 in returns.

boys' enrollment improves significantly by around 8 per cent if variability of male wages increases by 1 standard deviation, using both all and eldest children samples. However, for girls (especially for eldest girls), the variability in female wages can reduce their enrollment probabilities, but the estimated impacts remain statistically insignificant. In case of single children sample, the impacts of variability in male and female wages are statistically insignificant on both boys' and girls' enrollment respectively. Thus, I can say that the parents with multiple children behave differently to variability of female and male wages when deciding for girls' and boys' schooling respectively.

The difference in parents' behaviour towards variability in male and female wages can be explained in the following way: Labor market imperfection, uncertainty of getting an opportunity to work and secure remuneration may influence schooling decisions in a positive way that is parents would like to continue children's education longer in the prevalence of such imperfections. Parents probably perceive that without completing school education it would become hard for a child to get a good job in the presence of labor market uncertainties, and that higher education gives more confidence to individuals to try their luck in highly paid jobs. However, this perception does not hold for parents with girl child, these parents rather get discouraged by the variability for females in the labor market. When parents already perceive the pro-male labor market policies and opportunities, higher variability for females reduces their incentives for girls' schooling further.

On the similar line, one question can arise that if female success stories are known, will it influence parents' mindset towards girls' schooling and parents will be more likely to keep the daughters in school in expectation of similar success for own daughters? Therefore, to examine whether the exposure to local females in top positions can influence parents' decision to girls' education differently, I use the information of gender-disaggregated maximum wage in a location for an education level. Maximum wage returns can be used as a proxy for success in the labor market. The estimated enrollment probabilities are given in table A1.7 in appendix of this chapter. The results indicate that the information of maximum wage return for an education level would not influence parents' enrollment de-

cision positively, but rather it may significantly decline the enrollment chances of children (Row 2 in table A1.7).

Using interaction term of gender-disaggregated maximum wage returns with gender of the child, I find that the maximum female wage returns can influence parents' decision for schooling of girls in general and for single girl child in a positive manner, but these results are statistically insignificant. For single boys, the information of maximum returns that a male with an education level can earn surprisingly discourages parents' decision of boy's enrollment. This can be intuitively explained as parents are more concern about uncertainties of labor market, few success stories do not improve their decision of schooling significantly. Also, from the perspective of reciprocity expectation, it can be argued that even if daughters are expected to be highly successful, still mothers do not expect much help from a married daughter due to patrilocality custom and also often parents are ridiculed if financial help is taken from married daughters when parents have a son. Thus, it can be expected that the information of female role models would not influence girls' schooling significantly.

Further, I examine the impact of labor market variables on enrollment of children across different education levels. Table 1.7 shows the marginal effects of two labor market variables on enrollment of children and how these impacts are different between girls and boys after completion of different education levels. In general, the impact of return to education remains positive (around 2 per cent) and insignificant after completion of primary and secondary level, but negative and insignificant upon completion of elementary level (Panel A row 1). After adding the interaction terms of gender dis-aggregated returns to education with gender of the child, the results remain insignificant for boys' enrollment that is 1 unit increase in male returns to education have no significant impact on boys' enrollment probability across completion of any schooling levels. Alternatively, the impact of female returns to education on girls' enrollment probability is much larger in size compared to the impact of the male returns on boys across the schooling completion levels. One unit increase in female returns to education increases the probability of girls' enrollment by as large as 18 per cent after completion of elementary level, keeping

all other things constant. However, the impacts of increase in female returns on girls' enrollment are insignificant after completion of primary and secondary levels.

In table 1.7, the first result row in Panel B shows that increase in variability in wages increases the enrollment probability of children in general, significantly by 10 per cent beyond primary level, significantly by 6 per cent after completion of elementary level and by 1 per cent but insignificant after secondary completion. The size of the effect of variability in wages on enrollment declines with increase in education level. This can be explained as at earlier level of education variability in wages encourages parents to continue children's education further, but this incentive declines as level of education increases. We have seen in data that variability in wages increases with increase in education years. Though wage variability of lower educated workers may encourage parents to continue children's education for longer years but higher variability of wages at higher levels of education may not have the same motivations for parents to continue children schooling further.

In estimations of rows 2 and 3 of table 1.7 Panel B, I use interaction terms of gender-disaggregated standard deviation of wages with gender of the child. The results reveal that one unit increase in standard deviation of male wages improves the chances to continue boys' schooling beyond elementary level significantly by 6 per cent, keeping other things same. However, due to 1 unit increase in standard deviation of female wages, girls' enrollment probability may decrease after elementary and secondary school, but these results are insignificant. It again indicates that labor market imperfections encourage parents of boys to continue their schooling for longer period to secure a better future for them. But parents of girls can be demotivated by labor market uncertainties and can discontinue daughter's education beyond elementary schools.

Further, I also examine the impact of labor market variables on children's school enrollment after the age of 14. On average, returns to education have positive (2.9 per cent) but insignificant influence on enrollment probability of the children beyond age 14 in general (Panel A row 1). Among boys above age 14, one unit increase in male's return to education increases boys' probability to continue schooling by 2 per cent though the

Table 1.7: Labor market impact on enrollment across education levels:

Dependent Variable:	After Completion of:			Beyond
Enrollment	Std V	Std VIII	Std X	Age 14
Panel A				
Return to Education	0.0215	-0.0027	0.0199	0.0292
	(0.024)	(0.012)	(0.020)	(0.0201)
Using interaction term between Return to education and girl child				
Return to education	0.0192	-0.0034	0.0160	0.0200
	(0.0212)	(0.0120)	(0.0205)	(0.0202)
Return to education X Girl Child	0.1762	0.1830***	0.0282	0.3443***
	(0.1109)	(0.0706)	(0.1118)	(0.1295)
Panel B				
Standard Deviation of wages	0.1053**	0.0648***	0.0110	0.2039***
	(0.055)	(0.019)	(0.018)	(0.0336)
Using interaction term between Std Deviation of wages and girl child				
Standard Deviation of wages	0.0518	0.0607***	0.0119	0.1971***
	(0.0495)	(0.0208)	(0.0187)	(0.0339)
Standard Deviation of wages X Girl Child	0.2558	-0.0045	-0.0172	0.1737
	(0.2728)	(0.1047)	(0.0690)	(0.1347)
N	862	5680	1171	3964

Note: Controls used in estimations are number of siblings and male siblings, returns to educations, standard deviation of wages, mother's reciprocity expectation, child's age, urban/rural, HH consumption, parents' education, religion (muslim) and scheduled caste/tribes dummies. Standard errors clustered at HH level are given in parentheses.
* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$.

impact is statistically insignificant (Panel A row 2). For girls above age 14, improvement in female returns to education by one unit has a very large impact on their enrollment, improves girls' enrollment probability significantly by 34 per cent, keeping other things same (Panel A row 3). In Panel B of table 1.7, the results imply that an increase in variability of labor market improves the probability of parents' decision in favor of continuing child's education significantly by 20 per cent in general and for boys, above the age of 14 years. Among girls of age beyond 14 years, the variability of female wages increases girls' probability of continue schooling by 17 per cent, though the impact is statistically insignificant. Education returns have a very large and positive impact on girls' schooling decision after age 14 years, however, the labor market imperfection has positive and highly significant impact on boys' schooling decision beyond age 14, but for girls the impact is not statistically significant.

1.5.5 Reciprocity and Enrollment

Parent's reciprocity expectation from children can influence their schooling decision. Parents spend in child's education with the expectation that the child will grow-up to an

Table 1.8: Impact of parents' reciprocity expectation on enrollment $Pr(E_{Si}|R_i)$

Dependent Variable: Enrollment	All Children Sample		Eldest Children Sample		Single Children Sample	
Explanatory Variable: Different Reciprocity Expectation Indicators						
Reciprocity from child	-0.0790**	-0.0101	-0.1321**	-0.0654	-0.0922	-0.0400
	(0.0340)	(0.0309)	(0.0673)	(0.0478)	(0.0900)	(0.0745)
Reciprocity from daughter	0.0201**	0.0215***	0.0223	0.0323**	0.0121	0.0173
	(0.0094)	(0.0078)	(0.0161)	(0.0129)	(0.0393)	(0.0317)
Reciprocity from child - Financial	-0.0140	0.0167	-0.0145	0.0010	0.0278	0.0703**
	(0.0221)	(0.0184)	(0.0341)	(0.0260)	(0.0479)	(0.0358)
Reciprocity from daughter - Financial	0.0106	0.0088	0.0119	0.0110	-0.0015	0.0068
	(0.0087)	(0.0073)	(0.0152)	(0.0124)	(0.0371)	(0.0292)
Recipro4	0.0066	0.0087**	0.0075	0.0115*	0.0160*	0.0032
	(0.0048)	(0.0041)	(0.0083)	(0.0068)	(0.0085)	(0.0165)
Parent(s) in govt job	0.0385***	0.0255***	0.0480***	0.0367***	-0.0108	-0.0095
	(0.0089)	(0.0077)	(0.0156)	(0.0127)	(0.0278)	(0.0268)
N	12049	12049	3841	3841	875	875
Controls	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2

Note: No control variables are included in Set 1 except reciprocity expectation of parents from child. Set 2 further adds gender of the child and number of siblings (excluded in single children sample), child's age, urban/rural, HH consumption, both parents' education, religion and caste dummies. Clustered standard errors at HH level are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

educated successful individual and will provide old-age support to parents in future. To examine the impact of reciprocity expectation on enrollment, probit regression is used where enrollment status of a child is the dependent variable and the main explanatory variable is reciprocity expectation. Reciprocity expectation is a binary variable, it takes value 1 if mothers have expectation from a child to provide old-age care, otherwise it takes value 0. Four questions were asked about the reciprocity expectation of mothers, such as (i) *with whom she expects to live with when she gets old?*; (ii) *does she expect to live with daughter when she gets old?*; (iii) *from whom she expects to get financial help in old age?* and (iv) *does she expect to get financial help from daughter when she gets old?*. The answer categories for questions (i) and (iii) are son, daughter, both or none. If a mother expects reciprocal behavior from her child, whether son or daughter or both gets a value of 1, otherwise 0. The answer categories of (ii) and (iv) are yes and no, and thus directly form the binary variable of reciprocity expectation from daughter. In the analysis, I use all these 4 variables separately in regression to examine the impact of reciprocity on enrollment of children. Apart from these, I also construct another variable *Recipro4* using information of the four reciprocity questions together, such as *Recipro4*

gets value 0 if none of the 4 responses takes value 1, *Recipro4* gets value 1 if at least one of the responses takes value 1, *Recipro4* gets value 2 if at least two of the responses take value 1 and so on. Table 1.8 shows the estimated marginal effects with two different set of independent variables, set 1 includes only reciprocity as explanatory variable and set 2 adds full set of control variables used in the regression. I also use all children, eldest and single children samples in the estimations of table 1.8.

The marginal effects of reciprocity expectation from a child on enrollment indicate that reciprocity expectation negatively influence the enrollment of a child, statistically significant in all and eldest children samples without any control variables, but loses its significance after adding all controls. For a child in general, it can be the case that parents have reciprocity expectation but may or may not be from the child in question, but from his/her brother. However, in single children sample if parents have reciprocity expectation from a child it definitely is from the child in question, surprisingly it remains negative in case of single children as well, though insignificant. In case of reciprocity expectation from daughter, the marginal effects are not only positive but significant in case of all and eldest children sample using full set of controls. It indicates that if parents have expectation from daughter it can influence the children enrollment decision in a positive way.

Financial expectation from child have positive influence on probability of enrollment of a child and significant in case of single child, given all other variables remain constant. Parents financial expectation from child increases the probability of enrollment of single children by 7 per cent on average, keeping other variables same. The financial expectation from daughters has a positive impact on child's enrollment, however it is statistically insignificant.

Further, the results using *Recipro4* indicate that on average reciprocity expectation can positively and significantly improve enrollment chances of children (both all and eldest children sample). But among single children, the impact of reciprocity expectation remains positive but becomes insignificant when using full set of control variables.

In addition, I also use parents with government job as an counter indicator of reciprocity

expectation. In India, the government jobs are considered to be secured than private sector jobs and all government employees are entitled to monthly pension payment after retirement. So people in government job can be assumed as financially less dependent on grown-up children at the retired age, and therefore are expected to have lower reciprocity expectation from children. I construct the government job variable as if at least one parent has government job, the variable will take value 1, otherwise takes 0. The marginal effects of parent's government job have positive and highly significant impact on children of all and eldest children samples, but negative and insignificant among single children. Again, the results for all and eldest children samples are not intuitive as parents reciprocity expectation could not be precisely directed to the child in question. Rather, the positive significance due to government job indicates that higher job security for at least one parent leads to significant increase in children's schooling chances. However, the negative effect on enrollment of single children (though insignificant) indicates that parents with government job have lower reciprocity expectation and therefore it negatively impacts the enrollment decision of single children.

Including the interaction terms of reciprocity with girl child, the results in table 1.9 indicate that for single girl child, parents' reciprocity expectations influence her enrollment significantly by as large as 20 per cent higher probability. Parents with single girl child are more likely to continue her schooling as they perceive that education will make their daughter successful and independent in future and she may provide old-age care to parents even after marriage. For single boy child, parents reciprocity expectation reduces the probability of enrollment largely, by 24 per cent, this is not what I expected. It can be the case that parents with higher reciprocity expectation are more likely to stop boy's schooling as they want him to join the labor market as soon as possible and start providing monetary help to the family. For girl child in general and for eldest sisters the results remain positive, but become insignificant. Reciprocity expectation from daughter interacted with girl child has positive impact in all samples, such as 3.4 per cent significantly higher probability of enrollment among girls in general, 1.8 per cent higher probability among eldest girls and 2.7 per cent higher for single girl child.

Table 1.9: Gender difference in imapct of reciprocity on enrollment

Dependent Variable: Enrollment	All Children	Eldest Children	Single Children
Reciprocity Expectation from child -			
Reciprocity X Girl	0.0015	0.0128	0.2003**
	(0.0614)	(0.0901)	(0.0931)
Reciprocity	-0.0108	-0.0702	-0.2400***
	(0.0407)	(0.0685)	(0.0583)
N	11884	3780	856
Reciprocity Expectation from daughter -			
Reciprocity X Girl	0.0338**	0.0179	0.0271
	(0.0173)	(0.0284)	(0.0780)
Reciprocity	0.0135	0.0277*	0.0124
	(0.0089)	(0.0154)	(0.0357)
N	9151	2800	491
Financial Expectation from child -			
Reciprocity (f) X Girl	-0.0450	-0.1448**	-0.0275
	(0.0416)	(0.0639)	(0.0742)
Reciprocity (f)	0.0304	0.0354	0.0828*
	(0.0208)	(0.0276)	(0.0437)
N	12049	3841	875
Financial Expectation from Daughter -			
Reciprocity (f) X Girl	0.0171	0.0189	0.0102
	(0.0157)	(0.0269)	(0.0786)
Reciprocity (f)	0.0046	0.0061	0.0052
	(0.0083)	(0.0147)	(0.0317)
Reciprocity 4 together -			
Reciprocity4 X Girl	0.0127	0.0108	-0.0054
	(0.0090)	(0.0146)	(0.0388)
Reciprocity4	0.0054	0.0088	0.0044
	(0.0047)	(0.0081)	(0.0189)
N	8438	2550	443
At least a parent with govt. job -			
Govt job X Girl	-0.0035	-0.0016	-0.0077
	(0.0163)	(0.0266)	(0.0582)
Govt job	0.0264***	0.0372**	-0.0080
	(0.0087)	(0.0155)	(0.0297)
N	12049	3841	875

Note: Controls used in estimations are number of siblings and male siblings (excluded in single children sample), returns to educations, std dev of wages, mother's reciprocity expectation, child's age, urban/rural, HH consumption, parents' education, religion (muslim) and scheduled caste/tribes dummies.
Standard errors clustered at HH level are given in parentheses.
* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$.

Expectation of financial reciprocity from a child in general can motivate parents to continue boys' education, with significantly 8.3 per cent higher probability for single boy child. However, financial expectation from a child does not influence parents to continue girl's education even not for single child. Parents financial expectation from a child significantly reduces the enrollment probability for eldest daughter by 14 per cent, as parent expect financial help only from sons. On the other hand, the financial expectation from daughter increases enrollment probability of girls, 1.7 per cent for girls in general, 1.9 per cent for eldest girls and 1 per cent for single girls, but these effects are statistically insignificant.

The results of interaction terms of parent's government jobs with girl child indicate that it adversely impacts their enrollment probability. The same impact on single boys can be observed as well. However, parent's government job significantly and positively influences boy's enrollment probability when estimations are done for all and eldest children samples.

1.5.6 Decomposition of Gender gap in school enrollment and contribution of the identified factors:

So far the analysis shows the difference in enrollment probability among girls compared to boys and provides evidences that parents decide differently between girl and boy child, especially when they have to decide among multiple children. Further, the results show how labor market variables and parents' future expectation impact schooling decision in general, and also how the impacts differ between boys and girls. Then, it becomes important to quantify the gender gap in schooling decision and how these aforementioned factors are responsible for the gap. Therefore, I use the decomposition method to quantify the gap and the contributions of different factors into it.

As the outcome variable for the analysis that is enrollment status of children is binary in nature, the classical Oaxaza-Blinder (1973) decomposition approach would not be appropriate. Instead, for the decomposition of gender gap in enrollment, I use the multivariate decomposition method for nonlinear response models proposed by Powers, Yoshioka and Yun in 2011. This approach provides two-component decomposition that is the explained

and unexplained components of the gap. The explained difference is contributed by the differences in the predictors among boys and girls and is called the 'endowment effect'. In other words, the endowment effect measures the average change in enrollment of boys if they have the same predictors as the girls. The other component quantifies the differences in the coefficients and the intercept between boys and girls, and is called the 'coefficient effect'. In other words, the coefficient effect measures the expected changes in boys' average enrollment if boys have the same coefficient as girls. The difference in coefficient is generally attributed to discrimination and other potential effects of differences in unobserved variables and thus called unexplained component of the raw gap. Table

Table 1.10: Decomposition of Gender Gap in Enrollment: All and Eldest Children Samples

	All Children Sample			Eldest Children Sample		
Enrollment	Model I	Model II	Model III	Model I	Model II	Model III
Raw Gap	0.0258***	0.0259***	0.0266***	0.0304**	0.0304**	0.0327***
	(0.0073)	(0.0072)	(0.0061)	(0.0129)	(0.0126)	(0.0101)
Explained part	0.0109***	0.0140***	-0.0101***	0.0144***	0.0251***	-0.0181***
	(0.0021)	(0.0022)	(0.0019)	(0.0040)	(0.0043)	(0.0033)
As percentage of raw gap	42.11	54.11	-37.92	47.20	82.46	-55.38
Unexplained part	0.0150*	0.0119	0.0367***	0.0161	0.0053	0.0508***
	(0.0077)	(0.0077)	(0.0062)	(0.0138)	(0.0137)	(0.0102)
As percentage of raw gap	57.89	45.88	137.92	52.80	17.54	155.38
Contributions of different factors into Endowment						
Return to Education	0.0001	0.0001	0.0001	0.0001	0.0001	0.00003
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0001)
As percentage of raw gap	0.54	0.55	0.33	0.20	0.36	0.09
Std deviation of wages	0.0106***	0.0101***	0.0068***	0.0130***	0.0115***	0.0063***
	(0.0020)	(0.0021)	(0.0008)	(0.0038)	(0.0040)	(0.0016)
As percentage of raw gap	40.90	38.96	25.64	42.65	37.93	19.39
Reciprocity (financial)	0.0002	0.0003	0.0003	0.0013	0.0015	0.0005
	(0.0005)	(0.0005)	(0.0003)	(0.0012)	(0.0012)	(0.0005)
As percentage of raw gap	0.67	1.20	1.23	4.35	4.94	1.39
Siblings		0.0035***	0.000004		0.0090***	0.0008
		(0.0005)	(0.0003)		(0.0017)	(0.0005)
As percentage of raw gap		13.40	0.02		29.53	2.60
N	51399	51399	51399	15405	15405	15405

Note: In Model III full set of control variables are used, such as number of siblings and number of male siblings, returns to educations and standard deviations of wages across education levels, mother's reciprocity expectation, child's age, location type of the households (urban or rural), HH consumption, both parents' education and religion and caste dummies.
Standard errors clustered at HH level are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

1.10 shows the decomposition of gender gap in enrollment, the estimated coefficients, its standard errors and the percentage of contributions of the endowment effect (explained part) and coefficient effect (unexplained part) to the raw gap. Further, it also shows the

breakup of contributions by the predictors to the explained part of the gap. The decomposition method is used in all children and eldest children sample. Model I includes only the labor market variables and reciprocity expectation as the predictors, Model II adds sibling composition, and Model III includes the full set of control variables. The values of the estimated coefficients are not informative in the nonlinear estimations, but give idea about the direction and statistical significance of the estimation.

The estimated coefficients show that there is a statistically significant positive gender gap in enrollment decision. In model I, using only labor market variables and reciprocity expectation as predictors, the decomposition reveals that 42 per cent and 47 per cent of the gaps in all children and eldest children samples respectively can be explained significantly by these predictors and 58 per cent and 53 per cent remain unexplained due to discrimination effect and unobserved heterogeneity among children in the samples of all and eldest children respectively.

After adding the sibling composition in Model II, there is a considerable increase in the percentage of explained part of the gap by predictors. For all children and eldest children, 54 per cent and 82 per cent gaps are explained respectively by the predictors, such as labor market variables, parents reciprocity expectations and sibling composition of children. In the explained part, the contribution of returns to education is around 0.5 per cent in all children and around 0.4 per cent in eldest children, and reciprocity expectation contributes to 1.2 per cent of the gap in all children sample and around 5 per cent of the explained part in eldest children. The variability of wages contributes as large as around 30 per cent of the endowment effect of the gap in all children, and almost same among the eldest children. Sibling composition also plays a crucial role in explained gender gap, around 13 per cent for children in general, and around 30 per cent in eldest children sample.

In Model III after inclusion of full set of variables, the explained part becomes -38 per cent in all children sample and -55 per cent for eldest children sample. The negative percentage is caused due to negative contribution by some predictors to the gap, especially by age of the children (around -60 per cent in all children and -70 per cent for eldest children

samples), muslim religion (-4 to -5 per cent) etc. These negative contributions by the variables imply that, for example as age increases the enrollment probability decreases more for boys compared to girls. This can also be examined by using interaction terms of these variables with girl child in the probit estimations. However, the variability of labor market still plays a significant role in the explained part of the gap, around 25 per cent among all children and 19 per cent contributed to the explained part of the gender gap among eldest children.

1.5.7 Other Crucial Factors and gender gap in enrollment

Table 1.11 shows the marginal effects of full set of variables and its interaction terms with girl child in different samples, such as all, eldest and single children. The first, third and fifth columns show the results without the interaction terms and the second, fourth and sixth columns provide the effects of both variables and their interaction terms with girl child. The inclusion of all interaction terms with full set of control variables does not provide a good fit statistics of the model, when checked with information criteria (both BIC and AIC)[10]. However, I still want to perform the analysis with interaction terms between variables and girl child to examine how these variables impact the enrollment of girls and boys differently.

Sibling Composition: If parents have more than one child then parents with higher number of children will be less likely to provide education equally to all children irrespective of gender compared to parents with a single child. Number of siblings therefore adversely impacts the enrollment probability of children. In all children sample, the marginal effects show a negative relation, but it is statistically insignificant, but among eldest children having one more siblings lowers the chances of enrollment by 1.4 per cent (highly significant). For boys number of siblings has very low impact on enrollment but

[10]The results from the fitstat of information criteria provide strong support without the inclusion of interaction terms of girl child with age, household consumption, father's education, variability of wages, reciprocity, muslim and caste dummies and positive support for without the inclusion of interaction terms of girl child with urban, mother's education, sibling composition and returns to education. Overall, the fitness statistics do not provide support of interaction terms in general.

for girls with one more siblings will have significantly lower chances of enrollment by 1.3 per cent compared to girls with fewer siblings. Similarly, among the first born boys the impact of siblings is insignificant and low, but eldest sisters with 1 more siblings have on average 2.4 per cent lower enrollment probability compared to eldest sisters with fewer siblings, keeping other things same. However, male siblings impact differently on enrollment than expected. Having male siblings impact boys' enrollment negatively but girls' enrollment probability positively. It is found that in general among boys with one more siblings the enrollment probability declines significantly by 1.2 per cent, though the impact on eldest sons is low and insignificant. Among girls in general, having one more male siblings increases their enrollment probability by 1.6 per cent significantly, and first born girls with more male siblings continue schooling by 3.8 per cent compared to girls with fewer male siblings (Table 1.11). Intuitively, it can be the case that more male siblings gives a perception of secured future to parents and therefore the family wealth or earning can be invested in girls' education either in expectation of better marriage prospects or career for girls.

Age of the child: The age of children in general negatively affect the enrollment probability of the children, highly significantly by 4 per cent on average in all, eldest and single children samples. For boys, increase in age by one year reduces their chances to continue schooling significantly by 4 per cent compared to younger boys in all samples, keeping other things same. However, for girls the impact of age is much lower in size (less than 1 per cent) and statistically insignificant in all types of sample of children. From decomposition analysis, we have seen that age of the children negatively contributes to the gender gap in enrollment, that is the probability of enrollment decreases among boys compared to enrollment probability of girls as age increases (Table 1.11). In table A1.8 in appendix, the impacts of age on enrollment probability across different age-groups indicate that the girls within the age-group of 6-14 years have positive chances to continue schooling but from age 15, their probability of enrollment becomes negative. The children in the age-group of 6-14 years are entitled to get compulsory free education and parents are more likely to send the girls to school. But when the girls pass the age of 14, the schooling is

not free anymore and parents have to bear the schooling costs to send a child to school. It is observed that this adversely impacts the enrollment probability of girls beyond the age of 14. This pattern for girls' enrollment chances across age-groups was different before the amendment of RTE in 2009. Using IHDS 2005 data, it is found that the girls enrollment probability were negative for the age-group of 6-14 years and only becomes positive above age 14 (table A1.8 in appendix). Though these results are not statistically significant but provides an evidence that the implementation of RTE and making school education free until 14 years of age has a positive impact on girls' enrollment within that age-group in comparison with time before the RTE Act. The Act does not change the trend of probabilities of boys' enrollment. Among boys in all age group, the enrollment probability is negative and becomes highly significant beyond age 10 using 2012 data. As the age increases the chances of dropping out from school significantly increases. The boys' probability of dropping out from schools is lower in size in 2012 in comparison with 2005 data across the age-groups. Therefore, for boys the Act only declines its probability of dropping-out but doesn't change the direction of impact across age-groups.

Urban: The urban location has negative impact on parents' decision of children's schooling. The negative impact is almost double in case of girls' enrollment compared to the same on boys' enrollment.

Household's economic status: The household consumption is considered as a proxy for household economic status. The wealthy households are more likely to continue children's schooling compared to poor households. The positive impact of household wealth is significant and higher in case of boys both in all and eldest children samples compared to lower and insignificant impact on girls' enrollment probability (negative in case of first born girls). However, among parents with only one child, increase in household consumption by 1 unit increases the probability of girls' enrollment significantly by 8 per cent compared to poor household.

Parents' Education: Parents education positively influences children's education and schooling achievement (Davis-Kean (2005); Chevalier (2004); Magnuson and McGroder (2002)). Educated parents perceive higher value of education and may decide to provide

Table 1.11: Marginal effects on enrollment of boys and girls

Explanatory Variables:	All Children Sample		Eldest Children Sample		Single Children Sample	
Girl Child:	-0.0364***	0.0601	-0.0521***	0.3135**	-0.0057	-0.2971
	(0.0063)	(0.0708)	(0.0108)	(0.1360)	(0.0249)	(0.2263)
Siblings	-0.00381	-0.00007	-0.01401***	-0.0083		
	(0.0027)	(0.0031)	(0.0047)	(0.0056)		
Siblings X G		-0.0131**		-0.0238**		
		(0.0060)		(0.0103)		
Male Siblings	-0.00815**	-0.0120***	-0.00139	-0.0081		
	(0.0034)	(0.0039)	(0.0062)	(0.0071)		
Male Siblings X G		0.0157*		0.0380***		
		(0.0082)		(0.0148)		
Return from Education	0.01470*	0.0111	0.00430	0.0045	0.00482	0.0074
	(0.0076)	(0.0077)	(0.0147)	(0.0144)	(0.0245)	(0.0247)
Return X G		0.1157**		0.0576		0.0490
		(0.0464)		(0.0759)		(0.3258)
Std. Dev. of Wage	0.08503***	0.0855***	0.07805**	0.0781**	0.04821	0.0449
	(0.0155)	(0.0160)	(0.0337)	(0.0348)	(0.0429)	(0.0423)
Std Dev of Wage X G		-0.0523		-0.0622		0.0324
		(0.0546)		(0.0743)		(0.4593)
Reciprocity	0.01753	0.0272	0.00097	0.0292	0.07034**	0.0780*
	(0.0184)	(0.0204)	(0.0260)	(0.0270)	(0.0358)	(0.0433)
Reciprocity X G		-0.0358		-0.1582**		0.0244
		(0.0438)		(0.0768)		(0.0932)
Age	-0.04363***	-0.0427***	-0.04537***	-0.0431***	-0.04193***	-0.0412***
	(0.0011)	(0.0013)	(0.0020)	(0.0025)	(0.0049)	(0.0060)
Age X G		-0.0036		-0.0098		-0.0052
		(0.0031)		(0.0066)		(0.0141)
Urban	-0.02838***	-0.0209***	-0.02794**	-0.0196	0.01579	0.0216
	(0.0072)	(0.0074)	(0.0116)	(0.0124)	(0.0256)	(0.0264)
Urban X G		-0.0429**		-0.0496*		-0.0500
		(0.0193)		(0.0303)		(0.0715)
HH consumption	0.01892***	0.0184***	0.01349*	0.0145*	0.02775	0.0248
	(0.0044)	(0.0045)	(0.0076)	(0.0081)	(0.0198)	(0.0189)
Consumption X G		0.0045		-0.0058		0.0798*
		(0.0121)		(0.0166)		(0.0480)
Mother's Education	0.01096***	0.0100***	0.01257***	0.0113***	0.00997***	0.0086***
	(0.0009)	(0.0010)	(0.0015)	(0.0016)	(0.0026)	(0.0028)
Mother's Edu X G		0.0043*		0.0047		0.0323**
		(0.0023)		(0.0037)		(0.0138)
Father's Education	0.00687***	0.0063***	0.00774***	0.0067***	0.00792***	0.0063***
	(0.0007)	(0.0007)	(0.0011)	(0.0012)	(0.0018)	(0.0019)
Father's Edu X G		0.0029*		0.0059**		0.0122
		(0.0018)		(0.0028)		(0.0097)
Muslim	-0.06324***	-0.0676***	-0.07406***	-0.0752***	-0.07084*	-0.0953***
	(0.0078)	(0.0083)	(0.0128)	(0.0137)	(0.0376)	(0.0370)
Muslim X G		0.0267		0.0363		-
		(0.0206)		(0.0340)		-
Scheduled caste and tribe	-0.00619	-0.0048	-0.01228	-0.0021	-0.01088	-0.0279
	(0.0071)	(0.0075)	(0.0119)	(0.0128)	(0.0231)	(0.0235)
Sch caste/tribe X G		-0.0132		-0.0776**		0.2632***
		(0.0189)		(0.0338)		(0.0956)
N	12049	12049	3841	3841	875	875

Standard errors clustered at HH level are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

better education to children. Therefore, mother's and father's education increases the likelihood of children school enrollment. I examine the impacts of mother's and father's education on children's schooling and how these impacts are different among boys and girls (Table 1.11). In general, mother's education has larger effect on children's enrollment probability than father's education. Among children in general, one year increase in mother's education increases the enrollment probability of children by 1 per cent whereas the same improvement in father's education has an impact of 0.6 per cent on enrollment probability. This impact is larger in case of eldest children and little lower in case of single child. In both among all children and eldest children, the impact of one year improvement in mother's education is much larger for boys (around 1 per cent) compared to girls (around 0.4 per cent) enrollment probability. However, for single children, one year improvement in mother's education years increases the probability of girls' enrollment by 3.2 per cent, which is almost 4 times larger than impact on boys enrollment chances (0.86 per cent). Increase in father's education by one year improves the chances of boys' schooling by 0.6 per cent in general and by 0.7 per cent among eldest boys significantly compared to boys whose fathers are lower educated. The same increase in father's education causes 0.3 per cent and 0.6 per cent improvement in chances of enrollment among girls in general and among first born girls respectively compared to girls whose fathers are lower educated. In sample of single children, the impact of father's education on boy's enrollment is highly significant but statistically insignificant for girl's schooling. Studies have shown that mother's education has higher positive impact on children's schooling, especially on girls compared to father's education years (Glick and Sahn (2000); Hoddinott and Haddad (1995); Thomas (1994); Lincove (2009)). In this analysis, I find that impact of mother's education is larger in size compared to father's education on enrollment of children, though the size of the impact is larger on schooling probability of boys compared to girls in all children and eldest children samples. But among single children, mother's education plays a significant and larger role in girl's enrollment decision compared to enrollment of single boys.

Mother's bargaining status in the household : Literature has shown that mothers

play an important role in decision of children's education especially of girls' education (Hoddinott and Haddad, 1995). Motivated by this research, I examine how mother's say and/or her bargaining power in the household can influence enrollment decision of a child and how it is different between a girl and a boy child. Using the information on say in household decision-making from IHDS 2012, the variable 'mother's say' is constructed based on whether mothers have say in expensive purchases, land and/or property purchases and in decision of number of children. A categorical variable of mother's say is constructed such as if she has no say then gets value 0; if has say in any one matter gets value 1; if in two matters then 2 and if she has say in all three matters then gets value 3. I assume that higher value in this indicator means higher bargaining status of women in household decisions. In table A1.9 in appendix, the first column provides results from estimation including mother's say as a control variable, and second column presents estimation results using mother's say and its interaction term with girl child as predictors. The next four columns use samples of children whose mother has no say, say = 1, say = 2 and full say in household decision (i.e, say = 3), instead of including say as a variable in the regression. Including say as a predictor of enrollment in probit estimation, the result shows that mother say do not significantly impact a child's enrollment decision (column 1), but including interaction term indicates that *say* plays a positive role in girls' enrollment decision but not in boys' enrollment (column 2). However, these results remain insignificant. Then using the samples of children according to mother's say index, the gender impact on girls' enrollment compared to boys' improves as mother's have higher say within household. For example, in sample where the mothers have no say in household matters, their girl child has significantly 3.8 per cent lower probability to continue schooling. When mothers have say in some matters, girl child has 2.8 per cent lower probability of enrollment than a boy. And when mothers have higher and full say in household matters, the negative impact on enrollment of their girl child decreases and becomes statistically insignificant. Therefore, the results imply that higher bargaining status of mothers within the household reduces girls' school dropping-out chances or parents will be indifferent between boy and girl child's schooling.

Religion and Caste: Only muslim religion dummy is used in the regression, and dummy for hindu religion is excluded as muslim and hindu are highly correlated. Table 1.11 shows that in general muslim children are significantly 6 to 7 per cent less likely to continue schooling compared to children in other religion, keeping other things same. Average muslim boys are 6.7 per cent and the eldest muslim boys are 7.5 per cent less likely (significant in both cases) to continue schooling compared to boys in any other religion. However, muslim girls are around 2.7 and 3.6 per cent more likely (though statistically insignificant) to continue schooling compared to girls in any other religion in all and eldest children samples respectively. The caste dummy is used for the children from scheduled caste and tribes (SC/ST). The children who belong to SC/ST families are less likely to continue schooling compared to children in other castes and tribes. The impact of caste on enrollment of boys remains statistically insignificant. For first born girls, being in a SC/ST household lower their chances of schooling significantly by 7.7 per cent than first born girls in any other castes/tribes. But using single children sample, it is found that single girls from SC/ST communities are significantly 26 per cent more likely to continue schooling compared to single girls in any other caste and/or tribes.

Schooling choices: Public versus Private schooling and No schooling

Table 1.12: Ordered Probit Regression Results - Marginal effects of gender on schooling choice between No schooling, Public / Private schooling

Dependent Variable: Schooling choice	No Schooling	Public Schooling	Private Schooling
Explanatory Variable: Girl Child			
All Children	0.0545***	0.0247***	-0.0792***
	(0.0049)	(0.0027)	(0.0072)
Eldest Children sample	0.0671***	0.0286***	-0.0957***
	(0.0084)	(0.0044)	(0.0121)
Single Children sample	0.0144	0.0075	-0.0220
	(0.0186)	(0.0098)	(0.0284)

Note: Control variables used in estimations are siblings composition, average returns to educations, std dev of wages, parent's reciprocity expectation, child's age, urban/rural, HH consumption, parents' education and religion and caste dummies.
Standard errors clustered at HH level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The schooling choices can be categorised by public schooling, private schooling and no schooling, based on these categories *schooling choice* variable is constructed where no schooling (not enrolled in school) gets value 0; going to public school gets 1 and private school gets 2. Making *schooling choice* the dependent variable, I use the ordered probit

regression method to examine how gender of a child can impact the schooling choice of parents. The marginal effects of gender (that is being a girl) on three types of choices are shown in table 1.12. It is found that girls are significantly more likely to have no schooling, 5.4 per cent in all children sample and 6.7 per cent in eldest children sample compared to similar boys, *ceteris paribus*. On the one hand, the girls have significant and positive chances (around 2.5 per cent) to get public schooling compared to similar boys and on the other hand, girls have significant and large negative probabilities to get private schooling compared to boys, given other things same. The first born girls are 9.6 per cent less likely to go to private school compared to similar boys, and girls in general (all children sample) have 8 per cent lower chance to go to private school compared to similar boys. The result for single girls also show negative probability for private schooling but this result is statistically insignificant. Under RTE Act, studying in public school is free for the children in the age-group of 6-14 years. Beyond age 14, the public schooling is either free or requires very nominal fees; many public schools provide books and uniform at a very low subsidised rate if not free, most of the time tuitions are waived etc. Parents with preference for sons are more likely to send their girls to such free or low-costs school if they decide to continue daughter's education. Alternatively, private schools charge high cost, even at the elementary level. It is found that boys are more likely to go to costly private schools whereas girls go to public schools with higher probability, if they didn't drop-out yet. I also analyse the impacts of average schooling cost and school distance (considering the distributions of schooling cost and distance at the PSU level) on enrollment decision of children (results are not given in the paper but can be available on request), but the results were not significant and informative. However, the analysis of schooling choices provides the intuition towards the impact of schooling (direct) cost on parents' decision towards children education based on gender. This analysis is in line with another paper by Azam and Kingdon (2013) which shows that parents spend lower on girls compared to boys by sending girls to free schools and sending boys to expensive private schools.

1.5.8 Robustness Checks

The robustness of the main results is checked using different methods of estimations. Table 1.13 shows results using logit regression (Panel A) and linear probability model (Panel B). In addition, the results are re-estimated using the sample weight that is used in survey sample design of IHDS-2012 and with standard errors clustered at household level - the results are presented in Panel C[11]. Due to prevalence of joint-family structure in Indian households, the observations of children within the same household would be correlated and for variables, such as religion, location etc the observations would have the same values. Considering this, the standard errors of all the estimations in this paper are clustered at household level. Further, I can treat the cross-section data as panel data to control the possibly correlated household invariant heterogeneity without observing it. I also include different location (state or districts) specific dummies to take into account further location-invariant heterogeneity in the data. The household fixed effect is not included due to its large number and for many household there are very few observations, only one or two in many cases. Panel probit regression results are shown in table 1.13 in Panel D and the marginal effects of full set of control variables using panel regression and with state dummies are shown in table A1.10 in appendix.

The results in table 1.13 in Panels A, B, C and D using different techniques of regression are found robust with the main results of table 1.3 that is on average 4 per cent lower enrollment probability for girls compared to boys in all children sample and around 5 percent lower enrollment chances for eldest sisters compared to similar boys.

I also examine the results using first round data of IHDS 2005 (Panel E in table 1.13) and found that the results remain similar as IHDS 2012 results in tables 1.3 and 1.4. The estimations of enrollment probability across completion of school levels, such as primary, elementary and secondary levels and also across age-groups using both IHDS-2005 and IHDS-2012 are given in table A1.11 in appendix.

[11]The full results with marginal effects of full set of control variables are given in appendix in tables A1.2 and A1.3

Table 1.13: Enrollment of children: Using other estimation methods and IHDS-2005

Dependent Variable: Enrollment		All Children	Eldest Children	Single Children
Panel A: Logit Regression:				
Girl Child:	Odds ratio	0.5973***	0.4971***	0.9208
		(0.0529)	(0.0743)	(0.3748)
	dy/dx	-0.0371***	-0.0511***	-0.0054
		(0.0063)	(0.0107)	(0.0265)
Panel B: Linear Probability Model:				
Girl Child:	Coefficient	-0.0427***	-0.0586***	-0.0148
		(0.0072)	(0.0123)	(0.0263)
Panel C: Probit Regression - Using survey weights				
Girl child:	dy/dx	-0.0435***	-0.0517***	-0.0138
		(0.0081)	(0.0142)	(0.0342)
Panel D: Panel Probit Regressions				
Without location dummies:				
Girl child:	dy/dx	-0.0318***	-0.0538***	-0.0058
		(0.0060)	(0.0110)	(0.0246)
With State fixed effects:				
Girl child:	dy/dx	-0.0290***	-0.0516***	-0.0101
		(0.0060)	(0.0113)	(0.0261)
With District fixed effects:				
Girl child:	dy/dx	-0.0337***	-0.0526***	-0.0140
		(0.0075)	(0.0137)	(0.0358)
N		12049	3841	875
Panel E: Using IHDS-2005				
Girl child:	dy/dx	-0.0294***	-0.0410***	-0.0247
		(0.0054)	(0.0084)	(0.0208)
N		18238	7108	1077

Note: Controls used in the estimations of samples of children are siblings composition (except single children sample, average returns to educations, std deviations of wages, mother's reciprocity expectation, child's age, urban/rural, HH consumption, parent's education and religion and caste dummies.
Standard errors are clustered at household level and presented in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Using IHDS-2005 data, the estimations reveal that the enrollment probability of girls remains negative and highly significant across completion of schooling levels. In comparison to boys, the chances of dropping out for girls increase from 2.6 per cent after standard V to around 4 per cent after VIII, and then 3.6 per cent after completion of standard X. Using IHDS-2012 data, in comparison to boys, the probability for girls to discontinue schooling was statistically insignificant after completion of primary level, but beyond elementary level their chances to drop-out increase significantly to 3.7 per cent and then further increase to 6.7 per cent after completion of secondary schooling.

From the similar analysis across age-groups of children, it is found that the enrollment probability of girls in the age-group of 6-10 years is negative but statistically insignificant in comparison with similar boys in both 2005 and 2012. The age-group of 11-14 years follows the similar pattern in both 2005 and 2012, though the chances of dropping out

of girls increases in size from 6-10 years to 11-14 years and becomes highly significant both in cases of 2005 and 2012 data. For the age-group of 15-18 years, that is the age-group in which children are not anymore entitled to free and compulsory schooling under RTE, the chances of dropping-out from school becomes 4.3 per cent for girls using IHDS-2005. The same probability calculated from IHDS-2012 is much bigger (-7.8 per cent) in 2012 in comparison with IHDS-2005. The large increase in dropping out probability among girls compared to similar boys beyond the free schooling indicates the effect due to implementation of RTE Act in 2009.

1.6 Discussion and Policy Implications

Though access to free and compulsory education has been implemented upto elementary level and for children between 6 to 14 years of age in India, gender gap in enrollment beyond primary level persists and widens with the increase in level of education. In spite of multiple interventions by Indian government to improve education situation, the policies so far have not gone beyond making the schools accessible to all but failed to adequately improve the usage of education system beyond elementary level and even beyond primary level, especially for girls. The lower presence of girl in secondary schools could not be sufficiently explained by lack of girl-friendly infrastructure in schools and/or by lower intelligence and incapacity among girls to cope with secondary level of education.

This paper quantifies this difference in usage of schooling system by boys and girls and also identifies the factors and their contribution to the difference in parents' decision in child's education based on the gender of the child. It is found that a girl child is on average around 4 per cent less likely to continue schooling compared to a boy child. Within household, the difference in providing schooling becomes an important question when parents have more than one children and thus have to decide continuing whose education, for how long, to which school etc. The paper finds that the girls in the higher birth order face more discrimination. An eldest sister is significantly around 5 per cent less likely to continue school education compared to an eldest brother. This indicates

that in access to schooling, it is worse to be eldest daughter than a daughter in general. The results from sibling composition also confirm this, having more siblings in general increases the probability of discontinuing schooling for girls compared to boys. Further, examining the difference in decision of parents with single child based on the gender of child reveals that parents' decisions on schooling remain statistically indifferent between parents of a girl and a boy child. Based on this, I infer that parents discriminate based on gender of the child, especially when they have to choose between multiple numbers of children, they provide lower education to girls compared to boys whether it is due to inherent preference for a particular gender or it is due to income constraints in the household or some other reasons. Smaller family size improves access to education for daughters and therefore policies towards restricting family size can be helpful. However, population control policies may have adverse demographic impact in terms of pro-male biased sex ratios and can exacerbate the gender biasedness in many perspectives. Therefore, population control policies should be adopted through stricter adoption of laws restricting sex selective family planning practices. Awareness campaigns to increase parents' awareness about importance of girl child and her education can help to reduce gender differential behavior among parents within households. In such campaigns, success stories of girls through education should be highlighted. Similar actions were undertaken by current government, such as *Beti Bachao Beti Padhao* (Save girl child, educate girl child) scheme was launched in 2015 with the objective of providing importance to the girl child, ensuring survival and protection of the girl-child, promoting equality between boys and girls in education and also opposing dowry and early-age marriage of girls.

The constitutional commitment[12] and Right to Education (RTE) Act (2009-10) incorporate free and compulsory education to all children up to the age of 14 years, that is up to elementary education (MHRD). However, secondary education is not free yet. If parents decide to continue education of children beyond elementary level they have to bear the expenses. This paper examines the pattern of difference in schooling decision

[12]Article 21(a) of Indian Constitution makes it clear that both boys and girls have an equal opportunity to attend school from the age of six through fourteen, and that education is a fundamental right to all.

among girls and boys across the schooling levels. The data reveals that in the transition from one level to another more and more children are withdrawn from the formal education system. The estimations across education level imply that there is a significant and positive probability of dropping out among girls after elementary level compared to boys and it increases further among girls after completion of secondary level. Post-RTE implementation, it is found that girls within the age of 6-14 years are more likely to continue schooling, and above the age of 14 the girls are significantly less likely to continue schooling. A girl above age 14 years is 8 per cent less likely to continue schooling and if she is the eldest sister among siblings her probability to discontinue becomes as large as 11 per cent (significantly). Pre-RTE implementation, the data analysis of IHDS-2005 provides evidence that this pattern was different for girls, they were more likely to dropout as age increases and after age 14 it becomes reverse, they are more likely to continue schooling if they already went to school till the age of 14. Though not always statistically significant, these results imply that the amendment of RTE has worked positively for girls in the compulsory schooling age while it was free but for boys the act only reduces the probability of dropping-out with increase in age.

Beyond the elementary level, schooling of children requires expenditures on direct (tuition fee, uniform, books etc) and indirect (opportunity) costs. The direct costs of education do not differ between boys and girls, however when parents have low incentive to spend in education of daughters than sons, even minimal cost of education may seem burden. If the present value of returns from educating girls is lower than the present value of returns from educating boys, parents are more likely to invest lower in daughter's education than in son's education, keeping other things same. When parents are already reluctant to send girls to schools, if secondary education bears costs, parents will be more convinced to discontinue girl's education at the level when it is not anymore free. This also implies that as education cost increases with levels of education, the probability of girls' withdrawn from educational institution increases. Therefore, government interventions should be extended towards free and/or subsidized education up to higher secondary level (XII), it would help to increase girls' presence at the higher schooling levels. Further, it

is also found that girls are around 2.5 per cent more likely to go to public schools and around 8 to 9 per cent less likely to go to private schools. Public schools generally provide free and/or subsidised education, and therefore parents are found to send the daughters to such low cost schools with higher probability rather sending to costly private schools. Increase in number of public schools and providing quality education in such schools may help improving girls' presence in schooling system and their performance.

Further, poor families may discontinue child(ren)'s education, when the choice is between son and daughter parents may stop daughter's education with higher probability if money is a constraint. In the analysis, the results do not provide significant evidence that higher household wealth increases the probability of girls education, rather it provides evidence for increasing enrollment chances among boys in wealthy household compared to poor households. But when parents have inherent preference towards boys over girls, then the discrimination can aggravate further if schooling cost is a concern. In 2013, the revised *Rashtriya Madhyamik Shiksha Abhiyan* included *National Incentive to Girls* by transferring a sum of Rs.3,000/- to eligible girls as fixed deposit for encouraging girls in secondary education. The girls are entitled to withdraw the sum along with interest upon reaching 18 years of age and on passing secondary examination. Recently, Uttar Pradesh (UP) state government announced a reward of Rs 10,000 for girls who passed class X exams in UP. Similar initiatives can be extended in other states as well to encourage girls in secondary education. However, whether this fund will be used for girls' higher education beyond school or as dowry payment is a matter of concern and therefore demands for more critical analysis.

Apart from direct cost the indirect or opportunity cost of schooling can also cause girls' withdrawal from schooling. Often adolescent girls are used as household labor; such as helping mothers at household chores, taking care of younger siblings etc. Due to these responsibilities girls get fewer hours to study, and eventually may dropout from schools. The results using children having very young sibling support this hypothesis. It shows that a girl child with very young sibling(s) is 4.3 per cent significantly less likely to go to school compared to girls who do not have very young siblings. Under *Sarva Shiksha*

Abhiyan (Universal education) scheme, children care centers are built near schools to free girls from sibling care responsibilities. Further, flexibility in schooling hours, evening and night schools for working girls would help to continue schooling of girls who are not able to attend school in usual hours.

Household investment in school education primarily depends on parents' incentive and willingness to spend. Decisions regarding whether all the children in the family will be sent to school, which school and for how many years are made by either one or both the parents. Therefore to understand this difference in demand for a girl and a boy child, it becomes crucial to identify the factors and their contributions that influence the parents' preference for child's schooling. The primary motivations for educating a child come from labor market returns and opportunities and also parent's expectation of reciprocal behavior from a child in future in terms of old-age care. The decomposition method to quantify the gender gap in school education reveals that the labor market factors and reciprocity expectations have significant contribution to the explained part of the gender gap. It also provides significant evidence that parents have inherent preference for son's education compared to daughter's education.

The paper finds that variability in the labor market opportunities (captured in terms of standard deviation of wages) plays an important and significant role. If variability of labor market increases, parents are encouraged to provide more education to child to secure a better future for them as they perceive more education may reduce the uncertainties. The results support this hypothesis and found that in general variability of wages increases the enrollment probability of children by around 8 per cent. However, this is not always true for girls, the variability in wages may discourage parents to continue their education for longer. The data indicates that variability of wages is lower for females, because intuitively females are hesitant to try their luck in more competitive and highly paid jobs, they only look for jobs where the probability of getting it is higher based on their educational qualification and other skills. Parents also perceive the same for girls and therefore variability and uncertainty have adverse impacts on girl's schooling decisions. However, the returns to education contribute positively and significantly to

girl's enrollment, the impact is much higher in size compared to the impact of the same on boys' schooling. The prevalence of gender wage gap in all education level discourages parents to continue daughter's education upto higher level compared to son's education. In India, there is legal provisions towards equal wage to both male and female workers with similar qualification in the same occupation. The Equal Remuneration Act (ERA) was passed in 1976, with the purpose to ensure that employers do not discriminate on the basis of gender, in wages, transfers, training and promotion. Additionally, the Sixth Pay Commission has undertaken norms to limit discrimination of employees of different gender. However, gender wage gap still persists at all education levels. An educated female worker earns lower than a similar male for doing the same job. The responsibilities of enforcing labor related laws lie with labor inspectors and staffing of inspectors is widely inadequate. Inspectors deem legislation such as the Minimum Wages Act, Factories Act, Employees Insurance Act with higher priority, hence enforcement of the ERA is minimal. Thus, inadequate staffs and lower priority to ERA lead to weak enforcement. The ambiguous language of the ERA, such as the terms like 'same or similar work' and 'equal pay for work of equal value' give opportunities to the employers to take advantage of, and pay different wages to different gender for doing exactly the same work. Removal of such weaknesses and strict enforcement of ERA would be essential. If gender wage gap persists, parents will remain reluctant to invest in girl's education and the country will not realize the true potential of its demographic dividend. Labor market variabilities do not encourage girl's education in contrast to boy's education. Higher variability and uncertainties in job market encourage parents of sons but discourage parents of daughter. Therefore, favorable labor market opportunities, gender-neutral job evaluation schemes, and in cases where a specific gender has comparative advantage over others, a special tool to evaluate performance of the workers can reduce the variability in female wages and opportunities further and thus would encourage girls' schooling.

It is essential to change mindset of society by creating awareness of importance of women's economic contribution to society. The role of social media is crucial to increase acceptance towards women's economic role. The laws that secure women's equal employment

opportunity would be useful. Labor market should help to dilute the gender roles within families, providing equal advantages to both the parents at the onset of child-bearing, such as parental paid-leave, flexible working hours/work from home facility for childcare to both the parents without the loss of seniority or pension. If job opportunities become equal for both male and female across education level, gender difference in parent's decision to provide education to children will tend to equalize.

Parents' expectation of getting old-age support from child in the future motivates them to provide education to child and to make them successful individual in future. Reciprocity expectation plays a positive role in child's schooling decision and if parents have only one child and expect to get old-age support from that child, then this expectation strongly and significantly motivates them to continue that child's education, 7 per cent higher probability to continue schooling of the single child. Parents with government job and higher job security are positively motivated to provide better/longer education to children. But having government jobs can also be counteractive and can have adverse effects on schooling of children. Government jobs also secure pension payment after retirement and thus reduce the reciprocity expectation of parents. This is visible in the result for single child, parents with government job may discontinue (single) child's schooling with higher probability though the result remains insignificant.

In Indian patrilocal societies it is more likely that parents expect reciprocal behavior from sons rather than daughters. Parents perceive higher benefit to make sons more able for the labor market than daughters. Parents educate daughters to get better groom when they grow up rather than make them a better labor market candidate. This leads to parent's lower incentive to spend in daughter's education compared to sons.

In recent times, the probabilities of getting old-age security from sons have lowered due to current economic and social changes; the likelihood that sons will live at the same place as parents has reduced largely. At the same time, women are becoming more economically independent and aware about their rights and roles. The equal inheritance rights over parent's property have strengthened daughter's economic status further. Thus, the probability that daughters will take care of old parents increases. It can be expected that

the pension schemes may reduce such reciprocity expectations, and also discrimination between boy and girl child in providing schooling can also decline. The government interventions towards pensions schemes, transfers directed to old age security addressing health, disability issues will be useful to reduce parent's dependence on offspring at old ages. These would reduce parent's biasedness towards allocating resources differently among son and daughter.

Appendix 1

Table A1.1: Enrollment of children at school-going age

Explanatory Variables:		All Children Sample Model I	All Children Sample Model II	All Children Sample Model III	Eldest Children Sample Model I	Eldest Children Sample Model II	Eldest Children Sample Model III
Girl Child:	Coefficient	-0.126***	-0.129***	-0.275***	-0.00371	-0.0621	-0.389***
		(0.0159)	(0.0320)	(0.0483)	(0.0297)	(0.0570)	(0.0814)
	dy/dx	-0.02155***	-0.02665***	-0.03623***	-0.00060	-0.01255	-0.05168***
		(0.0027)	(0.0066)	(0.0063)	(0.0048)	(0.0115)	(0.0108)
No. of Siblings		0.0183**	0.0172	-0.0287	-0.247***	-0.214***	-0.106***
		(0.00875)	(0.0139)	(0.0205)	(0.0152)	(0.0252)	(0.0353)
No. of Male Siblings		-0.144***	-0.172***	-0.0601**	-0.0734***	-0.0992***	-0.00904
		(0.0116)	(0.0183)	(0.0260)	(0.0212)	(0.0342)	(0.0469)
Return from Education			0.0531	0.113*		0.0531	0.0290
			(0.0388)	(0.0580)		(0.0696)	(0.111)
Std. Dev. of Wage			0.287***	0.657***		0.339***	0.614**
			(0.0632)	(0.121)		(0.123)	(0.262)
Reciprocity				0.132			0.00848
				(0.139)			(0.196)
Age				-0.330***			-0.341***
				(0.0107)			(0.0200)
Urban				-0.220***			-0.208**
				(0.0551)			(0.0885)
HH consumption				0.144***			0.100*
				(0.0333)			(0.0569)
Mother's Education				0.0835***			0.0942***
				(0.00703)			(0.0110)
Father's Education				0.0520***			0.0582***
				(0.00528)			(0.00865)
Muslim				-0.480***			-0.552***
				(0.0598)			(0.0984)
Scheduled caste and tribe				-0.0462			-0.0962
				(0.0540)			(0.0896)
Constant		1.482***	1.263***	5.467***	1.869***	1.583***	5.967***
		(0.0168)	(0.0293)	(0.226)	(0.0327)	(0.0577)	(0.370)
N		49318	15829	12059	14945	4727	3844

Standard errors clustered at HH level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A1.2: Probit Regression: Marginal effects on enrollment of children - Taking into account Survey Weights

Dependent Variable: Enrollment	All Children Sample			Eldest Children Sample		
Explanatory Variables:	Model I	Model II	Model III	Model I	Model II	Model III
Girl Child	-0.0283***	-0.0174*	-0.0435***	-0.0244	-0.0096	-0.0517***
	(0.0103)	(0.0098)	(0.0081)	(0.0189)	(0.0184)	(0.0142)
No. of Siblings	-0.0004	0.0003	-0.0037	-0.0482	-0.0448***	-0.0131*
	(0.0041)	(0.0039)	(0.0036)	(0.0085)	(0.0080)	(0.0069)
No. of Male Siblings	-0.0304***	-0.0289***	-0.0039	-0.0161	-0.0149	0.0046
	(0.0054)	(0.0051)	(0.0044)	(0.0115)	(0.0109)	(0.0079)
Return from Education		0.0070	0.0024		-0.0005	-0.0116
		(0.0106)	(0.0097)		(0.0210)	(0.0222)
Std. Dev. of Wage		0.069***	0.0982***		0.1048***	0.1196***
		(0.0178)	(0.0185)		(0.0403)	(0.0344)
Reciprocity			0.0163			-0.0097
			(0.0205)			(0.0321)
Age			-0.0450***			-0.0457***
			(0.0014)			(0.0022)
Urban			-0.0292***			-0.0248*
			(0.0082)			(0.0130)
HH consumption			0.0222***			0.0172**
			(0.0055)			(0.0086)
Mother's Education			0.0110***			0.0133***
			(0.0012)			(0.0021)
Father's Education			0.0077***			0.0074***
			(0.0009)			(0.0014)
Muslim			-0.0620***			-0.0803***
			(0.0096)			(0.0147)
Scheduled caste and tribe			0.0137			0.0058
			(0.0087)			(0.0147)
N	12049	12049	12049	3841	3841	3841

Note: Standard errors are clustered at HH level, and also take into account survey weights.
SEs are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A1.3: Marginal effects of gender on enrollment decision of Single children - Taking into account Survey Weights

Dependent Variable: Enrollment	Model I	Model II	Model III
Explanatory Variables:			
Girl Child	-0.0179	0.0014	-0.0138
	(0.0432)	(0.0429)	(0.0342)
Return from Education		-0.0223	-0.0181
		(0.0398)	(0.0378)
Std. Dev. of Wage		0.1419*	0.0781
		(0.0757)	(0.0582)
Reciprocity			0.0875**
			(0.0396)
Age			-0.0418***
			(0.0053)
Urban			0.013
			(0.0306)
HH consumption			0.0342
			(0.0246)
Mother's Education			0.0097***
			(0.0034)
Father's Education			0.0095***
			(0.0026)
Muslim			-0.1050*
			(0.0553)
Scheduled caste and tribe			-0.0026
			(0.0275)
N	875	875	875
Controls	Set 1	Set 2	Set 3

Standard errors clustered at HH level and using survey weights are given in parentheses. * $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$.

Table A1.4: Missing values of labor market variables are substituted with District and then State average

	All Children	Eldest Children
Substituted with District average		
Girl Child	-0.0103**	-0.0148**
	(0.0040)	(0.0067)
Return to education	0.0022	-0.0084
	(0.0066)	(0.0124)
Std Dev of wages	0.0792***	0.0682***
	(0.0097)	(0.0154)
N	28869	9386
Further Substituted with State average		
Girl Child	-0.0082**	-0.0127**
	(0.0036)	(0.0059)
Return to education	0.0014	-0.0111
	(0.0066)	(0.0112)
Std Dev of wages	0.1256***	0.1278***
	(0.0099)	(0.0157)
N	34116	11140

Full set of control variables (as Model III in table 1.3) is used in all estimations. Standard errors in parentheses are clustered at HH levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A1.5: Marginal impact of having very young siblings on enrollment

Dependent Var:	All Children		Eldest Children	
Enrollment	Model I	Model II	Model I	Model II
Girl child	-0.0363***	-0.0327***	-0.0520***	-0.0465***
	(0.0063)	(0.0066)	(0.0108)	(0.0116)
Young sibling	-0.0089	0.0053	-0.0088	0.0034
	(0.0122)	(0.0144)	(0.0163)	(0.0204)
Young sibling X Girl Child		-0.0430*		-0.0385
		(0.0221)		(0.0295)
N	12049	12049	3841	3841

Model I includes having very young sibling as control variable and Model II includes young sibling and its interaction with girl child. Other control variables used in these estimations are siblings composition, average returns to educations and standard deviations of wages across education levels, parent's reciprocity expectation, child's age, urban/rural, HH consumption, both parents' education and religion and caste dummies.
Standard errors are clustered at HH level and given in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A1.6: Descriptive statistics of the families with single child and multiple children:

Average Values	Single child HH	Multiple children HH
HH consumption (scaled down)	1.45	1.22
Mother education (years)	5.6	3.9
Father education (years)	5.9	4.8
Mother's earning (INR)	9207.7	5836.0
Father's earning (INR)	37074.9	33044.2
Reciprocity expectation	0.92	0.97

Table A1.7: Labor market success on enrollment of children

Dependent Variable: Enrollment	All Children	Eldest Children	Single Children
Girl child	-0.0412***	-0.0499***	-0.0089
	(0.0081)	(0.0143)	(0.0344)
Maximum Return to Education	-0.0192***	-0.0223	-0.0727***
	(0.0066)	(0.0137)	(0.0259)
Using interaction term between Return to Education and girl child			
Girl child	-0.0353***	-0.0515***	-0.0010
	(0.0064)	(0.0110)	(0.0251)
Maximum Return to education	-0.0086	-0.0044	-0.0464**
	(0.0064)	(0.0104)	(0.0217)
Maximum Return to education X Girl Child	0.0231	-0.0014	0.0489
	(0.0204)	(0.0302)	(0.0966)

Full set of control variables (as Model III in table 1.3) is used. Clustered standard errors at HH level are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A1.8: Impact of age on enrollment of children across different age-groups

Explanatory Variable	All	6-10	11-12	13-14	15-18
IHDS-2012					
Age	-0.0427***	-0.0005	-0.0292***	-0.0618***	-0.0889***
	(0.0013)	(0.0010)	(0.0097)	(0.0126)	(0.0061)
Age X Girl	-0.0036	0.0047*	0.0015	0.0423*	-0.0104
	(0.0031)	(0.0029)	(0.0161)	(0.0229)	(0.0155)
Explanatory Variable	All	6-10	11-12	13-14	15-18
IHDS-2005					
Age	-0.0718***	0.0103***	-0.0347***	-0.0530***	-0.1139***
	(0.0018)	(0.0040)	(0.0119)	(0.0128)	(0.0052)
Age X Girl	-0.0004	-0.0052	-0.0044	-0.0080	0.0104
	(0.0029)	(0.0062)	(0.0170)	(0.0179)	(0.0083)

Control variables used in these estimations are gender of the child, sibling composition average returns to educations and standard deviations of wages, reciprocity expectation, location type of the households (urban or rural), household consumption, both parents' education and religion and caste dummies. Standard errors are clustered at HH level and given in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A1.9: Marginal Impact of Mother's say on enrollment of children

Dependent Var: Enrollment	Say as Control Var	Say and its interaction with girl child	Say = 0	Say = 1	Say = 2	Say = 3
Girl child	-0.0365***	-0.0386***	-0.0383***	-0.0283**	-0.0115	-0.0142
	(0.0063)	(0.0070)	(0.0073)	(0.0137)	(0.0424)	(0.0689)
Mother Say	-0.0036	-0.0050				
	(0.0043)	(0.0046)				
Mother Say X Girl Child		0.0076				
		(0.0107)				
N	12049	12049	9068	2393	355	220

Control variables used in these estimations are number of siblings and number of male siblings average returns to educations and standard deviations of wages across education levels, mother's reciprocity expectation, child's age, location type of the households (urban or rural), household consumption, both parents' education and religion and caste dummies.
Standard errors are clustered at household level and given in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A1.10: Panel Probit Regression Results with state dummies:

Dep Var: Enrollment	All Children	Eldest Children	Single Children
Marginal effects (dy/dx)			
Girl Child	-0.0290***	-0.0516***	-0.0101
	(0.0060)	(0.0113)	(0.0261)
No. of Siblings	-0.0047*	-0.0169***	
	(0.0027)	(0.0051)	
No. of Male Siblings	-0.0087***	-0.0025	
	(0.0033)	(0.0065)	
Return from Education	0.0213***	0.0114	0.0006
	(0.0074)	(0.0153)	(0.0205)
Std. Dev. of Wage	0.0658***	0.0713**	0.0693*
	(0.0141)	(0.0350)	(0.0380)
Reciprocity	0.0152	-0.0052	0.0635*
	(0.0172)	(0.0281)	(0.0383)
Age	-0.0433***	-0.0496***	-0.0453***
	(0.0012)	(0.0026)	(0.0071)
Urban	-0.025***	-0.0353***	0.0160
	(0.0069)	(0.0121)	(0.0257)
HH consumption	0.0159***	0.0108	0.0202
	(0.0042)	(0.0078)	(0.0201)
Mother's Education	0.0103***	0.0131***	0.0104***
	(0.0009)	(0.0017)	(0.0028)
Father's Education	0.0067***	0.0082***	0.0086***
	(0.0007)	(0.0011)	(0.0019)
Muslim	-0.0636***	-0.0696***	-0.1054***
	(0.0079)	(0.0136)	(0.0381)
Scheduled caste and tribe	-0.0077	-0.0235*	-0.0172
	(0.0073)	(0.0131)	(0.0238)

Standard errors are clustered at household level and given in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

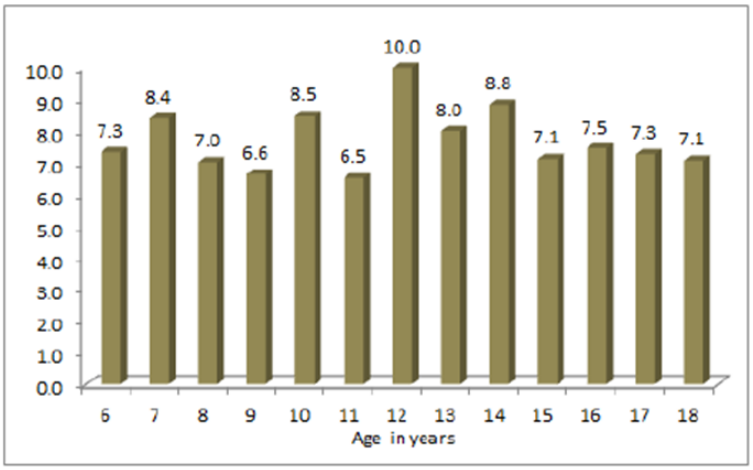
Table A1.11: Enrollment across age-groups and completion levels:

Using IHDS-2005 and IHDS-2012

Dependent Variable: Enrollment	After Completion of:		
IHDS-2005	Std V	Std VIII	Std X
Girl child	-0.0264***	-0.0430***	-0.0359**
	(0.0064)	(0.0107)	(0.0177)
N	10122	5175	2012
IHDS-2012			
Girl child	-0.0288	-0.0375***	-0.0667**
	(0.0192)	(0.0101)	(0.0194)
N	861	5675	1171
	Age-group wise:		
IHDS-2005	6-10 Years	11-14 years	15-18 years
Girl child	-0.0008	-0.0155***	-0.0430***
	(0.0071)	(0.0058)	(0.0098)
N	1243	8387	8608
IHDS-2012			
Girl child	-0.0016	-0.0256***	-0.0783***
	(0.0025)	(0.0077)	(0.0169)
N	3417	4678	3964

Full set of control variables (as Model III in table 1.3) is used in estimations.
Clustered standard errors at HH level are given in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure A1.1: Age distribution of children at school-going age



10.14754/CEU.2019.02

10.14754/CEU.2019.02

CHAPTER 2

Impact of 'Having a Son' on Women's Intra-household Status: Evidence from India [1]

2.1 Motivation

In intra-household settings, individual bargaining power is crucial for positions and control over decisions that influences resource allocation and individual well-being. Bargaining power also impacts individual labor allocation in different activities within household, his or her consumption basket, access to provisions in household, and children's education and health. Literature has measured bargaining power by individual's earning capacity, his or her monetary contribution to the family and asset holding. Bargaining power may differ between men and women for various reasons, often due to differential access to education and occupation.

Bargaining power may also differ due to unequal social norms in terms of class, caste and gender. Patriarchal societies commonly see men as superior to women even when

[1] I thank Andrea Weber and Gábor Kézdi for valuable advice.

both have same education and working status. Such societies also put higher values to having a son than having a daughter. Preference for son comes from the perception of higher utility gains from son(s) compared to daughter(s). Higher utility from son is seen in terms of future investment and insurance for parent's old ages. If it is so, then individual's fertility outputs can play an important role in bargaining power. Also, in such societies with prevalence of patriarchy and preference for son, child-bearing and childcare are considered to be women's virtue and sole responsibility and often women are blamed to not giving birth to the preferred gender composition of children. Then, fertility output may play more important role for women's bargaining power than men. In Asian societies, women who do not have a son often experience humiliation and negligence by family members and even abandoned by the family. Thus, the question arises what happens to women's position and status if they succeed to provide a son to the family. It would be interesting to examine whether having a son has any implications in women's intra-household bargaining power. Therefore, the research questions of this paper are:

- Does having a son impact women's status in family?
- Does having a son change women's bargaining power and their role in household decision making?

The rest of the paper is organised as follows: Section 2.2 reviews relevant literature on the topic; Section 2.3 gives description on data and Section 2.4 explains the identification strategy and estimation methodology; Section 2.5 presents the results as well as its robustness tests; and Section 2.6 concludes with discussion on results.

2.2 Literature Review

Intra-household bargaining power is often measured by earning (Friedberg and Webb (2006); Bittman, England, Sayer, Folbre, and Matheson (2003)) asset holding (Fafchamps and Quisumbing (2002); C. Doss (2013); Gray (1998)) land holding and savings (C. R. Doss et al. (1996); Deere and Doss (2006); Lise and Seitz (2011)). Bittman et al. (2003) paper

used Australian and United States data to show that bargaining has increased with the rise in income and women decrease their time in housework as their earnings rise, up to the point where both spouses contribute equally to income.

A paper by Pollak (2005) argued that it is not earning but individual's wage rate impacts bargaining power in the marriage. He explained that the individual whose earning is higher due to higher wage rate will gain higher bargaining power at home instead when his/her earning is higher due to longer working hours at the job and less at home. He also puts more importance on individual's household production while measuring bargaining power within household.

Women's bargaining power also depends on her education level compared to other family members, and if she comes from higher social background (Beegle, Frankenberg, & Thomas, 2001). Frempong, Stadelmann, et al. (2017) examined the impact of female education on their bargaining power within household using individual-level data on Ghana and Uganda and found that though female education has positive effects on household welfare but it is not through the channel of improvement in bargaining power due to higher education. Agarwal (1997) explains the complex nature of intra-household dynamics and links them with household bargaining power. She has pointed out that social norms play a crucial role in determining bargaining power.

After pointing out the determinants of the bargaining power, it is also important to understand the impacts of bargaining power on individuals. Bargaining power of an individual within household affects his or her access to consumption basket, control over expenses, access to other provisions in the household, labor allocation in variety of work, and child's education and health.

As consistent with the exchange-bargaining theory, Bittman et al. (2003) has shown that women's household-labor hours are negatively related with her earnings (using data on Australia and United States). Using UK family expenditure survey data, Lise and Seitz (2011) paper shows that the difference in earnings between the couples translates into differences in consumption allocations within the household. This paper also indicates that increases in marital sorting on wages and labor-hours explain the decline in intra-

household inequality.

Within household, women's bargaining power is an important parameter to influence various social and economic decision-makings. It is often claimed that men spend larger share of their earnings on alcohol, cigarettes, and on entertainment; whereas women, if have earning or access to the household earning/saving, are more likely to spend on general household welfare and children's education and health. Using information on Cote D'ivoire, Hoddinott and Haddad (1995) show that the increasing bargaining status for women would influence household expenses and would lead to higher welfare to children and household all-together.

Brown (2009) has shown that bargaining power instrumented with dowry payment is related to women's individual welfare and wellbeing in China. Using information on Ghana, Tolhurst, Amekudzi, Nyonator, Squire, and Theobald (2008) paper found that treatment seeking behavior for children was influenced by bargaining power in decision making and ownership of children, access to and control over resources to pay for treatment, norms of responsibility for payment, marital status, household living arrangements, and the quality of relationships between mothers, fathers and elders.

In respect to Indonesia, a paper has linked women's bargaining power with prenatal and delivery care positively (Beegle et al., 2001). A paper by Schmidt (2012) found a positive correlation between children's health and mother's role in decision making within household. He also found that in case of Bangladesh, mother's role in decision making, in terms of say in child's healthcare is associated with larger child height-for-age z-scores. A study on India has found positive relation between women's bargaining status with children health and household nutritional status (Imai, Annim, Gaiha, Kulkarni, et al., 2012).

Therefore, there are no shortages of literature which examined different determinants and dynamics of bargaining power, and also have explored the consequences of unequal bargaining power, the impacts of gaining higher status in decision making on individual and household welfare. In spite of the wide range of research in respect of intra-household bargaining power, only few have examined the links between bargaining status of indi-

viduals' and their fertility output. More in patriarchal societies and in moderate extent everywhere in general, composition of children is a crucial factor to influence resource allocation, division of income and expenses in the household. Therefore, the hypothesis is that families which have preference for any particular gender would provide higher status to the individuals, especially women, who could provide that preferred composition of children. Therefore, it is important to examine women's bargaining power and role in household decision-making in respect to the gender composition of children the women have.

A paper by Raley and Bianchi (2006) shows that in the United States in families with sons; fathers invest higher money and time into family; mothers face lower risk of marriage disruptions and enjoy higher marital happiness than in families with daughters/without sons. Using data on the United States, another paper by Morgan, Lye, and Condran (1988) found that parents with one or two children, and at least a son have lower divorce risks than parents with only daughters. Another similar kind of research on Sweden by Andersson and Woldemicael (2001) found that among couples with mixed child, divorce risks are slightly lower than couples with two daughters.

Using Chinese data, J. Li and Lavely (2003) investigated why women prefer son and why women give more importance to son, so the authors have used importance of son as outcome variable and found that the importance of a son is negatively associated with respondents education, personal autonomy, the extent to which husband shared housework, and exposure to the world beyond the village. The paper also finds that sex specific rates of infant mortality vary systematically with mother's response on the importance of son.

Warner (1991) considered women in the United States and Canada and men in the United States, and found that having female children is associated with more egalitarian views than women and men with male child. Washington (2008) used data on the voting behavior of the United States House of Representative between 1991 to 2004 to show that parenting of an additional girl child improves a representative's voting propensity towards egalitarian rights, conditional on total number of children. A descriptive paper

by Mason (1984) argues that women's position or "status" seems likely to be related to the supply of children because of its links with age at marriage.

Though there are many works which have shown that how gender composition of children influences marital happiness/life, father's investment in the family and parents' egalitarian views, but so far, as best of my knowledge, only one paper examines women's role in decision making in respect of their fertility output. (L. Li & Wu, 2011) investigated women's role in purchasing consumer durable items for household with respect to gender of the 1^{st} born child and found women with a 1^{st} born son have a 3.9 percentage points greater role in household item purchases than a woman with a 1^{st} born daughter (in China).

This chapter can be called an improvement from Li and Wu's paper as it explores broader definitions of both gender composition of children and household status of women. In the analysis it estimates the impact of having at least a son and also 1^{st} born son on women's say. As the dependent variables, I use larger number of indicators describing multiple aspects of bargaining power within the household (HH), such as large purchase, spending own and husband's money, visiting relatives, contraception use etc. Further, I also examine impact of having son(s) on domestic violence to women, as violence within household can be a proxy of individual's household status. As samples of analysis, this paper uses women who are still in the reproductive cycle and also women who have completed/opted out from the cycle. As gender composition of women's fertility output can be a determinant of their value in the family and can have important implications on their bargaining status within household; and as literature in this direction is scarce, the paper contributes to fill this gap in the literature.

2.3 Data

Data for this paper is taken from the National Family and Health Survey of fourth round (NFHS-4) conducted during 2015-16 in India. NFHS-4 collected information from a nationally representative sample of 568,200 households, 699,686 women of age 15-49, and

74,369 men. The NFHS-4 sample represents the characteristics of 99 per cent of India's population living in all 36 states. For the analysis I primarily use 'the couple data' from the survey that covers randomly selected 63,696 couples (NFHS-4, 2015). Under the survey the husbands and wives were interviewed separately. I use husbands' responses as a proxy for household's view.

The survey provides information on household member's say in various household decisions and also on household attributes; individual characteristics and spouse's background; reproductive behavior, childbearing and children's birth history; marriage and cohabitation; use of contraception, general health and nutrition; antenatal, delivery and postnatal care; child health and child-rearing practices; and status of women and spousal violence.

To address the under-reporting issue, the module was specially designed to allow the interviewer to be concerned about the interruptions by other household members and/or neighbors during the interview. They took precautions that during the interview interruptions and interferences remain nil or low from the family and outside the family. The survey records information on the interruptions during interviews. As the interruptions can significantly influence the responses, the empirical analysis of this paper further controls for such interruptions.

One concern may arise that the responses can be biased if the women respondents themselves are with patriarchal views, and therefore themselves do not put importance to their own voices in decision-making. However, as the research investigates how '(importance of) women's say' in household decision-making may change due to their fertility output; the respondents' biased views are less likely to influence the objective responses when they were asked about the reality in households that is whether they have say over such and such decisions. The subjective views are more likely to influence the responses if they were asked that whether they should have a say or not in the family decision making. Instead, women were asked about their actual position in decision making, such as:

- *Who decides how the money you earn will be used: mainly you, mainly your husband, or you and your husband jointly?*

- *Who decides how your husband's earnings will be used: mainly you, mainly your husband, or you and your husband jointly?*
- *Who usually makes decisions about making major household purchases: mainly you, mainly your husband, you and your husband jointly, or someone else?*
- *Who usually makes decisions about visits to your family or relatives: mainly you, mainly your husband, you and your husband jointly, or someone else?*
- *Would you say that using contraception is mainly your decision, mainly your husband's decision, or did you both decide together?*

Here, four cases may arise, such as (i) women who are in-support of the patriarchal values and have lower say in household decisions would truly reveal whether they have any say in decisions or not as they don't have any low self-esteem due to lack of say; (ii) women who are supporters of patriarchal norms but enjoy higher bargaining status in the household would more likely report the true status as they don't have any incentives to hide; (iii) women who are of more liberal view and have higher say in decisions also have no incentives to hide the actual state; but (iv) women who have liberal view but do not enjoy equal say in household decisions may or may not hide their true status, may not hide if they believe that revealing their true status can bring some change or may hide if reporting their low status is embarrassing for them. In anyways, I assume that the miss-reporting or hiding cases are rare when asked about their objective bargaining status, mostly the women would report their true status in decision making within the household. But this can be a problem when considering husband's responses as they were asked that:

In a couple, who do you think should have the greater say in each of the following decisions: the husband, the wife or both equally:

- *Deciding what to do with the money the wife earns from her job?*
- *Making major household purchases?*
- *Deciding about visits to the wife's family or relatives?*

- *Deciding how many children to have?*

Thus, husbands were asked more subjective and in general questions, then husbands responses may not reflect the scenario within the household but their views in general and that can be polluted by own and societies patriarchal values.

In the total sample of 63,696 couples around 5596 couples are without a child, these observations are dropped from the total sample for the purpose of this analysis and the total sample size becomes 58,100. Also, I use the individual sample of women for robustness checks, and among total 699,686 women of age 15-49, 223,067 observations are dropped as these women didn't give birth to any children.

2.4 Empirical Strategy

As estimation method, I use probability model to estimate the equation:

$$Pr(Y_i = 1|X_i, C_i) = \phi(\beta X_i + \gamma C_i + \epsilon_i) \tag{2.1}$$

i: individual respondents;

X: women having a son and

C: Other variables that may influence women's bargaining status in the household.

In equation (2.1), the dependent variable Y refers to different indicators for women's bargaining power in household decisions, such as women's say in:

- Large purchases in the household;
- Spending own money;
- Spending husband's money;
- Visiting relatives;
- Contraception use;
- Number of children.

Dependent Variable: Women's Say within household - Primarily, all the say indicators are constructed as binary variables. For the analysis, responses from women and men on say in different decisions have been considered.

Women were asked whether they have a say in the household matters, such as spending, large purchases, spending husband's money, visiting relatives and/or friends, and contraception use. The reply categories were: *'Only she, herself has a say', 'Only her husband has a say', 'Both she and her husband have joint/equal say',* and *'Someone else in the household has a say'*. Based on these responses, say variables are constructed as binary variables by the rule that if a woman has a say independently or jointly with husband, she will get value 1 and 0 otherwise.

Husbands were asked that in a couple who should have the greater say when deciding about the similar matters such as, spending husband's money, large purchases, spending wife's earning; visits to wife's family and/or friends; and in decision of number of children. The reply categories of husbands were: *'Only husband should have a say', 'Only wife should have a say', 'Both husband and wife should have equal say',* and *'Someone else should have a say'*. Similarly, the say indicators are constructed as binary variable if a husband thinks that in a couple wife should have a say independently or jointly with husband will get value 1 and 0 otherwise.

Explanatory Variable: Women having at least a Son - I construct the main explanatory variable 'women having son(s)' as 'women who have at least a son' (X in equation 2.1) from children ever born. As the purpose of this research is to examine whether having a son influence women's status than not having one, the comparison between having higher number of sons than only one or two becomes less relevant. Therefore, using at least a son would justify the question of interest.

Women having at least a son can be endogenous due to fertility decisions and preference for son. For example, women who have higher bargaining power may decide to keep the girl child and to stop having another child after having a daughter. But a woman, who lacks importance in household decision-making, may not have the power to decide about choices of childbearing and may be compelled into giving more births until she succeeds

to provide the preferred choices of children to the family.

Though the gender of a child is a random selection, the choices of child-bearing and actual fertility output are not random and often polluted by not only mother's choices but also depend on family members' demands. Thus gender composition of children is far from being random and needs critical attention when using as a predictor. I identify three mechanisms that can possibly influence the gender composition of children ever born.

Women's biological factors may influence the gender of the child. But literature has shown gender of a child as a natural phenomenon and random event. Literature found that the probability of having a boy is 0.513 and sex ratio at birth is about 1.05, irrespective of race, societies and nations (Teitelbaum (1972); Johansson and Nygren (1991); Waldron (1983); Waldron (1993)).

The second mechanism is sex-selective abortion to get preferred gender composition of children. Couples may decide to not keep the child after knowing the sex of the fetus, abort it depending on the preference for a particular gender. There is a ban on sex detection test in India by Pre-Conception and Pre-Natal Diagnostic Techniques (PCPNDT) Act since 1994. Revealing sex of the fetus is a punishable offense in India. Further, one can infer that sex selective abortion is rare if the sex ratio of children is statistically indistinguishable from 50 per cent or sex ratio of boy:girl is 51:49, which is agreed as the natural in many literature. Thus, according to the order of birth, such as for children born at first, second, third and so on, sex ratio at each birth should be similar as the natural one (51:49) when sex-selective abortion is rare.

The third mechanism that can influence composition of children in a family is differential stopping behavior (DSB) among couples. On the one hand, parents can stop having children after attaining the desired gender composition of children and family size. And, on the other hand, parents may continue having more children until they reach the desired number of children of preferred gender, even if have reached desired family size. Therefore, DSB depends both on desired number of children and desired gender composition of children.

Followed by the above discussion, the main explanatory variable 'women having son(s)'

and/or 'women having at least a son' is endogenous in nature and therefore will provide a biased estimator if not treated for endogeneity. One potential instrument for women having son(s) can be 'gender of the 1^{st} born'.

Let's consider 'gender of the 1^{st} born' as an instrument Z for the main explanatory variable 'having at least a son', earlier defined as X in equation 2.1. I can write,

$$\begin{aligned} \text{Z} &= 1 \text{ if the } 1^{st} \text{ born is a son} \\ &= 0 \text{ if the } 1^{st} \text{ born is a daughter.} \end{aligned}$$

Considering having at least a son is similar as a treatment (let's say treatment X), Z is a binary instrumental variable for this treatment. The potential treatment indicators are thus,

$$\begin{aligned} \text{X}(0) &= 1 \text{ if woman having at least a son but } 1^{st} \text{ born a daughter (Z} = 0) \\ \text{X}(1) &= 1 \text{ if woman having at least a son and also } 1^{st} \text{ born a son (Z} = 1) \end{aligned}$$

The actual treatment indicator is X = Z X(1)+(1-Z) X(0). Based on this definition, I can categorize the total sample as,

- X(0) = 0: The women whose 1^{st} born is a daughter (Z = 0), and don't have at least a son, even if they have continued child-bearing. Around 94.8 per cent of women who don't have at least a son fall under this category of compliance.

- X(0) = 1: The women whose 1^{st} born is a daughter (Z = 0), may still have at least a son if had continued child-bearing and had a son(s) later. In the sample who have at least a son, around 37.7 per cent are always takers.

- X(1) = 1: The women whose 1^{st} born is a son (Z = 1), and thus have at least a son. Around 62.2 per cent of women have at least a son as compliers.

- X(1) = 0: The women whose 1^{st} born is a son (Z = 1), but don't have at least a son. This can only happen if the first son has died and the women didn't have a son afterwards. The never-takers are around 5.2 per cent of women who don't have at least a son.

The last category of never takers is not applicable for this analysis, as I consider only children *ever born*.

I include both X = 1 for women in the categories of compliers (Z = 1) and always takers (Z = 0), and X = 0 only for women in compliers (Z = 0) to estimate the treatment effect on the outcome. The choice of being in the treatment group is endogenous, that is having at least a son is a choice the couple can make. Further, the gender of the 1^{st} child can be a valid instrument if and only if two conditions are satisfied. The instrument, gender of the 1^{st} born (Z) should be correlated with having at least a son (X). The condition can be proved by the non-zero correlation coefficient of Z with X, and it is found that this condition holds [2].

The other condition is that the instrument must be exogenous that is $Cov(Z, \epsilon)$ should be equal to zero. Therefore, the binary variable, 'gender of the 1^{st} born' should be a random event. And, having at least a son is the only way by which gender of the 1^{st} born can influence the outcome variables which implies exclusion condition of Z. We can justify this condition intuitively.

Table 2.1 shows the ratio between boy child and girl child according to the birth order without fixing the total number of children of the respondents. The ratio of having son to having daughter in the first born is 52.5 : 47.5, little higher than the ratio 51 : 49 shown by various literature as natural sex ratio at birth. Therefore, the question remains, 'is the gender of the 1^{st} birth random?' Fixing for parity the sex ratio of children in the data are given in table A2.1 in appendix of this chapter.

To get children of preferred gender, sex selective abortion is a tool couples may use, however intuitively, in case of 1^{st} pregnancy the chances of sex selective abortion can be considered as rare. It can be assumed that couples usually start adopting tools to get preferred gender composition of children from the second pregnancy, conditional on the gender of the first birth. At the first pregnancy, couples are typically more excited about the pregnancy than strictly adhering to their preference, unless the couples bear a very

[2]Correlation coefficient between at least a son and 1^{st} born a son: 0.3242*** (0.00294). [Standard error is in parentheses and *** implies statistical significance at 1 per cent level.]

Table 2.1: Sex ratio of Children across the order of birth

Order of Birth	Ratio (Boy:Girl)
1st born	52.5 : 47.5
2nd born	51.3 : 48.7
3rd born	52.5 : 47.5
4th born	52.5 : 47.5
5th born	51.6 : 48.4
6th born	51.6 : 48.4
7th born	50.0 : 50.0
8th born	53.1 : 46.9
9th born	50.6 : 49.4
10th born	45.3 : 54.7

strict preference for a gender that they do not want the other gender child at all. To examine this, I look into the abortion pattern in respect of gender composition of the children.

Table 2.2: Abortion and Gender composition of children

Parity	First Born	Abortion done (Per cent)
Without fixing for parity	B	16.23
	G	17.26
Parity 1	B	16.35
	G	16.11
Parity 2	B	16.25
	G	17.41
Parity 3	B	16.03
	G	17.79

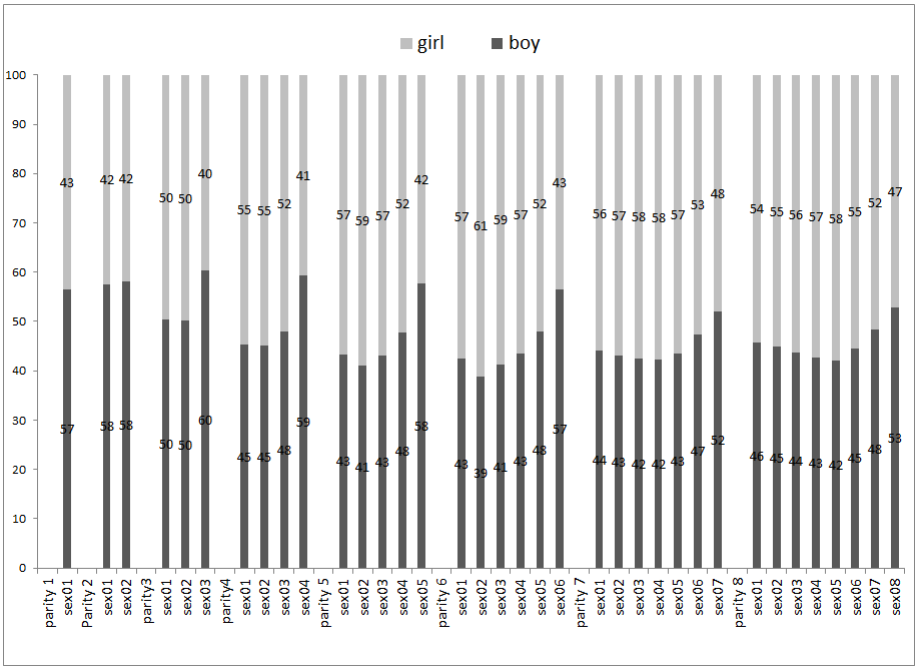
Gender Composition of children by order of birth		Abortion done (Per cent)
Without fixing for parity	last born B	16.53
	last born G	16.98
Parity 1	B	16.35
	G	16.11
Parity 2	BB	16.63
	BG	15.83
	GB	16.47
	GG	19.13
Parity 3	BBB	16.01
	BBG	16.15
	BGB	15.41
	BGG	16.87
	GBB	16.12
	GBG	18.15
	GGB	17.81
	GGG	20.93

The abortion data in the NFHS do not provide information on the time of abortion across birth order of children, instead the women were asked if 'she ever had terminated a pregnancy'. In spite of this lack in the data, I can still get an idea in general about what proportion of women ever had an abortion with respect to gender composition of their children. Table 2.2 shows the abortion rate among women according to the children composition they have. Looking at the abortion rate as per the 1^{st} born child (left side table), it is evident that the abortion rate is higher in case of 1^{st} born daughter (17.26)

than 1^{st} born son (16.23) without fixing for parity size. Taking into account of the parity size, for parity one, the percentage of women ever aborted a pregnancy is similar (16 per cent) between women with a son and women with a daughter. However, starting from parity two onwards, I find that on the one hand, the percentage of women who have ever terminated a pregnancy increases among women with 1^{st} born a daughter and on the other hand, the percentage of women who have aborted decreases for women with 1^{st} born a son.

Table 2.2 also exhibits the abortion rate of women across the gender composition of children by birth order (right side table). Generally, the proportion of women had an abortion is similar (16.53 per cent in case of last child a boy and 16.98 per cent if the last child is a girl) irrespective of the gender of the last child without fixing for parity. Fixing for parity, similar proportion of mothers who have single child have terminated a pregnancy irrespective of the gender of the child they already had (16.35 per cent of mothers with boy child and 16.11 per cent with a girl child). However, starting from parity two, it becomes evident that abortion pattern is influenced by the gender composition of children, especially percentage of abortion is higher when couples have more daughters compared to sons. In parity two, higher proportion of mothers (19 per cent) has terminated pregnancy if having two girls compared to mothers who have either at least a boy or two boys (around 16 per cent). Similarly, in parity three, higher proportion of mothers have gone through abortion who have three girls (20.93 per cent), compared to mothers who have either three or two boys. This abortion pattern in respect of gender composition of children indicates that abortion rate is higher if couples have more number of girls, possibly abortions are done in fear of getting another girl. It can also be said that women after having more daughters are less willing to have another child, than women having a mixed composition. This implies that abortion is used as more of a method of stop having more daughters than stop having daughter at all. Thus, abortion probability depends on expected family size and expected gender composition of children. Parents do not mind to have more sons even if it exceeds their planned family size, but will strictly adhere to the planned family size if daughter is expected to born, rather will abort the

Figure 2.1: Parity-wise Sex ratio in Birth-order



fetus, probably without knowing the gender of the fetus.

As we are interested in the gender of the 1^{st} born, table 2.2 for women had abortion without fixing the parity and in parity 1 indicate that sex selective abortion is not a case for the 1^{st} born. Further, if I examine the parity-wise (that is holding the number of children of the respondent constant) sex ratio across birth orders in figure 2.1, I find that sex ratio is biased towards male child at the last birth in case of each parities and from parity three onwards this biasedness is clearly visible. The couples are more likely to decide to stop having another child if the last born is a son than a daughter, thus this implies that the stopping behavior is influenced by the sex of the last child. Therefore, on the one hand, the abortion pattern implies that couples, who already have as many daughters as their desired number of children in total, are more likely to have an abortion than who have mixed children, probably they fear to have another daughter exceeding their expected family size. On the other hand, parity-wise sex ratio of birth orders indicates that couples are more likely to stop having more children as soon as their demand for son is met. Both these together indicate that stopping behavior is related to both family size and preference for son; but do not provide evidence of sex selective

abortion at the 1^{st} birth. Rather, abortion behavior is more influenced by the family size and to stop getting more daughters. Thus, the skewedness in sex ratio is caused more by differential stopping behavior rather than sex-selective abortion.

The birth of the 1^{st} child is not influenced by differential stopping behavior. Referring to the exclusion condition, that is having at least a son seems the only channel through which gender of the 1^{st} born influences the outcome variable of women's bargaining status in the household after controlling for the other covariates which can influence women's bargaining power. Based, on the above arguments, gender of the 1^{st} born, or 1^{st} born a son can be a valid instrument for women having at least a son.

The control variables used throughout the analysis are different demographic variables of household, such as, located in rural/urban areas, female-headed, wealth index, sex ratio within household, religion, age difference between husband and wife, education difference between husband and wife, wife earning or not, whether interview was interfered or interrupted. To control for state level unobserved characteristics, such as social norms and customs, people's relation and dynamics, which can also influence individual's role and status within household, I use state fixed effects (FEs) in the estimation process.

2.5 Results

2.5.1 Descriptive Statistics

Table 2.3 shows summary of household characteristics. Around 30 per cent of the couples live in urban areas and 70 per cent in rural areas. By wealth index the households are categorized; around 40 per cent of the households fall in rich categories (19 per cent are richest and 20 per cent are richer); 21 per cent of the total sample belong to middle income/wealth group; 21 per cent and 18 per cent are poorer and poorest respectively. Majority of the households are Hindu, followed by Muslim and Christian households. Among total, 94 per cent household have male household head. Table 2.4 shows summary statistics of individual characteristics. The survey has covered women between age of 15-49 years. The mean age of women in the sample is 33 years and mean education

Table 2.3: Summary of Household Characteristics

Households Characteristics	in Percent
Type of Place of Residence	
Urban	29.60
Rural	70.40
Household Head	
Male	94.48
Female	5.52
Wealth index categories	
Poorest	18.13
Poorer	21.15
Middle	21.19
Richer	20.18
Richest	19.35
Religion	
Hindu	75.36
Muslim	12.99
Christian	6.79
Other	4.86

Table 2.4: Summary of Individual Characteristics

Variables	Obs	Mean	Std. Dev.	Min	Max
Wealth Index	58100	-0.0085	1.0001	-2.4115	2.8261
Sex ratio (No of women/No of men)	58100	0.5671	0.4083	0.0769	5
Women's Age (in years)	58100	33.6732	7.7457	15	49
Husband's Age (in years)	58100	38.3259	8.2130	17	54
Women's education (in years)	58100	5.7073	5.1079	0	20
Husband's education (in years)	58100	7.4182	4.9162	0	20
Whether Women earn	58100	0.2410	0.4330	0	1
Whether Husbands earn	58100	0.8837	0.3206	0	1
Children born	58100	2.7719	1.5432	1	15
Son	58100	1.3365	0.9535	0	9
Daughter	58100	1.2397	1.1098	0	10

is around 6 years. The husbands are of age between of 17 to 54 years, the mean age of husbands is 38 years and their mean education is 7 years. Considering the total sample of the survey which also includes the childless women, the average number of children born from a woman is 2.5, however excluding the childless women the average number of child per women is 2.8. Mean number of boy child ever born from a women is 1.3 and average number of girls ever born from a women is 1.2.

Table 2.5 reports that 82.9 per cent women have at least a son. Among women respondents, 82 per cent said that they have some say (independently and/or jointly with husband) and 15 per cent have independent say in spending their own money; 72 per cent have independent and/or joint say in spending their husband's money and only 5 per cent have independent say in spending husband's money. Around 75 per cent women have independent and/or joint say in any large purchase in the household whereas only 6 per cent women have independent say in large purchases. In visiting relatives, 76 per

cent women have a say independently and/or jointly with husband and 6 per cent have independent say in the same. In contraception use, around 91 per cent of women reported to have either independent and/or joint say with husband and 7.6 per cent can decide on their own about the same.

Using husbands' responses, table 2.5 shows that around 74 per cent and 84 per cent husbands think that wife should have either an independent or joint say with husband in spending husband's earning and wife's own earning respectively. Around 75 per cent husbands said that a wife should have a say independently or jointly with husband in household's large purchases and 79 per cent husbands think that wife should have an independent or joint say in visiting her family and friends. In deciding about number of children, 90 per cent husbands responded that wife should have a say either independently or jointly with husband. Among husbands, 8 per cent and 18 per cent of husbands think that wife should have an independent say in spending husband's money and wife's own earning respectively. Only 3 per cent husbands recognize that wife should have an independent say in decision of number of children to have.

Table 2.5: Women having at least a son and having say in HH decisions

Variables	in Percent	
Women having at least a son	82.9	
Interrupted by someone when interviewed	17.3	
Women's Say in Household	Independently/ jointly	Only
From Women's responses	with husband	Independently
Spending	82.24	15.05
Large purchases	75.37	6.26
Spending husband's money	72.57	5.17
Visiting relatives	76.45	6.55
Contraception use	91.62	7.61
Husband thinks that women should have a say in:	Independently/ jointly	Only
From Husband's responses	with husband	Independently
Spending	74.57	8.04
Large purchases	74.80	6.46
Wife's own money	84.75	18.07
Visiting relatives	79.27	9.48
No Of Children	90.01	3.06

2.5.2 Probit Estimation results:

Full sample of couples who have at least a child

This section presents the estimation results using the full sample of 58100 couples who have been parents at least once, and discusses the implications of these results. In all the estimations, the state FEs are used to control for the state level patriarchal values and other unobserved factors that may influence bargaining power of women within the household. Bargaining power of women in the same locality can be correlated with each other due to the factors that are common in a locality. Therefore the standard errors are clustered at the level of primary sampling units (PSU) in all the estimations. Also, bargaining power can be related to religious values and therefore, the religion dummies for three main religions of India (Hindu, Muslim and Christian) are used. Other control variables used in the estimations are location (rural or urban) of household; wealth index; women's earning status; gender of the household head; difference of age between husband and wife; difference in education years between husband and wife; ratio of female and male members in the household; and whether interrupted by household members during interview[3].

Table 2.6 shows the probit estimation results using 'women having a say' in various household decisions as the dependent variables in five separate regressions and 'having at least a son' as the main explanatory variable. The sample sizes of these estimations with five different indicators of say are not equal, because some questions were not eligible for all women, and not all women were asked all the questions of decision-makings. Mainly 'spending own money' question was asked conditional on women's earning status, and only 24 per cent women earn money in the total sample and were eligible for the question. Also, the question of spending husband's money was asked conditional on husbands' earning status. Similarly, husbands' were asked spending money questions conditional on husband's and wife's earning status. Due to different attributes of eligibility in the

[3]The multicollinearity between the covariates are tested with correlation coefficients between them and also with variance inflation factor, and multicollinearity is not found.

decision-making questions, sample sizes differ between estimations and I do not treat for missing observations in these estimations to equate the sample size.

The coefficients and marginal effects in table 2.6 panel A are estimated using women's responses and in panel B using husbands' responses. The marginal effects measure the impact of discrete change in the explanatory variable that is having no son to having at least a son on the probabilities of women having say in household decision. The panel A

Table 2.6: Probit Regression Results

Panel A. Women's responses about own say					
Dependent Variable:		Spending	Large	Husband's	Visiting	Contraception
Women's Say in:		Own Money	Purchase	Money	Relatives	Use
At least a son	Coefficients	-0.00819	0.0880***	0.0879***	0.0827***	0.0451
		(0.0416)	(0.0181)	(0.0179)	(0.0185)	(0.0360)
	Marginal effects	-0.0020	0.0265***	0.0241***	0.0280***	0.0067
		(0.0102)	(0.0054)	(0.0054)	(0.0057)	(0.0053)
Controls #		Yes	Yes	Yes	Yes	Yes
Observations		11085	43648	43319	43648	26706
Panel B. Husband's perception about women's say						
Dependent Variable:		Spending	Large	Wife's	Visiting	Number of
Women's Say in:		Husband's Money	Purchase	Money	Relatives	Children
At least a son	Coefficients	-0.0104	-0.0346*	0.0178	-0.0106	0.00327
		(0.0182)	(0.0184)	(0.0541)	(0.0190)	(0.0231)
	Marginal effects	-0.0032	-0.0107*	0.0040	-0.0029	0.00055
		(0.0056)	(0.0057)	(0.0122)	(0.0052)	(0.0039)
Controls#		Yes	Yes	Yes	Yes	Yes
Observations		42932	43648	6833	43648	43648

Note: Marginal effects represent change in outcome variables (women's say) due to discrete change in binary explanatory variable from 0 (having no son) to 1 (having at least a son).
: Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview.
Standard errors clustered at PSU level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

probit (uninstrumented) regression results indicate that women who have at least a son have significantly 2.6 per cent and 2.4 per cent higher probabilities to have say in large purchases and in spending husband's money respectively, compared to women who do not have a son. Also, the women with at least a son have significantly 2.8 per cent higher probabilities to have a say in visiting friends and relatives compared to women who have no son. In contraception use, women with at least a son have higher probabilities to have a say independently or jointly with husband compared to women without a son, but this coefficient remains statistically insignificant. In spending own money, women with a son do not have a higher say compared to women with no son. In Panel B, the (uninstrumented) regression results using husbands' responses indicate that husbands

with at least a son feel that wife should have a higher say either independently or jointly with husband in matters of spending their own earning and the number of children they want to have compared to husbands without a son, but these results are insignificant. However, in cases of decisions like spending husbands money, large purchases and visiting relatives, husbands with at least a son do not feel that wife should have higher say compared to husbands without a son. The full results, the marginal effects of full set of variables on women's say are given in appendix tables A2.2 and A2.3, using women's and husbands' responses respectively.

However, these results are suspected to be biased due to the potential endogeneity in the main explanatory variable 'having at least a son'. Therefore, to resolve the endogeneity issue, I use gender of the 1^{st} born or '1^{st} born a son' as an instrument for 'having at least a son'. Table 2.7 shows the instrumental variable probit (IV Probit) regression results. In panel A, IV Probit results using women's responses indicate that women with at least

Table 2.7: IV Probit Regression Results

Panel A. Women's responses about own say						
Dependent Variable:		Spending	Large	Husband's	Visiting	Contraception
Women's Say in:		Own Money	Purchase	Money	Relatives	Use
At least a son	Coefficients	-0.131	0.0814**	0.123***	0.142***	0.0101
		(0.107)	(0.0412)	(0.0404)	(0.0415)	(0.0953)
	Marginal effects	-0.0324	0.0245**	0.0391***	0.0413***	0.00149
		(0.0265)	(0.0124)	(0.0129)	(0.0121)	(0.0141)
First stage regression: *Dependent Variable: At least a son*						
Gender of the 1^{st} born		0.274***	0.328***	0.329***	0.328***	0.240***
		(0.00673)	(0.00362)	(0.00364)	(0.00362)	(0.00427)
Controls #		Yes	Yes	Yes	Yes	Yes
Observations		11085	43648	43319	43648	26706
Panel B. Husband's perception about women's say						
Dependent Variable:		Spending	Large	Wife's	Visiting	Number of
Women's Say in:		Husband's Money	Purchase	Money	Relatives	Children
At least a son	Coefficients	0.0505	-0.00459	0.0493	0.00847	0.0649
		(0.0407)	(0.0411)	(0.137)	(0.0427)	(0.0506)
	Marginal effects	0.0157	-0.00143	0.01112	0.00231	0.01091
		(0.0126)	(0.0128)	(0.0308)	(0.0116)	(0.0085)
First stage regression: *Dependent Variable: At least a son*						
Gender of the 1^{st} born		0.328***	0.328***	0.283***	0.328***	0.328***
		(0.00365)	(0.00362)	(0.00857)	(0.00362)	(0.00362)
Controls#		Yes	Yes	Yes	Yes	Yes
Observations		42932	43648	6833	43648	43648

Note: Marginal effects represent change in outcome variables (women's say) due to discrete change in binary explanatory variable from 0 (having no son) to 1 (having at least a son).
: Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview.
Standard errors clustered at PSU level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

a son have significantly higher probability to have say in different matters such as, 2.4 per cent higher chance to have say in large purchases; 3.9 per cent in spending husband's money and 4.1 per cent in visiting relatives and friends; compared to women with no son. Women with son(s) also may have higher say in contraception use, but this result is statistically insignificant. The impact of having at least a son on women's say in spending own earning is negative and insignificant.

When we consider intra-household bargaining over resources, it is the joint income and/or partner's income in which individual's bargaining should matter, as generally, individual would have higher say in own earning than anyone else. Therefore, when we want to examine whether having a son can impact women's say in household matters, it should be more appropriate to test on say in husband's money rather than say in own money. Because women, who have the freedom to work and earn, are expected to also have say in own money irrespective of having son or not. Thus, the outcome variable 'spending own money' can be considered as a placebo check for the household bargaining status of women. Therefore, the insignificant impact of having at least a son on 'spending own money' supports the placebo check. And, the positive and highly significant impact of having at least a son on women's say in husband's money implies that women's say significantly improves on average by 4 per cent due to having at least a son compared to having no son, keeping other things same.

The result on say in large purchases implies that the chances of having say in household purchases improve (by 2.4 per cent) for women with at least a son compared to women without a son. The survey didn't specify what can be considered as large purchases, the perception of large purchases may differ between men and women. For women it may happen that they consider consumer durables, jewelry, car etc as large purchases and women with son(s) may have higher say in these purchases compared to women who do not have a son. For men, they may consider purchase of stock, land, property etc as large purchases and they may or may not recognize women's voice in such purchases, irrespective of children composition they have.

Having a role in decisions of going to visit relatives/friends can be an indicator of women's

autonomy. Women with at least a son may have significantly higher freedom in mobility, on average by 4 per cent in comparison to women without a son. Women's autonomy can be restricted not only by her status in the household but can be due to general safety concern of going out. It may also have an implication on autonomy of women, who often take their children along with while going outside. In patriarchal society, women accompanying a boy child may gain higher autonomy compared to women with girl child. Taking daughters while going anywhere outside home may cause higher restrictions due to both safety concern and patriarchal norms, depending on the societal and local characteristics.

In the decision of child bearing, women with at least a son have higher probability of having a say in contraception use though the result is statistically insignificant. If the say in contraception use improves then it can be seen as after providing a son, decision to have more children or not and about family size can be trusted upon women as the demand for a son is already satisfied. But a woman, who does not have a son so far, also does not have a say in family planning matters until she succeeds to provide a son to the family. Using women's responses, the marginal effects of all covariates on say in decision-making from IV probit estimations are presented in appendix table A2.4.

In table 2.7 panel B, using husbands' responses, I find that the estimated marginal effects of having at least a son on husbands' views about women's say in a couple remain insignificant in all cases of household decisions. Earlier in section 2.3, I mentioned that women and husbands were asked the household decision questions in different manner, women were asked more direct questions about actual situation in the household, such as *'Who usually makes decisions about making major household purchases: mainly you, mainly your husband, you and your husband jointly, or someone else?'*, whereas the husbands were asked not in particularly about their wife, but wife's say in general, such as *'In a couple, who do you think should have the greater say in making major household purchases: the husband, the wife or both equally?'*. The insignificant impact of having at least a son on husbands' subjective views thus implies that his views about women's say in couples remain same irrespective of having son(s) or not. In other words, I can infer

that having son(s) can improve women's intra-household status on average by 4 per cent compared to women without a son, but this does not have an impact on husbands' views in general. Using husbands' responses, the (IV-Probit) estimated marginal effects of the full set of variables on women's say in different matters are presented in appendix table A2.5.

Table 2.8: Reduced Form Regression Results

Panel A. Women's responses about own say						
Dependent Variable:		Spending	Large	Husband's	Visiting	Child
Women's Say in:		Own Money	Purchase	Money	Relatives	Bearing
1^{st} born a son	Coefficients	-0.0362	0.0266**	0.0406***	0.0462***	0.00248
		(0.0295)	(0.0135)	(0.0133)	(0.0136)	(0.0228)
	Marginal effects	-0.00892	0.00802***	0.01295***	0.01350***	0.00037
		(0.0073)	(0.0041)	(0.0042)	(0.0040)	(0.0034)
Controls #		Yes	Yes	Yes	Yes	Yes
Observations		11085	43648	43319	43648	26706
Panel B. Husband's perception about women's say						
Dependent Variable:		Spending	Large	Wife's	Visiting	Child
Women's Say in:		Husband's Money	Purchase	Money	Relatives	Bearing
1^{st} born a son	Coefficients	0.0166	-0.00119	0.0140	0.00291	0.0214
		(0.0134)	(0.0135)	(0.0386)	(0.0140)	(0.0166)
	Marginal effects	0.00517	-0.00037	0.00316	0.00079	0.00359
		(0.0042)	(0.0042)	(0.0087)	(0.0038)	(0.0028)
Controls#		Yes	Yes	Yes	Yes	Yes
Observations		42932	43648	6833	43648	43648

Note: Marginal effects represent change in outcome variables (women's say) due to discrete change in binary explanatory variable from 0 (having 1^{st} born daughter) to 1 (having 1^{st} born son).
: Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members; religion; state fixed effects and interrupted by HH members during interview.
Standard errors clustered at PSU level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.8 shows the reduced form regression where the dependent variable is women's say and the main explanatory variable is 1^{st} born a son. Panel A presents estimated coefficients and marginal effects using women's responses and panel B provides results using husbands' responses. The results are similar as the IV probit, but magnitude of the effects of having 1^{st} born son is lower than the effects of (instrumented) at least a son. Women with 1^{st} born son are significantly more likely to have a higher say by, 0.8 per cent in large purchases, 1.3 per cent in spending husband's money and 1.3 per cent in visiting relatives / friends, compared to women with 1^{st} born daughter. The effect of 1^{st} born son on say in spending own money remains insignificant as the effect of at least a son. Using husbands' responses, the effects of 1^{st} born son on women's say remain insignificant as panel B in table 2.7 as well. I can infer that having 1^{st} born son can improve women's

status in intra-household decisions by 1 per cent on average compared to the women with 1^{st} born daughter, and also that 1^{st} born son does not significantly change husbands' general view about individual bargaining among couples. From reduced form regressions, the full set of marginal effects of variables on women's say are shown in appendix tables A2.6 and A2.7 using women's and husbands' responses respectively.

Outside reproductive cycle sample of couples

Table 2.9: IV Probit Regression Results - Out of reproductive cycle sample

Panel A. Women's responses about own say					
Dependent Variable:	Spending	Large	Husband's	Visiting	Child
Women's Say in:	Own Money	Purchase	Money	Relatives	Bearing
			Marginal effects:		
IV Probit regression:	*Dependent Variable: At least a son Instrumented with* 1^{st} *born a Son*				
At least a son	-0.0421	0.0344	0.0554**	0.0539**	-0.00072
	(0.0456)	(0.0249)	(0.0257)	(0.0247)	(0.0240)
Reduced form regression:	*Dependent Variable:* 1^{st} *born a Son*				
1^{st} born a son	-0.0074	0.0063	0.0102**	0.0099**	-0.00011
	(0.0080)	(0.0046)	(0.0048)	(0.0046)	(0.0036)
Controls#	Yes	Yes	Yes	Yes	Yes
Observations	9274	33864	33593	33864	23492
Panel B. Husband's perception about women's say					
Dependent Variable:	Spending	Large	Wife's	Visiting	Child
Women's Say in:	Husband's Money	Purchase	Money	Relatives	Bearing
			Marginal effects:		
IV Probit regression:	*Dependent Variable: At least a son Instrumented with* 1^{st} *born a Son*				
At least a son	0.0102	-0.00056	0.0149	0.0040	0.0152
	(0.0257)	(0.0258)	(0.0526)	(0.0237)	(0.0171)
Reduced form regression:	*Dependent Variable:* 1^{st} *born a Son*				
1^{st} born a son	0.00197	0.00005	0.00283	0.00080	0.00291
	(0.0047)	(0.0047)	(0.0096)	(0.0044)	(0.0031)
Controls#	Yes	Yes	Yes	Yes	Yes
Observations	33251	33864	5723	33864	33864

: Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview. Standard errors clustered at PSU level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

So far the results have used the total sample that includes both women who have stopped child bearing, and women who are still in the age of reproduction and may have more children in future. When I examine women's bargaining power in respect of their fertility output, it may be the case that women who are still in the reproductive cycle and have only daughters may get higher importance in household decisions in expectation that they may provide son in future than women who have only daughters and stopped child bearing. Women who have already decided herself or by family to not have more children,

either sterilized or declared in-fecund can be considered as outside reproductive cycle sample. If these women have only daughters, or only sons or both son(s) and daughter(s) when outside reproductive cycle, their reported role in decision-making will give us more accurate realization of bargaining power as without any expectation for getting preferred gender composition of children in future.

The results using sample of couples where women are already outside reproductive cycle, table 2.9 panel A shows the marginal effects of having son on women's say using women's responses and panel B presents marginal effects using husbands' responses. The results include both IV probit and reduced form estimations and both the results remain robust with the full sample results in tables 2.7 and 2.8. The magnitude of the impact of having at least a son increases using the out-of reproductive cycle sample. From women's responses, women with at least a son are more likely to have say; 3.4 per cent in large purchases, 5.5 per cent (significant) in husband's money and 5.4 per cent (significant) more chances to have say in visiting relatives/friends, compared to women with no son. The impact of having at least a son on say in spending own money remains negative and insignificant and therefore is consistent with placebo check. In this case, the say in use of contraception doesn't have important implication as these women are already out-of reproductive cycle. Overall, the results imply that the son-effect on the intra-household status is higher for women who have already completed the reproductive responsibilities (5 per cent), compared to women who continue child-bearing (4 per cent).

The results using husbands' responses are also similar as earlier, and larger in magnitude, though remain insignificant. The reduced form results from the out-of reproductive cycle sample show that women with 1^{st} born son have significantly higher say by 1 per cent in spending husband's money and visiting relatives/friends and higher but insignificant say in large purchases compared to women with 1^{st} born daughter.

Say indicator constructed with stricter definition - Women's independent say

In this section, the say indicators are constructed with stricter condition, that is women's say variables get value 1 only if women independently have a say, otherwise 0. Table 2.10

Table 2.10: IV Probit Regression: Women have an independent say

Panel A. Women's responses about own say					
Dependent Variable:	Spending	Large	Husband's	Visiting	Child
Women's Say in:	Own Money	Purchase	Money	Relatives	Bearing
			Marginal effects		
IV Probit regression: *Dependent Variable: At least a son Instrumented with* 1^{st} *born a Son*					
At least a son	0.00044	0.01136*	0.00990	0.00785	0.02061
	(0.0242)	(0.0068)	(0.0063)	(0.0071)	(0.0136)
Reduced form regression: *Dependent Variable:* 1^{st} *born a Son*					
1^{st} born a Son	0.00025	0.00373*	0.00318	0.00244	0.00493
	(0.0066)	(0.0022)	(0.0020)	(0.0023)	(0.0032)
Controls#	Yes	Yes	Yes	Yes	Yes
Observations	11085	43648	43319	43648	26652
Panel B. Husband's perception about women's say					
Dependent Variable:	Spending	Large	Wife's	Visiting	Child
Women's Say in:	Husband's Money	Purchase	Money	Relatives	Bearing
			Marginal effects		
IV Probit regression: *Dependent Variable: At least a son Instrumented with* 1^{st} *born a Son*					
At least a son	0.00959	0.00335	-0.01043	-0.00986	-0.00124
	(0.0080)	(0.0071)	(0.0326)	(0.0086)	(0.0050)
Reduced form regression: *Dependent Variable:* 1^{st} *born a Son*					
1^{st} born a Son	0.00316	0.00115	-0.00293	-0.00326	-0.00042
	(0.0026)	(0.0023)	(0.0092)	(0.0028)	(0.0016)
Controls#	Yes	Yes	Yes	Yes	Yes
Observations	42878	43594	6813	43594	43594

: Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview. Standard errors clustered at (PSU) level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

shows both IV and reduced form estimated impacts of having at least a son and having 1^{st} born son respectively on women's independent say in household matters. In panel A from women's responses, I found that women who have at least a son are more likely to have an independent say in matters of spending own and husband's money, visiting relatives, contraception use and significantly higher independent say in large purchases compared to women without a son, keeping other things same. However from husbands' responses, the results in panel B imply that husbands with son(s) do not feel statistically different from husbands without a son about wife having an independent say in different household matters. The reduced form estimated marginal effects of having 1^{st} born a son on women's independent say also give similar results as the IV estimations. Women with

1^{st} born son have higher independent say in large purchases than women with 1^{st} born daughter, whereas husbands with a 1^{st} born son do not think any differently about wife's say on household matters from husbands with 1^{st} born daughter.

Individual Women Sample Results:

This section considers the individual women sample of the survey, including women selected as couples that is their husbands were interviewed separately, and also women selected as individuals which meant their husbands were not interviewed.

Table 2.11: Probit Regression Results - Women Sample

Dependent Variable: Women's Say in:	Spending Own Money	Large Purchase	Husband's Money	Visiting Relatives	Child Bearing
			Marginal Effects		
IV Estimation					
At least a son	-0.0355	0.0163	0.0320***	0.0393***	-0.0116
	(0.0238)	(0.0114)	(0.0118)	(0.0112)	(0.0139)
Reduced form estimation					
1^{st} born a son	-0.0099	0.0053	0.0106***	0.0127***	-0.0027
	(0.0066)	(0.0037)	(0.0038)	(0.0036)	(0.0032)
Controls #	Yes	Yes	Yes	Yes	Yes
Observations	13087	52364	51946	52364	30632

Note: Marginal effects represent change in outcome variables due to discrete change of binary explanatory variables from 0 (having no son) to 1 (having at least a son). # : Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview. Standard errors clustered at Primary Sampling Unit (PSU) level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The IV probit estimations in table 2.11 indicate that women with at least a son have higher probability to have say, 1.6 per cent higher probability of having say in large purchases, 3.2 per cent significantly higher probability in husband's money and 3.9 per cent higher in visiting relative/friends compared to women without a son. And, son-effect remains insignificant in case of women's say in spending own money and child bearing. The results with women's individual sample remain robust with the couple's result of the possibility of 4 per cent improvement in say for women due to having at least a son. Similarly, women with 1^{st} child son have significantly higher probability to have say in husband's money and visiting relatives by around 1 per cent compared to women with 1^{st} child daughter.

2.5.3 Ordered Probit Estimation: Using Categories of Responses

Table 2.12: Ordered IV Probit Regression: Couples Sample

Panel A. Women's responses about own say					
Dependent Variable: Women's Say in - At least a Son	Spending Own Money	Large Purchase	Husband's Money	Visiting Relatives	Child Bearing
Marginal effects: Pr[Say | At least a son]					
Only woman has a say	0.00380	0.01391***	0.00975***	0.01444***	0.00468
	(0.0161)	(0.0041)	(0.0037)	(0.0044)	(0.0071)
Woman has a joint say with husbands	0.00007	0.02156***	0.01996***	0.01980***	0.00054
	(0.0003)	(0.0064)	(0.0076)	(0.0060)	(0.0008)
Only husband have a say	-0.00357	-0.02743***	-0.02686***	-0.02770***	-0.00506
	(0.0152)	(0.0081)	(0.0102)	(0.0083)	(0.0077)
Someone else has a say	-0.00030	-0.00804***	-0.00285***	-0.00655***	-0.00016
	(0.0013)	(0.0024)	(0.0011)	(0.0020)	(0.0002)
Observations	43648	43648	43648	43648	43648
Panel B. Husband's perception about women's say					
Dependent Variable: Women's Say in - At least a Son	Spending Husband's Money	Large Purchase	Wife's Money	Visiting Relatives	Child Bearing
Marginal effects: Pr[Say | At least a son]					
Only woman has a say	0.00187	-0.00040	0.00098	-0.00210	-0.00040
	(0.0051)	(0.0042)	(0.0226)	(0.0057)	(0.0029)
Woman has a joint say with husbands	0.00212	-0.00062	-0.00013	-0.00145	-0.00062
	(0.0058)	(0.0066)	(0.0031)	(0.0039)	(0.0045)
Only husband have a say	-0.00386	0.00091	-0.00082	0.00340	0.00095
	(0.0106)	(0.0097)	(0.0190)	(0.0093)	(0.0069)
Someone else has a say	-0.00014	0.00010	-0.00002	0.00015	0.00006
	(0.0004)	(0.0011)	(0.0005)	(0.0004)	(0.0004)
Observations	43648	43648	43648	43648	43648

Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview. Standard errors clustered at PSU level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

So far, the 'say' variables were constructed considering that either women have an independent say and/or women have a say jointly with husband getting value 1, otherwise if only husband has a say or someone else in the household has a say then say variables get 0. For further analysis in detail, the estimations in this section keep the four answer categories intact, that is (i) only woman has a say; (ii) woman has a joint say with husband; (iii) only husband has a say; and (iv) someone else in the family has a say/others. For this analysis ordered probit regression method is used. Considering the endogenous nature of the main explanatory variable (having at least a son), IV probit regression method is used by instrumenting *having at least a son* with 1^{st} *born a son.*

Table 2.12 shows the ordered IV probit estimated marginal effects of having at least a

son on women's say, couple's joint say, husbands' say and someone else's say in different household matters. The panel A results are estimated using women's responses and panel B results use husbands' responses. The impact of having at least a son on say in different matters in the household can be explained as follows:

- *Large purchases*: Due to having at least a son, the probability of women's independent say (highly significantly) increases by 1.4 per cent; the probability of their joint say with husband also significantly increases by 2.2 per cent; but the probability of husband's independent say in the matter decreases by 2.7 per cent comparing with couples who do not have a son.

- *Spending husband's money*: Women who have at least a son, the likelihood of their independent say in spending their husband's earning significantly increases by around 1 per cent; the likelihood of having joint say with husband significantly increases by 2 per cent; whereas the likelihood of their husband's independent say in spending their own earning declines by 2.7 per cent and also the likelihood of someone else say declines by 0.3 per cent, when compared with the couples who have no son.

- *Spending wife's own money*: In this matter, the change in individuals' say in couples due to son-effect is very small in magnitude and remains statistically insignificant.

- *Visiting women's relatives and/or friends*: When comparing with couples without a son, women with son(s) are significantly more likely to have an independent say by 1.4 per cent; to have a joint say with husband by 2 per cent; and the probability of husband's independent say declines by 2.8 per cent.

- *Contraception use*: The couples who have at least a son, the likelihood of women's independent say in contraception use increases by 0.4 per cent; of having joint say with husband increases by very small amount and likelihood of husband's independent say in the matter decreases by 0.5 per cent when compared with couples without a son. These effects are statistically insignificant.

Using husbands' responses, the marginal effects of having at least a son on individuals' say in general among couples are small in magnitude and statistically insignificant and thus imply that husbands with at least a son do not feel differently from husbands with no son in matters of individual bargining status among couples in general. In table 2.13 panel

Table 2.13: Ordered Reduced Form Regression - Couple Sample

Panel A. Women's responses about own say					
Dependent Variable: Women's Say in - 1^{st} born a Son	Spending Own Money	Large Purchase	Husband's Money	Visiting Relatives	Child Bearing
Marginal effects: Pr[Say | 1^{st} born a son]					
Only woman has a say	-0.00395	0.00391***	0.00328***	0.00504***	0.00255
	(0.0051)	(0.0013)	(0.0012)	(0.0014)	(0.0023)
Woman has a joint say with husbands	-0.00042	0.00605***	0.00670***	0.00694***	0.00023
	(0.0005)	(0.0021)	(0.0025)	(0.0020)	(0.0002)
Only husband have a say	0.00401	-0.00770***	-0.00902***	-0.00969***	-0.00269
	(0.0052)	(0.0027)	(0.0033)	(0.0027)	(0.0025)
Someone else has a say	0.00036	-0.00226***	-0.00096***	-0.00229***	-0.00009
	(0.0005)	(0.0008)	(0.0003)	(0.0006)	(0.0001)
Observations	11085	43648	42884	43648	26706
Panel B. Husband's perception about women's say					
Dependent Variable: Women's Say in - 1^{st} born a son	Spending Husband's Money	Large Purchase	Wife's Money	Visiting Relatives	Child Bearing
Marginal effects: Pr[Say | 1^{st} born a son]					
Only woman has a say	0.00325*	0.00019	0.00021	-0.00087	0.00077
	(0.0017)	(0.0014)	(0.0072)	(0.0019)	(0.0009)
Woman has a joint say with husbands	0.00373*	0.00030	-0.00002	-0.00060	0.00120
	(0.0019)	(0.0022)	(0.0006)	(0.0013)	(0.0015)
Only husband have a say	-0.00674*	-0.00045	-0.00018	0.00141	-0.00185
	(0.0035)	(0.0032)	(0.0064)	(0.0031)	(0.0023)
Someone else has a say	-0.00024*	-0.00005	-0.00005	0.00006	-0.00012
	(0.0001)	(0.0004)	(0.0002)	(0.0001)	(0.0001)
Observations	42932	43648	6839	43648	43648

Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview. Standard errors clustered at PSU level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A, from the reduced form of ordered probit regression results using women's responses, it is evident that women with 1^{st} born son are more likely to have an independent say; such as 0.4 per cent (significant) in large purchases; 0.3 per cent (significant) in husband's money; 0.5 per cent (significant) in visiting relatives; and 0.2 per cent in contraception use when compared with women with 1^{st} born daughter. Also, women with 1^{st} born son have around 0.6 to 0.7 per cent higher probabilities to have a joint say with husband in matters of large purchases, spending husband's money, and visiting relatives/friends,

comparing with women having 1^{st} born daughter. However, husbands who have at least a son are significantly less likely to have an independent say such as, 0.8 per cent less in large purchases, 0.9 per cent lower in husband's own money and around 1 per cent lower in visiting relatives/friends comparing with them whose 1^{st} born is a girl. In panel B, using husbands' responses, the reduced form results indicate that husbands with 1^{st} born son are more likely to feel that women should have significantly higher independent say and also joint say with husband in spending husband's money, comparing with husbands who have 1^{st} born daughter. The rest of the results using husbands' responses remain insignificant.

2.5.4 Impact of having son on domestic violence to mothers

Domestic violence can be considered as an indicator of women's household status, and it is inversely related to women's position in the household. It can be perceived that women with higher status are less likely to face domestic violence compared to women who have lower status within the household. Literature has examined different determinants of intra-household violence, such as age (Ahmed, 2005), education, social and economic status, religion and caste, rural/urban location, customs (of patrilocality and dowry) (Koenig, Ahmed, Hossain, & Mozumder, 2003); property ownership (Panda & Agarwal, 2005), autonomy (Lamichhane, Puri, Tamang, & Dulal, 2011), and also infertility and not having a son (Ali and Bustamante Gavino (2007), Sami and Ali (2012) Ghouri and Abrar (2010)). However, the paper by (Koenig et al., 2003) menioned that "the effects of women's status on violence were found to be highly context-specific" and in more culturally conservative areas, women's higher autonomy can increase the risk of intra-household violence.

With the purpose of this paper, it would be interesting to examine the impact of having son on domestic violence to mothers. NFHS data have information on domestic violence from women respondents, and categorized the violent incidences into severe and less severe violence; and emotional and sexual violence. In this paper, I use incidences of severe and less severe domestic violence. Less severe violence includes incidences of push,

pull, punch, twist and slap, and severe violence includes incidences of kick, strangulation, burn and threatening with knife/gun. To estimate the son-impact on domestic violence, I use the following equation:

$$Pr(Y_i^D = 1|X_i, C_i) = \phi(\beta X_i + \gamma C_i + \epsilon_i) \tag{2.2}$$

In equation (2.2), the dependent variable Y^D refers to severe and/or less severe violence within household; i represents individual respondents; X denotes women having son(s), and C refers to other covariates.

Table 2.14 shows the impact of having son on probability of domestic violence. For these estimations, I use subsamples of all women, women inside the reproductive cycle, and women outside the reproductive cycle. The estimated marginal effects in the all

Table 2.14: Son-effect on Domestic Violence: Using Women's responses in Couple Sample

Dependent Variable:	Severe	Less Severe	Severe	Less Severe	Severe	Less Severe
Domestic Violence:	Violence	Violence	Violence	Violence	Violence	Violence
	All Women		Inside Reproductive Cycle		Outside Reproductive Cycle	
At least a son	-0.0019	0.0143	-0.0172***	-0.0064	0.0088	0.0244
	(0.0095)	(0.0171)	(0.0067)	(0.0145)	(0.0201)	(0.0358)
1^{st} born a son	-0.0011	0.0037	-0.0142**	-0.0062	0.0013	0.0038
	(0.0030)	(0.0054)	(0.0058)	(0.0112)	(0.0035)	(0.0063)
Controls #	Yes	Yes	Yes	Yes	Yes	Yes
Observations	43594	43594	9728	9728	33822	33822

Note: Full set of control variables are used in regressions. Standard errors in parentheses are clustered at PSU level, and take into account domestic violence survey weights. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

women sample indicate that the impact of having son(s) on domestic violence is insignificant. However, when I examine the women who are still in the reproductive cycle, having son(s) can reduce the likelihood to face any severe violence compared to women who do not have a son. More specifically, having at least a son can reduce the probability of severe violence by 1.7 per cent and having 1^{st} born son can reduce the likelihood of severe domestic violence by 1.4 per cent compared to women who have no son and women with 1^{st} born daughter respectively. Women with a son do not significantly differ in facing less severe violence than women with no son. For the out-of-reproductive cycle women the son-effect on domestic violence remains insignificant.

Table 2.15: Son-effect on Severe Domestic Violence: Across age of the last born

Dependent Variable:	Severe Violence			
Age of the last born (years)	less than 2	2 - 5	6 - 10	more than 10
Women in the Reproductive cycle				
At least a son	-0.0204**	-0.0199*	0.0002	-0.0111
	(0.0092)	(0.0107)	(0.0238)	(0.0235)
1^{st} born a son	-0.0172**	-0.0180*	0.0007	-0.0074
	(0.0084)	(0.0097)	(0.0173)	(0.0155)
Controls #	Yes	Yes	Yes	Yes
Observations	3928	3101	912	760

Note: Full set of control variables are used in regressions. Standard errors in parentheses are clustered at PSU level, and take into account domestic violence survey weights. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

As it is evident that son-effect is significant in case of women who are in the reproductive cycle, and becomes insignificant for women outside reproductive cycle, it may also indicate that son-effect disappears as child's age increases. To further examine this, I estimate the impact of having son(s) on domestic violence across the age of the last born child. The results in table 2.15 provide evidence that for very recent mothers the son-effect is highly significant and the effect disappears beyond last-born's age of 5 years. Apart from this, women's age has a highly significant impact on domestic violence that is as age increases women are more likely to face violence, compared to women at younger ages (results are not included in the paper, can be available on request).

In comparison to son-effect on women's intra-household say, it is found that the having son(s) can reduce the likelihood of violence, especially severe violence when women are recent mothers, but as women's age increases and children get older the son-effect on violence disappears, whereas due to having son(s) women's say improves whether they have completed reproductive cycle or still in the reproductive cycle.

2.5.5 Other important factors to women's bargaining power within household

This section discusses the impacts of other crucial factors on women's say in different decision making within households [4]. Also to examine the heterogeneity in the son-effect

[4]The marginal effects of the full set of variables are given in appendix tables A2.4 and A2.6 using women's responses and in tables A2.5 and A2.7 using husbands' responses.

with respect to other covariates, I use interaction terms of different control variables with having son(s). The results from IV-probit and reduced form estimations using women's responses are given in tables 2.16 and 2.17 respectively. The inclusion of all interaction terms with full set of control variables does not provide a good fit statistics of the models, when checked with information criteria (both BIC and AIC)[5]. However, I still want to perform the analysis with interaction terms between variables and having son(s) to examine how the son-effect acts with respect to these variables.

Location (Rural/Urban): In rural areas women have (highly) significantly lower bargaining power within household compared to urban areas. More specifically, village women have significantly lower chances to have say in different household decisions, such as 2.5 per cent lower probability in large purchases; 1.8 per cent lower probability in husband's earning; 2 per cent lower probability in visiting their friends and relatives; and even 3.7 per cent lower probability to have say in spending own earning; compared to women living in cities (tables A2.4 and A2.6). Compared to urban husbands, the rural husbands are significantly less likely to think that wife should have say in household matters, such as, around 2 per cent lower probability in large purchases and in spending wife's own money (tables A2.5 and A2.7). Due to having son(s), the rural women do not significantly differ in having say in household matters compared to the similar urban women. However due to having 1^{st} born daughter, the rural women have significantly lower say in household matters, such as 3.2 per cent lower say in large purchases, 2.5 per cent lower say in husband's money, 2.7 per cent lower say in visiting relatives/friends and even 4.6 per cent lower say in own money, when compared to urban women with 1^{st} born daughter (table 2.17).

Wealthy Households: Women in wealthy households have significantly higher bargaining power compared to women in poor households. From estimated results (tables A2.4 and A2.6), women in wealthy families have 2.6 per cent higher probability of having say

[5]The results from the fitstat of information criteria provide only support for interaction terms of having a son with rural and education difference between couples in case of say in husband's money, and for other say variables, fitstats do not support for inclusion of the interaction terms of control variables with having a son.

Table 2.16: Marginal Effects from IV-Probit Estimations with Interaction terms: Women Responses

	Dependent variable: Women's say in -				
Explanatory Variables	Spending Own Money	Large Purchase	Husband's Money	Visiting Relatives	Contraception Use
At least a son	-0.7091	-0.0459	0.3690	0.4249	-0.0673
	(0.5193)	(0.2713)	(0.2788)	(0.2626)	(0.2694)
Rural HH	-0.1562*	-0.0357	0.0189	0.0370	-0.0024
	(0.0849)	(0.0414)	(0.0425)	(0.0402)	(0.0490)
At least a son * Rural	0.1426	0.0133	-0.0454	-0.0701	0.0041
	(0.1006)	(0.0500)	(0.0512)	(0.0484)	(0.0557)
Wealth	-0.0188	0.0248*	0.0390***	0.0504***	0.0146
	(0.0247)	(0.0145)	(0.0149)	(0.0140)	(0.0202)
At least a son * Wealth	0.0521*	0.0076	-0.0169	-0.0236	-0.0002
	(0.0274)	(0.0169)	(0.0175)	(0.0164)	(0.0224)
Earning Women		0.0747***	0.1204***	0.0845***	0.0034
		(0.0167)	(0.0168)	(0.0160)	(0.0181)
At least a son * Earning Women		-0.0132	-0.0296	-0.0302*	0.0026
		(0.0193)	(0.0199)	(0.0188)	(0.0201)
Female Headed	-0.0872**	-0.0008	0.0185	-0.0050	0.0104
	(0.0435)	(0.0223)	(0.0230)	(0.0219)	(0.0235)
At least a son * Female Head	0.0857*	-0.0126	-0.0336	-0.0189	-0.0289
	(0.0493)	(0.0257)	(0.0263)	(0.0250)	(0.0256)
Age Difference	-0.0064	-0.0010	0.0026	0.0028	-0.0017
	(0.0054)	(0.0026)	(0.0027)	(0.0026)	(0.0028)
At least a son * Age Diff	0.0090	0.0007	-0.0039	-0.0039	0.0016
	(0.0064)	(0.0032)	(0.0033)	(0.0031)	(0.0033)
Education Difference	-0.0055**	-0.0052***	-0.0052***	-0.0045***	-0.0007
	(0.0025)	(0.0012)	(0.0013)	(0.0012)	(0.0013)
At least a son * Edu Diff	0.0017	0.0019	0.0021	0.0019	-0.0007
	(0.0026)	(0.0014)	(0.0014)	(0.0013)	(0.0014)
HH Sex Ratio	-0.0594	0.0366	0.0680**	0.0744**	-0.0180
	(0.0564)	(0.0310)	(0.0319)	(0.0299)	(0.0326)
At least a son * Sex ratio	0.0846	-0.0074	-0.0561	-0.0712*	0.0229
	(0.0712)	(0.0395)	(0.0408)	(0.0381)	(0.0390)
Hindu	-0.4526	-0.0605	0.2018	0.2204	-0.0223
	(0.3248)	(0.1617)	(0.1666)	(0.1575)	(0.1620)
At least a son * Hindu	0.5047	0.0654	-0.2547	-0.2862	0.0522
	(0.3794)	(0.1958)	(0.2013)	(0.1901)	(0.1858)
Muslim	-0.5006	-0.1143	0.1497	0.1710	-0.0442
	(0.3383)	(0.1677)	(0.1737)	(0.1638)	(0.1681)
At least a son * Muslim	0.5037	0.0724	-0.2501	-0.2825	0.0632
	(0.3949)	(0.2031)	(0.2094)	(0.1973)	(0.1929)
Christian	-0.4845	-0.0256	0.2414	0.2519	-0.0014
	(0.3299)	(0.1626)	(0.1670)	(0.1580)	(0.1648)
At least a son * Christian	0.5324	0.0574	-0.2666	-0.2750	0.0302
	(0.3830)	(0.1954)	(0.2007)	(0.1899)	(0.1892)
Interrupted by someone	-0.0741**	-0.0455***	-0.0536***	-0.0385***	-0.0266
	(0.0334)	(0.0147)	(0.0156)	(0.0148)	(0.0162)
At least a son * Interrupted	0.0270	0.0041	0.0012	-0.00305	-0.0205
	(0.0357)	(0.0162)	(0.0170)	(0.0160)	(0.0170)
Observations	11085	43648	43319	43648	26706

Note: State fixed effects are used and standard errors in parentheses are clustered at PSU level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.17: Marginal Effects from Probit Estimations with Interaction terms: Women Responses

	Dependent variable: Women's say in -				
Explanatory Variables	Spending Own Money	Large Purchase	Husband's Money	Visiting Relatives	Contraception Use
1^{st} born son	-0.0078	0.0006	0.0268	-0.0083	0.0412**
	(0.0419)	(0.0223)	(0.0232)	(0.0222)	(0.0170)
Rural HH	-0.0461***	-0.0323***	-0.0246***	-0.0273***	0.0038
	(0.0149)	(0.0084)	(0.0087)	(0.0083)	(0.0071)
1^{st} born son * Rural	0.0167	0.0142	0.0123	0.0130	-0.0045
	(0.0193)	(0.0103)	(0.0107)	(0.0100)	(0.0088)
Wealth	0.0219***	0.0273***	0.0230***	0.0293***	0.0158***
	(0.0069)	(0.0038)	(0.0040)	(0.0038)	(0.0033)
1^{st} born son * Wealth	0.0083	0.0063	0.0025	0.0016	-0.0022
	(0.0084)	(0.0046)	(0.0048)	(0.0046)	(0.0040)
Earning Women		0.0650***	0.0961***	0.0619***	0.0084
		(0.0075)	(0.0078)	(0.0074)	(0.0060)
1^{st} born son * Earning Women		-0.0015	0.0019	-0.0016	-0.0050
		(0.0098)	(0.0102)	(0.0096)	(0.0076)
Female Headed	-0.0105	-0.0153	-0.0228*	-0.0369***	-0.0097
	(0.0251)	(0.0130)	(0.0136)	(0.0127)	(0.0116)
1^{st} born son * Female Head	-0.0144	0.0070	0.0254	0.0292*	-0.0092
	(0.0333)	(0.0179)	(0.0187)	(0.0176)	(0.0157)
Age Difference	-0.0005	-0.0014	-0.0003	-0.0015*	0.0005
	(0.0016)	(0.0009)	(0.0009)	(0.0009)	(0.0007)
1^{st} born son * Age Diff	0.0028	0.0016	-0.0005	0.0022**	-0.0015
	(0.0020)	(0.0012)	(0.0012)	(0.0011)	(0.0010)
Education Difference	-0.0048***	-0.0036***	-0.0044***	-0.0037***	-0.0009
	(0.0013)	(0.0007)	(0.0007)	(0.0007)	(0.0006)
1^{st} born son * Edu Diff	0.0009	0.0003	0.0020**	0.0019**	-0.0008
	(0.0017)	(0.0009)	(0.0010)	(0.0009)	(0.0008)
HH Sex Ratio	0.0181	0.0390***	0.0320***	0.0267***	0.0079
	(0.0150)	(0.0084)	(0.0085)	(0.0081)	(0.0070)
1^{st} born son * Sex ratio	-0.0269	-0.0241**	-0.0209*	-0.0205*	-0.0187*
	(0.0219)	(0.0122)	(0.0125)	(0.0118)	(0.0102)
Hindu	-0.0173	-0.0078	0.0035	-0.0202	0.0337***
	(0.0292)	(0.0159)	(0.0170)	(0.0166)	(0.0126)
1^{st} born son * Hindu	-0.0089	0.0029	-0.0191	0.0091	-0.0186
	(0.0371)	(0.0194)	(0.0201)	(0.0196)	(0.0143)
Muslim	-0.0407	-0.0496***	-0.0413**	-0.0620***	0.0220
	(0.0358)	(0.0181)	(0.0191)	(0.0186)	(0.0147)
1^{st} born son * Muslim	-0.0579	-0.0067	-0.0241	0.0022	-0.0197
	(0.0456)	(0.0219)	(0.0225)	(0.0219)	(0.0167)
Christian	-0.0353	0.0282	0.0357	0.0187	0.0362*
	(0.0360)	(0.0218)	(0.0228)	(0.0221)	(0.0192)
1^{st} born son * Christian	0.0109	-0.0106	-0.0239	0.0104	-0.0195
	(0.0457)	(0.0258)	(0.0261)	(0.0259)	(0.0223)
Interrupted by someone	-0.0459***	-0.0439***	-0.0658***	-0.0411***	-0.0435***
	(0.0142)	(0.0082)	(0.0085)	(0.0080)	(0.0069)
1^{st} born son * Interrupted	-0.0035	0.0032	0.0237**	-0.0006	-0.0022
	(0.0183)	(0.0104)	(0.0108)	(0.0102)	(0.0086)
Observations	11085	43648	43319	43648	26706

Note: State fixed effects are used and standard errors in parentheses are clustered at PSU level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

in spending own money; 3.1 per cent more likelihood of having say in large purchases; 2.5 per cent higher probability to have say in husband's money; 3 per cent higher probability in say of visiting relatives and friends and 1.5 per cent higher probability in decision of contraception use compared to women in poor households. Wealthy husbands are also significantly more supportive to wife's say in a couple, such as 2.1 per cent more likely in spending husband's money; 2 per cent more likely in large purchases; 1 per cent more likely in spending wife's own earning; 2.7 per cent more likely in visiting her relatives and friends and 2.3 per cent highly likely in deciding number of children compared to husbands from poor households (tables A2.5 and A2.7). The marginal effects of wealth interacted with women having son(s) imply that in wealthy families women with son(s) have higher say in own money than women with son(s) in poor families. This can be explained that in wealthy families when resources are not constrained women with son(s) can enjoy higher say in own earning, but in poor families with constrained resources women can not enjoy full liberty to spending own money. But in wealthy families, even women without a son have significantly higher say in different matters, such as in 2.5 per cent higher say in large purchase, 3.9 per cent higher say in husband's money and 5 per cent higher say in visiting relatives/friends compared to women without a son in poor household (table 2.16). For women without a son in wealthy families, the marginal effect for spending own money remains insignificant, which serves as a placebo check, as in wealthy household women have higher say in own money anyway. The results remain similar for interaction terms between wealth and 1^{st} born son (table 2.17).

Earning Women: Women who earn money have significantly higher bargaining within household compared to women who do not earn. The estimated results indicate that earning women have significantly higher probability of around 6.3 per cent to have say in large purchases; 9.6 per cent higher probability to have say in husband's earning; 6 per cent higher say in visiting friends and relatives; and only 0.6 per cent (insignificant) more likely to have say in contraception use compared to women who do not earn (tables A2.4 and A2.6). Similarly husbands of earning women are more likely to feel that wife should have say in matters, such as 2.3 per cent significantly more likely to feel that in a couple

wife should have say in spending husband's money and 1.1 per cent significantly more likely think that wife should have say in large purchases compared to husbands whose wife is home-maker (tables A2.5 and A2.7). Even after having no son, earning women have 7.5 per cent significantly higher say in large purchases; 12 per cent significantly higher say in husband's money and 8.4 per cent significantly higher say in visiting relatives compared to non-earning women with no son (table 2.16). Using interaction terms between earning status and having son(s), I find that due to having son earning women's say in household does not statistically differ from non-earning women with son(s). The reduced form results in table 2.17 indicate similar impact of earning status of women on their say conditional on having 1^{st} born son.

Age-difference between couples: From women's responses, the results indicate that women, who are much younger than their husbands that is higher age difference from husband, are less likely to have much say in household matters compared to women who have husband of less age difference. These results are statistically insignificant. Husbands, who are much more elder than their wife, are less likely to recognise wife's say in matters like spending husband's money and/or wife's own money, visiting her friends and relatives etc compared to husbands whose wife is of similar age. Also, using interaction terms between having son(s) with age difference, the insignificant results imply that the age difference between couple is not statistically important determinant for women's say, conditional on having son(s).

Education difference between couples: The higher the education difference between husbands and wives, the significantly lower will be the bargaining power of women, compared to women with lower education difference with husbands. Women, who are lower educated than their husbands, are significantly less likely to have say in household matters, around 0.1 - 0.4 per cent lower probability to have say compared to women who are equally educated as their husband (tables A2.4 and A2.6). From husbands' responses it seems that the education differences between husband and wife do not significantly influence much of the husbands' thinking about who should have higher say in household matters (tables A2.5 and A2.7). Women with lower education compared to their

husbands, if having no son, have significantly lower probability to have say in different household matters compared to women without son but similarly educated as husband. And, having son does not significantly improve women's status, no matter women is lower educated than husbands or not (tables 2.16 and 2.17).

Ratio of women compared to men: The higher the ratio of women compared to men within households, the higher will be women's say within households compared to lower ratio of women to men. However, from husbands' responses it seems that sex ratio within household does not significantly influence husbands' perception about who should have higher say in household decision makings (tables A2.4 to A2.7). Using interaction terms, it is found that in household with larger number of women, even after having no son women may have significantly higher say, by 6.8 per cent in husband's money and 7.4 per cent in visiting relatives compared to women in household with lower female to male sex ratio (table 2.16). However, in households with larger number of women than men, women if having 1^{st} born son may have significantly lower say in household matters by 2 per cent compared to women with 1^{st} born son in a male majority (in number) household (table 2.17).

Female headed household: The female headed household does not improve women's intra-household say compared to male headed household, rather it is found that in female headed household, women have significantly lower probability to have say in visiting relatives (by 2 per cent) and in contraception use (by 1.5 per cent) compared to male dominated household (tables A2.4 and A2.6). Further, examining by interaction terms with having son(s), it is found that in female headed household, if women have 1^{st} born daughter then they have significantly 2.3 per cent lower say in husband's money and 3.7 per cent lower say in visiting relatives compared to women with 1^{st} born daughter in male-headed household. And, having 1^{st} born son in female headed household can significantly improve women's autonomy to meet relatives/friends by 3 per cent than women having 1^{st} born son in male-headed households (table 2.17). By having at least a son, women's say in own money can improve significantly by 8.6 per cent in female-headed household and not-having at least a son can reduce the say by 8.7 per cent in

female-headed household compared to male-headed household (table 2.16). Therefore, from these results of female headed household, we can say that female members may also promote patriarchal values and customs and female dominance in the household does not automatically indicate higher status for women in such household.

Religion: Muslim households have significantly lower bargaining power for women in household decisions compared to households in other religion. Among Hindu, Muslim and Christian households, muslim women have highly significantly lower say in many household matters, such as 7.1 per cent significantly lower probability to have say in spending own money; 5.4 per cent significantly lower probability in large purchases; significantly 5.6 per cent lower likelihood to have say in husband's money; and 6.2 per cent significantly lower chance in say of visiting relatives and friends; however 1.1 per cent more likely to have say in number of children compared to women in other religions. Hindu women also have lower probability to have say in different household matters, but these results are insignificant. However, it seems that Christian women have (around 2.2 to 3 per cent) more chances to have say in different household matters, but most of these results for Christian women remain insignificant (tables A2.4 and A2.6). The results estimated from husbands' responses, it is clear that Muslim husbands are less likely to feel that women should have an independent or joint say with husband in household decisions, such as, 2.7 per cent less likely in spending husband's earning; 1.5 per cent lower probability in large purchases; 6.1 per cent significantly less likely to think that wife should have say in spending their own earning; 5 per cent and 2.2 per cent significantly less likely to recognize women's say in decisions like visiting her relatives/friends and number of children respectively, compared to husbands of other religion. For Hindu and Christian husbands, the results are insignificant (tables A2.5 and A2.7). The results from interaction term between different religion and having at least a son remain insignificant for all religion. However, by having 1^{st} born daughter, Muslim women may have significantly lower say in household decisions, such as 5 per cent lower say in large purchase, 4 per cent lower say in husband's money and 6 per cent lower autonomy to meet friends/relatives compared to women with 1^{st} born daughter in any other religion

(table 2.17).

2.5.6 Robustness Check

Table 2.18: Linear Probability Model: Using Couple Sample

Panel A. Women's responses about own say					
Dependent Variable:	Spending	Large	Husband's	Visiting	Child
Women's Say in:	Own Money	Purchase	Money	Relatives	Bearing
IV regression:					
At least a son	-0.0351	0.0246**	0.0395***	0.0421***	0.00128
	(0.0263)	(0.0124)	(0.0129)	(0.0122)	(0.0142)
Reduced form regression:					
1^{st} born a son	-0.0096	0.0081**	0.0130***	0.0138***	0.0003
	(0.0073)	(0.0041)	(0.0043)	(0.0040)	(0.0034)
Controls #	Yes	Yes	Yes	Yes	Yes
Observations	11085	43648	43319	43648	26706
Panel B. Husband's perception about women's say					
Dependent Variable:	Spending	Large	Wife's	Visiting	Child
Women's Say in:	Husband's Money	Purchase	Money	Relatives	Bearing
IV regression:					
At least a son	0.0156	-0.0010	0.0108	0.0028	0.0121
	(0.0127)	(0.0126)	(0.0308)	(0.0117)	(0.00872)
Reduced form regression:					
1^{st} born a son	0.0051	-0.0003	0.0031	0.0009	0.0040
	(0.0042)	(0.0042)	(0.0088)	(0.0038)	(0.0028)
Controls#	Yes	Yes	Yes	Yes	Yes
Observations	42932	43648	6839	43648	43648

: Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview. Standard errors clustered at PSU level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

To examine the robustness of the results, I also estimate the impact of having at least a son on women's say using linear probability models. Table 2.18 shows both IV linear probability model (LPM) and reduced form LPM estimations. These results remain robust in size and in statistical significance level as the results in tables 2.7 and 2.8, respectively.

Table 2.19 shows estimation results using third round of NFHS, which was conducted in 2005-06 and published in 2009 (NFHS-3, 2009). Panel A presents estimation results using women's individual sample. It is found that women with at least a son have significantly higher chances to have say in large purchases (3.2 per cent), to go visiting friends and relatives (2 per cent) and contraception use (12 per cent), compared to women without a son. In spending husband's money, the women with a son have chances to have a better say than women without son, but this effect is statistically insignificant.

Panel B in table 2.19 shows the estimated results by using couple data of NFHS-3. The impact of having at least a son on women's say has results in similar direction but the effects are lower in size and statistically insignificant. The results in Panel C use husbands' responses from couple data in NFHS-3. Having at least a son has positive but statistically insignificant impact on husbands' perception about women's intra-household say, as similarly as I found with NFHS-4 data.

Table 2.19: IV Probit Regression Marginal Effects: NFHS-3 (2005-2006)

Panel A. Using Women's responses from Women individual data 2005					
Dependent Variable:	Spending	Large	Husband's	Visiting	Child
Women's Say in:	Own Money	Purchase	Money	Relatives	Bearing
At least a son	-0.0292**	0.0316**	0.0021	0.0204*	0.1176***
	(0.0114)	(0.0127)	(0.0116)	(0.0117)	(0.0127)
Controls #	Yes	Yes	Yes	Yes	Yes
Observations	58896	58896	58896	58896	58896
Panel B. Using Women's Responses from Couple data 2005					
Dependent Variable:	Spending	Large	Husband's	Visiting	Child
Women's Say in:	Own Money	Purchase	Money	Relatives	Bearing
At least a son	-0.0227	0.0134	0.0064	0.0108	0.0059
	(0.0265)	(0.0186)	(0.0164)	(0.0169)	(0.0163)
Controls #	Yes	Yes	Yes	Yes	Yes
Observations	8555	27933	27791	27933	18363
Panel C. Husband's perception about women's say from Couple data 2005					
Dependent Variable:	Spending	Large	Wife's	Visiting	Child
Women's Say in:	Husband's Money	Purchase	Money	Relatives	Bearing
At least a son	0.0164	0.0089	0.0081	0.0187	0.0099
	(0.0171)	(0.0160)	(0.0126)	(0.0159)	(0.0105)
Controls#	Yes	Yes	Yes	Yes	Yes
Observations	25174	27933	27933	27933	27933

: Control variables used in regressions are: rural/urban; wealth index; women earning or not; female headed HH; difference of age between husband and wife; difference in education between husband and wife; ratio of female and male members in HH; religion; state fixed effects and interrupted by HH members during interview. Standard errors clustered at PSU level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.6 Conclusion

In societies with patriarchal values and custom of patrilocality, sons get more importance than daughters with expectation of higher utility gain from sons in future. Traditionally in Indian society, on the one hand, a boy is widely viewed as an asset to the family, his birth is celebrated and his mother is praised and rewarded by the family/society. On the other hand, a girl is considered as '*paraya dhan*' that is someone else's assets or considered

that she belongs to in-laws family, thus, birth of a girl is not praised considering girls as liabilities to the family, and her mother faces humiliation and neglect until she provides a son to the family. Based on this, it seems reasonable to expect that mothers of son(s) get higher status in the household compared to mothers without any son.

This paper examines whether having a son improves women's bargaining status or say within household compared to women without a son. The probit estimation results show that having at least a son significantly improves women's say in economic decisions in household, such as large purchases and spending husband's money. Having a son also improves women's autonomy, mother of a son can independently or jointly with partner decides on whether to make a visit to friends and relatives. Impact of having son(s) is found insignificant on women's say of 'spending own money'. It can be considered that individuals should have a higher say in spending own money than anyone else and thus, an earning woman should have higher say in own money irrespective of gender composition of her children. Therefore, say in spending own money should not be different due to having son or not, and thus, the insignificant son-effect on spending own money is consistent with the placebo check. Further, in different household matters, such as large purchases, spending husband's money and visiting relatives, women with an eldest son have significantly higher chances to have a say compared to women whose first child is daughter. This paper also uses husbands' subjective responses about women's say in a couple in general and found that having son(s) does not significantly change husbands' general views.

When including the detail categories of responses in terms of women having independent say; having joint say with partner; only husband having say or someone else having say, the ordered probability model estimations reveal that especially in large purchases, spending husband's money and visiting relatives and friends, women with at least a son have significantly higher probability to have an independent say by around 1 per cent and having a joint say with probability of 2 per cent compared to women with no son. Similarly, for such women with son(s), it is less likely by around 2.7 per cent that only their husbands make decisions on household matters, compared to women with no

son. Women with first born son also have significantly higher probability to have a say in household matters, both independent say and joint say with husbands compared to women whose first born is a girl.

Overall, the results in this paper imply that women with a son can realize higher value or status in family decision making than women without a son. However, having son(s) is not effective to change husbands' view to more egalitarian values in terms of individual's say in various decisions in a couple. The examination of the women who stopped child-bearing and thus are out-of reproductive cycle also reveals the chances of improvements in their say, compared to women without any son. In addition, from the estimated son-effect on domestic violence, I find that recent mothers if having son(s) have lower likelihood to face severe violence within household and as children gets older the son-effect disappears. Among other factors, the one that has the most important implications on women's say in household matters is her earning status, it is found that earning status has significant and largest positive influence on women's say in household decisions. Apart from this, household wealth and urban location also have higher significant influences on women's household status. But, religion, especially Muslim religion can influence women's say in negative manner compared to women in any other religions. Additionally, it is found that household characteristics, such as rural locatoion, female headed household, Muslim families etc can punish women significantly for not providing a son by ignoring their say in household decisions and reducing their autonomy in meeting relatives/friends.

Appendix 2

Table A2.1: Parity-wise (upto Parity 3) gender composition of children

Parities:	Composition	Per cent
Parity 1	Boy(B)	56.5
	Girl(G)	43.5
Parity 2	BB	30.5
	BG	27.0
	GB	27.5
	GG	14.9
Parity 3	BBB	11.9
	BBG	12.9
	BGB	15.6
	BGG	10.0
	GBB	15.7
	GBG	9.6
	GGB	17.1
	GGG	7.3

Table A2.2: Marginal Effects from Probit Estimations - Using Women Responses

	Dependent variable: Women's say in -				
Explanatory Variables	Spending Own Money	Large Purchase	Husband's Money	Visiting Relatives	Contraception Use
At least a son	-0.0020	0.0265***	0.0280***	0.0241***	0.0067
	(0.0102)	(0.0054)	(0.0057)	(0.0054)	(0.0053)
Rural HH	-0.0370***	-0.0249***	-0.0183***	-0.0206***	0.0010
	(0.0107)	(0.0061)	(0.0063)	(0.0062)	(0.0049)
Wealth	0.0265***	0.0312***	0.0248***	0.0306***	0.0147***
	(0.0052)	(0.0029)	(0.0030)	(0.0028)	(0.0024)
Female Headed	-0.0182	-0.0105	-0.0085	-0.0211**	-0.0143*
	(0.0169)	(0.0091)	(0.0095)	(0.0089)	(0.0078)
$Age_{husband} - Age_{wife}$	0.0011	-0.0005	-0.0006	-0.0003	-0.0003
	(0.0011)	(0.0006)	(0.0006)	(0.0006)	(0.0005)
$Educ_{husband} - Educ_{wife}$	-0.0043***	-0.0036***	-0.0034***	-0.0028***	-0.0013***
	(0.0009)	(0.0005)	(0.0005)	(0.0005)	(0.0004)
Sex ratio in HH	0.0071	0.0316***	0.0258***	0.0201***	0.0001
	(0.0108)	(0.0061)	(0.0063)	(0.0060)	(0.0052)
Hindu	-0.0225	-0.0063	-0.0077	-0.0158	0.0231**
	(0.0205)	(0.0123)	(0.0125)	(0.0123)	(0.0094)
Muslim	-0.0721***	-0.0542***	-0.0561***	-0.0619***	0.0106
	(0.0248)	(0.0139)	(0.0142)	(0.0139)	(0.0109)
Christian	-0.0294	0.0226	0.0222	0.0240	0.0253*
	(0.0271)	(0.0173)	(0.0177)	(0.0174)	(0.0143)
Earning women		0.0635***	0.0967***	0.0605***	0.0056
		(0.0052)	(0.0055)	(0.0052)	(0.0042)
Interrupted during interview	-0.0483***	-0.0421***	-0.0530***	-0.0415***	-0.0450***
	(0.0096)	(0.0059)	(0.0061)	(0.0058)	(0.0047)
Observations	11085	43648	43319	43648	26706

Note: Standard errors in parentheses. State fixed effects are used and standard errors are clustered at the PSU level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.3: Marginal Effects from Probit Estimations - Using Husbands Responses

	Dependent variable: Women's say in -				
Explanatory Variables	Spending Husband's Money	Large Purchase	Wife's Money	Visiting Relatives	Number of Children
At least a son	-0.0032	-0.0107*	0.0040	-0.0029	0.0005
	(0.0056)	(0.0057)	(0.0122)	(0.0052)	(0.0039)
Rural HH	-0.0047	-0.0196***	-0.0245*	-0.0080	0.0051
	(0.0077)	(0.0077)	(0.0129)	(0.0071)	(0.0052)
Wealth	0.0212***	0.0197***	0.0105*	0.0273***	0.0233***
	(0.0032)	(0.0032)	(0.0059)	(0.0029)	(0.0023)
Female Headed	0.0005	-0.0170*	-0.0033	-0.0152*	-0.0044
	(0.0096)	(0.0094)	(0.0201)	(0.0088)	(0.0066)
$Age_{husband} - Age_{wife}$	-0.0017***	0.0001	-0.0026**	-0.0004	0.0004
	(0.0006)	(0.0006)	(0.0012)	(0.0006)	(0.0004)
$Educ_{husband} - Educ_{wife}$	-0.0005	-0.0005	-0.0007	0.00005	-0.0002
	(0.0005)	(0.0005)	(0.0010)	(0.0005)	(0.0003)
Sex ratio in HH	-0.0015	0.0011	-0.0057	-0.0011	0.0031
	(0.0064)	(0.0062)	(0.0126)	(0.0057)	(0.0043)
Hindu	-0.0036	0.0102	-0.0296	-0.0147	-0.0007
	(0.0151)	(0.0144)	(0.0255)	(0.0140)	(0.0109)
Muslim	-0.0265	-0.0149	-0.0613**	-0.0498***	-0.0222*
	(0.0170)	(0.0164)	(0.0301)	(0.0160)	(0.0126)
Christian	0.0137	0.0382*	-0.0366	0.0111	0.0016
	(0.0199)	(0.0209)	(0.0348)	(0.0190)	(0.0142)
Earning women	0.0238***	0.0116**		0.0068	0.0038
	(0.0054)	(0.0052)		(0.0048)	(0.0036)
Interrupted during interview	-0.0050	-0.0115*	-0.0025	-0.0148***	-0.0178***
	(0.0061)	(0.0061)	(0.0117)	(0.0056)	(0.0040)
Observations	42932	43648	6833	43648	43648

Note: Standard errors in parentheses. State fixed effects are used and standard errors are clustered at the PSU level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.4: Marginal Effects from IV-Probit Estimations - Using Women Responses

	Dependent variable: Women's say in -				
Explanatory Variables	Spending Own Money	Large Purchase	Husband's Money	Visiting Relatives	Contraception Use
At least a son	-0.0324	0.0245**	0.0391***	0.0413***	0.0015
	(0.0265)	(0.0124)	(0.0129)	(0.0121)	(0.0141)
Rural HH	-0.0368***	-0.0248***	-0.0184***	-0.0207***	0.0010
	(0.0107)	(0.0061)	(0.0063)	(0.0062)	(0.0049)
Wealth	0.0261***	0.0312***	0.0249***	0.0307***	0.0146***
	(0.0052)	(0.0029)	(0.0030)	(0.0028)	(0.0024)
Female Headed	-0.0202	-0.0106	-0.0078	-0.0200**	-0.0146*
	(0.0170)	(0.0091)	(0.0095)	(0.0090)	(0.0079)
$Age_{husband} - Age_{wife}$	0.0010	-0.0005	-0.0006	-0.0003	-0.0003
	(0.0011)	(0.0006)	(0.0006)	(0.0006)	(0.0005)
$Educ_{husband} - Educ_{wife}$	-0.0041***	-0.0036***	-0.0035***	-0.0029***	-0.0013***
	(0.0009)	(0.0005)	(0.0005)	(0.0005)	(0.0004)
Sex ratio in HH	0.0011	0.0313***	0.0277***	0.0231***	-0.0006
	(0.0119)	(0.0064)	(0.0067)	(0.0063)	(0.0056)
Hindu	-0.0226	-0.0063	-0.0077	-0.0158	0.0231**
	(0.0204)	(0.0123)	(0.0125)	(0.0123)	(0.0094)
Muslim	-0.0714***	-0.0541***	-0.0563***	-0.0624***	0.0107
	(0.0248)	(0.0139)	(0.0142)	(0.0139)	(0.0109)
Christian	-0.0301	0.0226	0.0223	0.0241	0.0252*
	(0.0271)	(0.0173)	(0.0177)	(0.0174)	(0.0143)
Earning women		0.0635***	0.0964***	0.0599***	0.0057
		(0.0052)	(0.0055)	(0.0052)	(0.0042)
Interrupted during interview	-0.0481***	-0.0421***	-0.0530***	-0.0415***	-0.0449***
	(0.0096)	(0.0059)	(0.0061)	(0.0058)	(0.0047)
Observations	11085	43648	43319	43648	26706

Note: Standard errors in parentheses. State fixed effects are used and standard errors are clustered at the PSU level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.5: Marginal Effects from IV-Probit Estimations - Using Husbands Responses

	Dependent variable: Women's say in -				
Explanatory Variables	Spending Husband's Money	Large Purchase	Wife's Money	Visiting Relatives	Number of Children
At least a son	0.0157	-0.0014	0.0111	0.0023	0.0109
	(0.0126)	(0.0128)	(0.0308)	(0.0116)	(0.0085)
Rural HH	-0.0049	-0.0197***	-0.0245*	-0.0081	0.0051
	(0.0077)	(0.0077)	(0.0129)	(0.0071)	(0.0052)
Wealth	0.0214***	0.0198***	0.0106*	0.0273***	0.0234***
	(0.0032)	(0.0032)	(0.0059)	(0.0029)	(0.0023)
Female Headed	0.0017	-0.0165*	-0.0028	-0.0148*	-0.0038
	(0.0096)	(0.0094)	(0.0202)	(0.0088)	(0.0067)
$Age_{husband} - Age_{wife}$	-0.0017***	0.0001	-0.0026**	-0.0004	0.0004
	(0.0006)	(0.0006)	(0.0012)	(0.0006)	(0.0004)
$Educ_{husband} - Educ_{wife}$	-0.0006	-0.0005	-0.0008	0.00002	-0.0002
	(0.0005)	(0.0005)	(0.0011)	(0.0005)	(0.0003)
Sex ratio in HH	0.0018	0.0028	-0.0042	-0.0002	0.0049
	(0.0067)	(0.0066)	(0.0138)	(0.0060)	(0.0045)
Hindu	-0.0036	0.0102	-0.0295	-0.0146	-0.0006
	(0.0151)	(0.0144)	(0.0255)	(0.0140)	(0.0109)
Muslim	-0.0269	-0.0151	-0.0615**	-0.0500***	-0.0224*
	(0.0170)	(0.0164)	(0.0302)	(0.0160)	(0.0126)
Christian	0.0139	0.0383*	-0.0361	0.0112	0.0018
	(0.0199)	(0.0209)	(0.0348)	(0.0190)	(0.0142)
Earning women	0.0232***	0.0113**		0.0066	0.0035
	(0.0054)	(0.0052)		(0.0048)	(0.0036)
Interrupted during interview	-0.0050	-0.0115*	-0.0025	-0.0148***	-0.0178***
	(0.0061)	(0.0061)	(0.0117)	(0.0056)	(0.0040)
Observations	42932	43648	6833	43648	43648

Note: Standard errors in parentheses. State fixed effects are used and standard errors are clustered at the PSU level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.6: Marginal Effects from Reduced form Estimations: Women Responses

	Dependent variable: Women's say in -				
Explanatory Variables	Spending Own Money	Large Purchase	Husband's Money	Visiting Relatives	Contraception Use
1^{st} born a son	-0.0089	0.0080**	0.0129***	0.0135***	0.0004
	(0.0073)	(0.0041)	(0.0042)	(0.0040)	(0.0034)
Rural HH	-0.0372***	-0.0245***	-0.0179***	-0.0202***	0.0010
	(0.0107)	(0.0061)	(0.0063)	(0.0062)	(0.0049)
Wealth	0.0266***	0.0309***	0.0244***	0.0302***	0.0145***
	(0.0052)	(0.0029)	(0.0030)	(0.0028)	(0.0024)
Female Headed	-0.0184	-0.0118	-0.0098	-0.0222**	-0.0146*
	(0.0169)	(0.0090)	(0.0095)	(0.0089)	(0.0078)
$Age_{husband} - Age_{wife}$	0.0011	-0.0005	-0.0006	-0.0003	-0.0004
	(0.0011)	(0.0006)	(0.0006)	(0.0006)	(0.0005)
$Educ_{husband} - Educ_{wife}$	-0.0043***	-0.0034***	-0.0033***	-0.0027***	-0.0013***
	(0.0009)	(0.0005)	(0.0005)	(0.0005)	(0.0004)
Sex ratio in HH	0.0060	0.0280***	0.0225***	0.0175***	-0.0008
	(0.0107)	(0.0060)	(0.0063)	(0.0059)	(0.0052)
Hindu	-0.0228	-0.0062	-0.0075	-0.0156	0.0231**
	(0.0204)	(0.0123)	(0.0126)	(0.0123)	(0.0094)
Muslim	-0.0726***	-0.0535***	-0.0551***	-0.0612***	0.0108
	(0.0248)	(0.0139)	(0.0142)	(0.0139)	(0.0109)
Christian	-0.0298	0.0225	0.0224	0.0242	0.0253*
	(0.0271)	(0.0173)	(0.0177)	(0.0174)	(0.0143)
Earning women		0.0643***	0.0975***	0.0611***	0.0057
		(0.0052)	(0.0055)	(0.0052)	(0.0042)
Interrupted during interview	-0.0482***	-0.0421***	-0.0530***	-0.0414***	-0.0449***
	(0.0096)	(0.0059)	(0.0061)	(0.0058)	(0.0047)
Observations	11085	43648	43319	43648	26706

Note: Standard errors in parentheses. State fixed effects are used and standard errors are clustered at the PSU level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2.7: Marginal Effects from Reduced form Estimations: Husbands Responses

	Dependent variable: Women's say in -				
Explanatory Variables	Spending Husband's Money	Large Purchase	Wife's Money	Visiting Relatives	Number of Children
1^{st} born a son	0.0052	-0.0004	0.0031	0.0008	0.0036
	(0.0041)	(0.0042)	(0.0087)	(0.0038)	(0.0028)
Rural HH	-0.0047	-0.0197***	-0.0244*	-0.0081	0.0052
	(0.0077)	(0.0077)	(0.0129)	(0.0071)	(0.0052)
Wealth	0.0212***	0.0198***	0.0104*	0.0273***	0.0233***
	(0.0032)	(0.0032)	(0.0058)	(0.0029)	(0.0023)
Female Headed	0.0009	-0.0164*	-0.0036	-0.0149*	-0.0044
	(0.0096)	(0.0094)	(0.0201)	(0.0088)	(0.0066)
$Age_{husband} - Age_{wife}$	-0.0017***	0.0001	-0.0026**	-0.0003	0.0004
	(0.0006)	(0.0006)	(0.0012)	(0.0006)	(0.0004)
$Educ_{husband} - Educ_{wife}$	-0.0006	-0.0005	-0.0007	0.00003	-0.0002
	(0.0005)	(0.0005)	(0.0010)	(0.0005)	(0.0003)
Sex ratio in HH	-0.0003	0.0029	-0.0060	-0.0005	0.0035
	(0.0063)	(0.0062)	(0.0123)	(0.0057)	(0.0042)
Hindu	-0.0035	0.0102	-0.0295	-0.0147	-0.0006
	(0.0151)	(0.0144)	(0.0255)	(0.0140)	(0.0109)
Muslim	-0.0265	-0.0151	-0.0611**	-0.0499***	-0.0221*
	(0.0170)	(0.0164)	(0.0301)	(0.0160)	(0.0126)
Christian	0.0139	0.0383*	-0.0367	0.0112	0.0018
	(0.0199)	(0.0209)	(0.0348)	(0.0190)	(0.0142)
Earning women	0.0237***	0.0112**		0.0067	0.0038
	(0.0054)	(0.0052)		(0.0048)	(0.0036)
Interrupted during interview	-0.0050	-0.0115*	-0.0025	-0.0148***	-0.0178***
	(0.0061)	(0.0061)	(0.0117)	(0.0056)	(0.0040)
Observations	42932	43648	6833	43648	43648

Note: Standard errors in parentheses. State fixed effects are used and standard errors are clustered at the PSU level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

10.14754/CEU.2019.02

CHAPTER 3

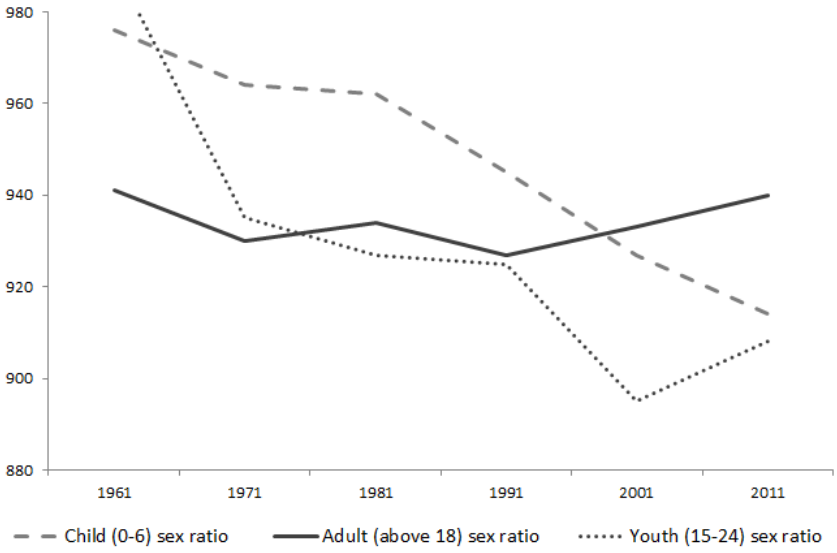
The Co-existence of Biased Sex Ratio and Crime against Women in India: Examining the Causality[1]

3.1 Motivation

Preference for sons has deep historical roots in many parts of the world, especially in Asia; stronger in patriarchal societies like India and China. Often, preference for sons can be evidenced in patrilocal societies, where this preference comes from the inter-related perceptions of higher economic, social and cultural utilities from sons. Indian parents consider sons as their old age care-giver and provider of secured shelter under the customs of patrilocality, as tools to carry forward family names and sons are culturally eligible to perform last rights of parents. Sons are preferred as it is more likely that they would contribute to increasing family wealth, whereas daughters are considered as economic burden to the family and take away family wealth in the form of dowry. Parents often feel that in spite of spending on her growing up, education, health and well being,

[1]I thank Gábor Kézdi and Andrea Weber for valuable advice.

Figure 3.1: Sex Ratios (1961-2011) in India



Source: Census of India.

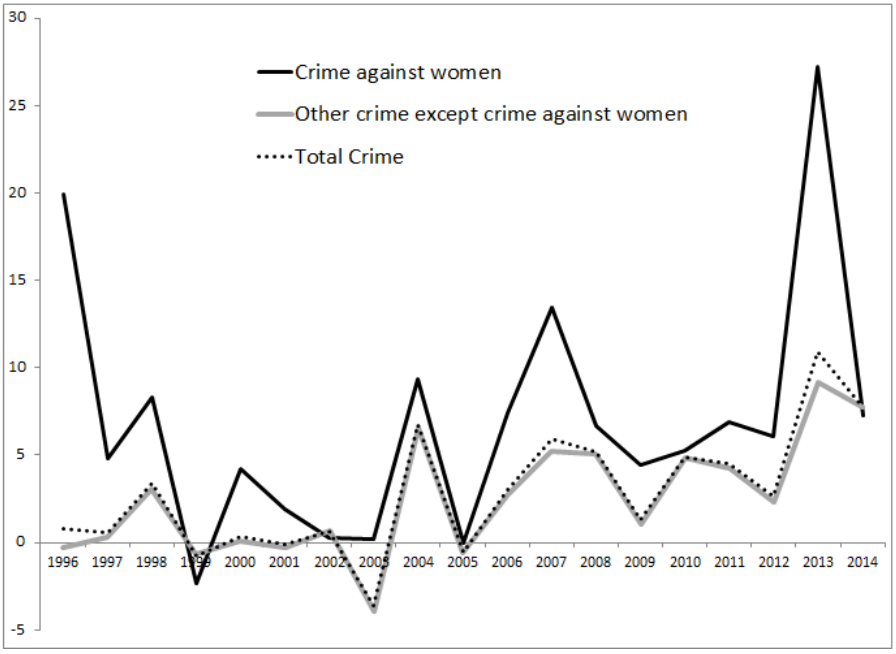
after marriage a daughter belongs to her husband's family and her labor value or income only benefits the in-laws family.

Therefore, Indian families historically prefer sons and do not hesitate to use different methods to fulfil their desire for sons. Sex selective family planning, abortions and infanticide of girl child may result in skewed child sex ratios in favor of boys, and eventually imbalance adult sex ratios as well.

With the prevalence of son preference and adoption of different techniques to get the desired composition of children, it is no wonder that highly skewed sex ratios prevail both in terms of child sex ratio and adult sex ratio in India. While adult sex ratio has shown a rising trend recently, the child sex ratio has dropped further and reached the lowest at 914[2] in the last census 2011, from 983 in 1951 (Census, 1951-2011). Youth sex ratio has also declined largely, from 990 in 1961 to the lowest at 895 in 2001 and recently increased to 908 (in 2011). Figure 3.1 shows the trends of child sex ratio, youth sex ratio and adult sex ratio during the last five decades in India from 1961 to 2001. India recently has been on news often for heinous rape incidences. It registers 27 per cent, the highest annual growth rate in crime against women in 2013 since last two decades, whereas the average

[2]In India, sex ratio is measured as number of females per 1000 males.

Figure 3.2: Growth in crime against women, other crimes and total crime (1996-2014)

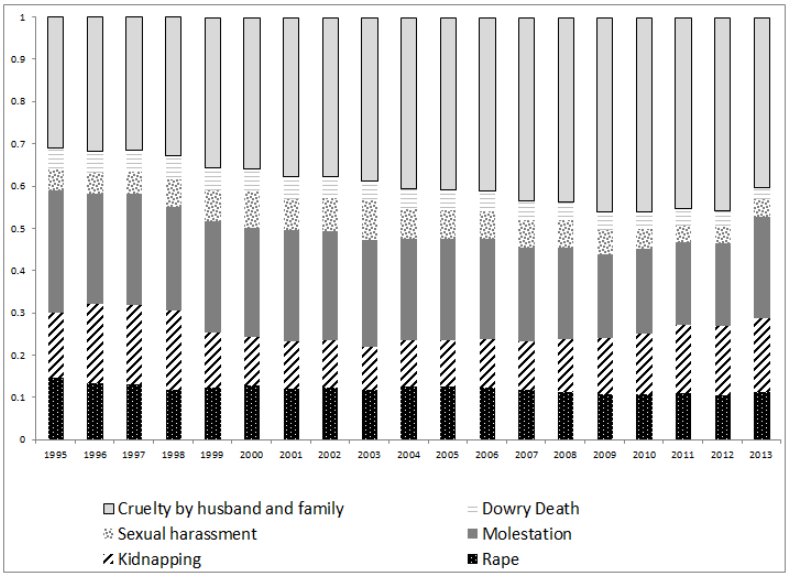


Source: Authors's calculation using NCRB data.

annual growth of total crime was 10.9 per cent in the same year (Bureau, 1995-2013). Figure 3.2 shows the annual change in different crime incidences under the category of crime against women, other crimes (except crime against women) and total crime over the years during 1995 to 2014. During this period, the growth rate of crime against women are greater than other crime and total crime in most of the years, except 1999 and 2002. The share of crime against women in total crime has increased from 5.4 per cent in 1996 to 11 per cent in 2013. The most recent crime data from NCRB (2016) reports on average 39 crimes against women in every hour in India, which was 21 incidences in an hour in 2007.

On the night of December 16, 2012, the gangrape of a paramedical student, Nirbhaya (named by media) in Delhi shocked the nation with the brutality of the incidence. Since then, particularly rape incidences have attracted a dramatic rise in media coverage. Regular news on rape incidences make us think that rape is growing disproportionately in India compared to any other crimes. However, one can also argue that it may not be the increase in incidences but increase in reporting. Rising media attention and increasingly concerned authorities have made it easier for rape victims to seek for justice today than earlier times. Yet, apart from rape, other crimes against women have not received much

Figure 3.3: Share of different crimes in Crime against women (1995-2013)



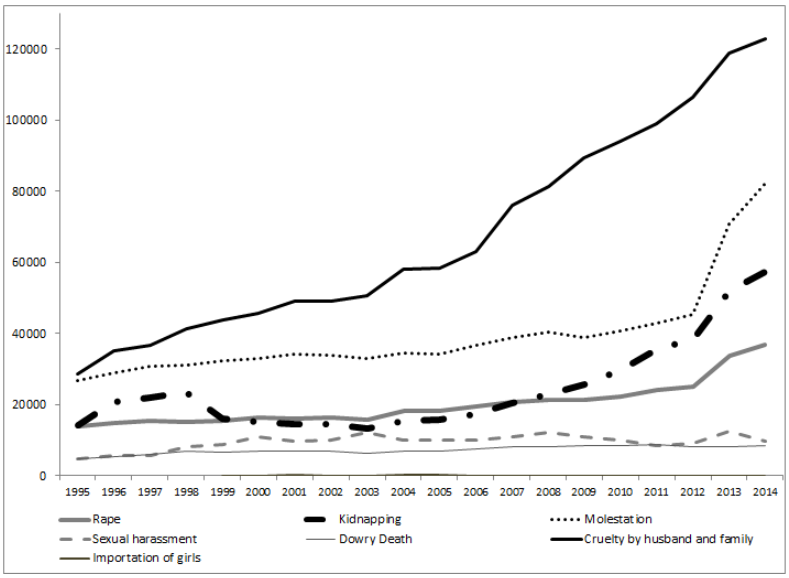
Source: Authors's calculation using Bureau (1995-2013) data.

attention so far. Then, it is questionable to assume that a large part of the total increase in crime against women is rather caused by increase in reporting than increase in crime incidences.

To explore the rising trend of crime against women (CRW) in more details, such as how different crimes have increased/decreased over time under the category of CRW and which crime mainly contributes to the overall increase in CRW, it is necessary to examine the crime-head-wise incidences under CRW. Figure 3.3 shows the crime-head wise shares of different crimes under the category of crime against women during the period of 1995 to 2013. It is the *cruelty by husband and his family members*, in general can be termed as domestic violence that registers the highest share in crime against women. The share of domestic violence in total CRW has increased from 30 per cent in 1995 to 46 per cent during 2009-2012 and recently declined to 40 per cent. *Molestation* has the second highest share in total CRW, around 20-25 per cent on average followed by *kidnapping* around 15-18 per cent share recently. *Rape* accounts for 11 per cent of incidences in total crime against women from 2007 to 2013, after decline from 14 per cent share in 1995.

Though the media spotlight towards rape may influence us to think that rape incidences would have the highest growth rate among different crime heads under CRW, but rape only accounts for 8 per cent average annual growth rate during the period of 1995-2013,

Figure 3.4: Crime-head-wise growth of different crime under CRW: 1995-2014



Source: Author's calculation using NCRB data.

whereas domestic violence by in-laws accounts for the highest average growth rate of 17 per cent followed by kidnapping (15 per cent), sexual harassment (9 per cent) and molestation (9 per cent).

Intuitively, domestic violence is the category that may have serious under-reporting issues, as often women do not want to report against the family members out of fear and/or compassion, and also difficult for the victims to report such crime when the accused and the victim live in the same household. Rather, for other types of CRW, the reporting would be easier than domestic violence. Figure 3.4 shows the trends of different crimes under the category of CRW and provides evidence in support that the increase in crime against women in last 19 years is primarily and consistently contributed by incidences of domestic violence with the steepest rise over the years, followed by molestation. Among others, the rape incidences have much flatter growth in comparison to domestic violence, but almost parallel growth rate as molestation. Therefore, it seems that rise in reporting should not be a matter of serious concern if majority of incidences against women are actually violence within household. In India, the coexistence of skewed sex ratio in favor of men and increasing crime against women creates a puzzle that when it seems logical by the economic theory that scarcity should increase the value of girls in society, surprisingly, instead they are increasingly treated with violence and harassment in their daily life either

by family members or by strangers. It can be called a puzzle as well from the perspective of Indian society, as due to lack of women cohort to marry in some states, they have to buy bride from other states and also sometimes multiple brothers get married to one girl to continue family lines, still we don't value the scarce women in our society, then it is puzzling for me. Daily news on crimes also make it evident that the regions that are more patriarchal in general also record higher crime incidences against women, such as Delhi, Chandigarh and some districts of Rajasthan, Haryana, Uttar Pradesh and Gujarat. Most of these states with imbalanced sex ratios are located at the northern and western parts of India and are considered to be more orthodox with their patriarchal mindset. Thus, it will be intuitive to examine whether this coincidence of male-biased sex ratios and higher crime against women bears any empirical relationship in case of India.

Therefore, the questions of interest in this paper are:

- Do biased sex ratios hold a relationship with the higher crime rates against women in India?
- Whether the wide-level preference for son in the past is responsible for the increase in crime rate against women today?

Studies have already shown that higher number of single men can increase crime in general, property crime and violence. Therefore, in this paper I will also examine:

- Do biased sex ratios impact crime against women and total crime differently?

Though there is no shortage of studies that examined preference for son and the effect of sex ratios on the overall crime rate, very few have examined its effect on crime against women and/or how the effects are different between crime against women and other crimes. Further, as per the best of my knowledge there is no attempt until recently to question the puzzle of coexistence of scarce women and increasing violence against women in India. Thus, this paper attempts to fill these gaps in the existing literature. I explain the contribution of this paper in more details at the end of the literature review.

The paper proceeds as follows: Section 3.2 provides a theoretical background to this research and intuitively explains the causal pathway that may have resulted in the coin-

cidence of biased sex ratio and increasing crime against women together. In section 3.3, relevant existing literature are mentioned followed by the discussion on data used for the analysis in section 3.4. Section 3.5 outlines the empirical framework of the research. In section 3.6, the results from the district-year panel analysis are shown and discussed. Finally, section 3.7 discusses about policy implications of the findings, and offers concluding remarks.

3.2 Theoretical Background

If potential parents do not adopt any tools based on their preference for children, gender of a new born child is a random phenomenon by nature. To achieve the preferred gender composition among children in the family, different methods are adopted at various stages in the process of child-bearing, even starting before conception to giving birth, and continue after birth as well. Family planning methods are adopted with the motive of having at least a son or two, and therefore child-bearing decision is not only influenced by preference for sons but also preferred composition of children and family size. Parents who already have a boy child may not desire/decide to have another child, whereas the parents who don't have at least a son, have wished for another child, and this desire may continue until they have a boy child.

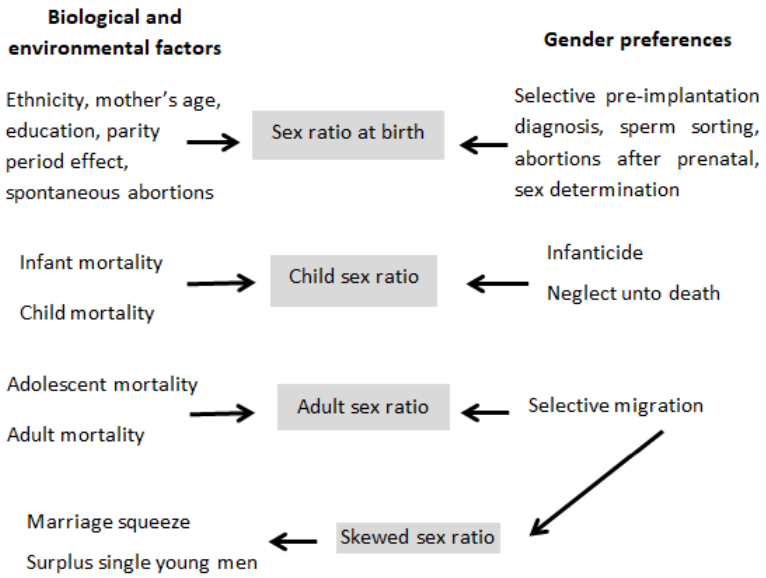
Since 1980s, with the advent of modern technology, the identification of sex of the fetus became easy for couples. This contributed to increasing number of sex selective abortions of female fetuses and resulted in an increasingly skewed sex ratio at birth in India. To control sex-selective abortions, Government of India amended the Pre-conception and Pre-natal Diagnostic Techniques (PCPNDT) Act[3] in 2002-03 that makes sex detection of fetus a punishable offence in the country.

High infant mortality rate among girls also indicates that a girl child gets less post-natal care and probably faces more discrimination and thus higher life risks compared to a boy

[3]The PCPNDT Act was enacted in 1994 and it was amended and effectively implemented in the year 2003.

child at early ages. The adoption of methods to achieve preferred composition and family size imbalances the sex ratio at birth, child sex ratio and eventually adult sex ratio in favor of males. Chart 3.1 shows the factors responsible for sex ratio and also the scopes of human interference in the process that may make the ratio of female to male imbalanced.

Chart 3.1: Factors responsible for sex ratios



Source: Author's compilation based on C. Guilmoto (2012).

In spite of the amendment of PCPNDT Act, 8 million female fetuses have been aborted between 2001 and 2011 in India, due to lack of enforcement measures (Pandey, 2011). Selective abortions of girls was about 4.2–12.1 million from 1980–2010, with a greater rate of increase in the 1990s than in the 2000s (Jha et al., 2011). The population control policies, such as two child norm provide further incentives towards having boys than a girl; if only two child then ideal family is at least one son if not two.

Preferences for boy child not only imbalance sex ratio among children but also cause dismal adult sex ratio in future. Skewed child sex ratio implies less number of matches between men and women, and more marriage squeeze in future (C. Z. Guilmoto, 2012). Therefore, historically existing preference for sons may have consequences of rising short-fall of supply of women as partner and increase the number of single surplus men. When it seems reasonable to think that shortages of girls in the marriage market will make the

girls dearer to Indian society (family), surprisingly, the evidence of increasing violence against women creates a puzzle. The puzzle that 'the scarcity of women in India does not increase their value' can be explained by culture, customs and the perceived value of a woman in the Indian family.

The correlation between skewed sex ratio and violence against women can be intuitively explained from two channels: (i) increasing frustration among surplus men due to lack of opportunities and (ii) by increasing vulnerability of women.

Surplus men and increasing violence

The Office of registrar general, Census of India stated that the number of surplus male of age 15-35 was 7 million in 1991 and reached 17 million in 2011. In total, India has 37 million excess men than women (Census, 2011). When society has a scarcity of women, there will be increased competition between men in the marriage market (Guttentag & Secord, 1983). Men become more competitive in acquiring higher education, better jobs and thus higher resources to make themselves more attractive in the marriage market. Studies on countries with low female to male ratio showed that the majority of the surplus men belong to the lower strata of the economy as higher competition in the marriage market favors the richer and higher educated males in high-paid jobs (Hudson & Boer, 2002). An analysis of the Chinese Census in 2000 shows that around 90 per cent of all unmarried men in China have less than senior high school education and that surplus men in China have lesser job opportunities, tend to be poor and can cause social unrest (J. D. Tucker et al., 2005).

Thus, men with lower resources will have higher risk to remain single, without any prospect of family making. These surplus men in Asia are different from single American men, who have options to be in a short-term relationship, but the Asian surplus men remain in lack of getting any partner or be in a relationship with a woman.

In a paper by S. Li, Zhang, Yang, and Attané (2010), the authors mentioned, "In the cultural context of China, singlehood is a state of frustration, and even of deprivation, for which it is difficult to find socially acceptable compensations: having children, living with a partner, having sexual relations, are aspects of life from which single men may be

excluded". The Indian surplus men are not different from Chinese surplus men, also suffer from similar deprivation and lack of opportunities, are often ridiculed by family and/or society and get less importance in property division. These generate (sexual) frustration among these men and can vent out in terms of violent behavior (against women).

Hesketh and Xing (2006) also argue that surplus men are predominantly of low socioeconomic class and their lack of access to resources; education and income earning opportunities may lead to antisocial behavior and violence, threatening societal stability and security. There are many other literature which show that young unmarried men are more crime prone than married men and women (Hudson and Boer (2002); Hudson and Den Boer (2004); (Edlund, Li, Yi, & Zhang, 2007)). With the increasing population of India, and her second position in the world population, the increasing number of single young men not only poses threat to India but to global society as a whole. Precisely, with no (less) chance to have a relationship with any woman, the sexual frustration of these single Indian men may cause insensitiveness towards women in the society and the patriarchal mindset has dominated their violent behavior towards women to exercise the power relations on the perceived weaker section of society.

Less women and increasing vulnerability

The other channel comes from the lower number of females in the population compared to males. The continuous shortage of women cohort to marry may cause widening of age-gap among couples[4], that is higher number of men will be matched with more younger women than earlier. Women, who marry at early ages, tend to remain less educated, less involved in labor market, begin having children earlier, and less bargaining power in the family (Jensen & Thornton, 2003). These young brides are more vulnerable and usually face higher abuse from the family members.

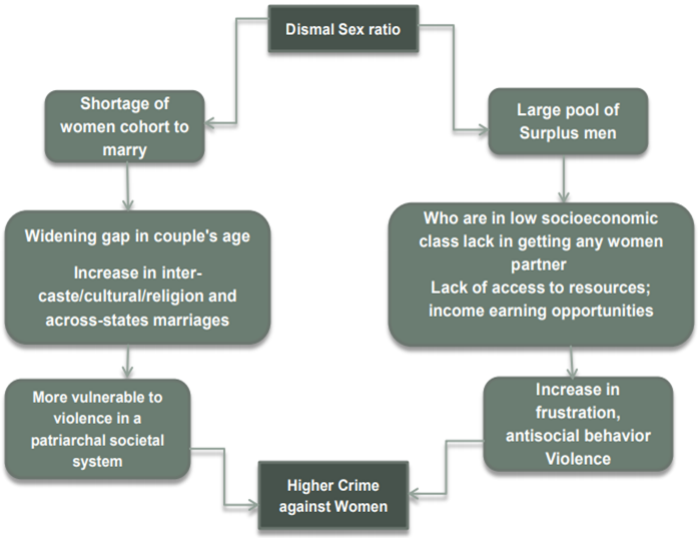
The patriarchal society of India emphasizes the need of marriage for sons to exercise the power relations of men being the primary bread-earner in the family; controlling the household resources, and producing sons to transfer the property according to the

[4]Couple's age gap= Husband's age - Wife's age

son-specific family lines. Therefore, in locations where female cohorts are much lower in number, the solution is to avail, even buy brides from poorer and/or lower-caste families and from distant places. A BBC Report "India's 'bride buying' country" documented few cases among thousands of girls from poor families of Jharkhand, Odisha, Assam and West Bengal are bought as bride in northern states like Haryana and Uttar Pradesh due to the skewed sex-ratios (Agal, 2006). The shortages of local brides therefore break the age-old customary norms, and increasingly accept the inter-caste, inter-religion, inter-state and inter-region marriages due to intense scarcity (Blanchet (2005); Davin (2005); Davin (2007); Kaur (2004); Kaur (2008); Kaur (2010); Kaur (2012); Kaur (2013); Ahlawat (2009); Chaudhry and Mohan (2011); D. Kumar (2012); Kukreja and Kumar (2013); Srinivasan (2017)).

The cross marriages bring the brides from either long-distance or different cultural set-up, therefore these girls need more efforts and time to adjust in the non-familiar culture, customs and location. These brides have less bargaining power within the marriage and are more likely to face higher discrimination and violence due to lack of support from husband's family members and from the neighborhood (Yang & Lu, 2010). Based on the discussion so far, I build a schematic diagram (Chart 3.2) that shows two possible channels to explain the hypothesis, 'more single young men in society may increase crime against women' and thus explains the puzzle, intuitively.

Chart 3.2: Two pathways explain the puzzle



The scarcity of women increases the number of early age, inter-caste and inter-religion marriages and raises the risk of domestic violence. The widely followed custom of Indian society that 'wife should be less educated and earning less than her husband' also establishes the male-dominance over wife. This explains that on the one hand, single men belong to the lowest strata in terms of resources, education and job and women at the apex with highest resources, education and job experience difficulty to find a match. On the other hand, less access to education and thus less income generating opportunities among married women decrease their economic value in the family; and patriarchal society with preference for son measures the social value of a woman in terms of son(s) provided by her to the family.

3.3 Literature Review

Studies have investigated causes behind dismal sex ratios (C. Z. Guilmoto and Attané (2007); Pande and Astone (2007); Angrist (2002); Klasen and Wink (2003); Singariya (2012); Chakraborty and Sinha (2006); Sen (1992); Sen (1990); C. Tucker and Van Hook (2013); Premi (2001)), especially pro-male biased sex ratios, and its direct consequences such as marriage squeeze (C. Z. Guilmoto (2012); Park and Cho (1995); Guttentag and Secord (1983); South and Trent (1988); Jiang, Feldman, and Li (2014)) and generating surplus men, who are named as *bare branches* (Hudson and Boer (2002); Hudson and Den Boer (2004); J. D. Tucker et al. (2005)) by many literature. The demographic factors that have been cited in the literature for having effect on sex ratio at birth are sibling composition, family size and birth order, parent's age and occupation, social and economic status of parents and most importantly preference for children of some particular gender and for some particular composition of children (Park and Cho (1995); Ding and Hesketh (2006); Jha et al. (2006)). Birth control policies may also have adverse impacts on national sex ratio. Ding and Hesketh (2006) showed that China's one child policy reduced total birth rate and family size, and intensified the desire for son and thus imbalanced the sex ratio of China further. A paper by Jha et al. (2006) shows

that India's preference for son and thus sex selective abortion is primarily responsible for lower sex ratio at birth nationally. Women who already have one or more daughters are at risk of going through selective abortions willingly or due to family pressure. The authors also highlighted, "Based on conservative assumptions, the practice accounts for about 0.5 million missing female births yearly, translating over the past 2 decades into the abortion of some 10 million female fetuses" (Jha et al., 2006).

The decreasing ratio of female to male generates increasing number of single marginalised men with little family prospects and almost no outlet for sexual energy, and these men react with frustration and violence towards society, in terms of increasing antisocial and violent behavior. Multiple literature have supported the hypothesis that more men in the population can increase violence (Hudson and Boer (2002) and also among others, Hudson and Den Boer (2004); Messner and Sampson (1991); Hesketh, Lu, and Xing (2011); Edlund et al. (2007); Steffensmeier and Allan (1996)). Further, the findings from studies also indicate that an increasing percentage of violent crime is committed by young, unmarried, low-status males (Messner and Sampson (1991); Oldenburg (1992)).

Using Indian data for the period of 1980-1982, Dreze and Khera (2000) found a strong correlation between murder rates and sex ratio, after controlling for urbanization, poverty and literacy rate. They found that the districts with higher female to male ratios have lower murder rates and emphasized that patriarchal societies are likely to be more violent. Hudson and Boer (2002) re-estimated the same relationship and found similar correlation between sex ratio and homicide rates in the late 1990s.

Using Interpol data for 70 countries, Barber (2000) has shown evidence for the alternative hypothesis that the countries with lower number of males compared to females may have more incidences of violent crimes, rape and assaults. He explained that societies with fewer males are likely to have more family conflicts, and therefore higher aggression within families and thus higher violent crimes. However, another paper by him, Barber (2003) found that the rate of violent crimes against persons increased with male to female ratio (15-44 years age-group) in Britain and Scotland and stated that it happended 'because men are much more likely to commit all types of violent crimes than women are'.

Edlund et al. (2007) found positive correlation between surplus men and crime level in China. The paper found 0.01 increase in sex ratio in favor of men raises violence and property crime rate by 6 per cent and that the increasing number of unmarried young male accounts for as much as one third of overall increase in crime.

Oldenburg (1992) has argued that in areas with high level of violence, preference for sons is higher as sons are valued as protection against violence and exercise of power (in reference to Northern states of India). He also found that murder rates are high in low sex ratio (F:M) districts of Uttar Pradesh. Messner and Sampson (1991) examined race-specific data on robbery and homicide rates for a sample of 153 American cities and finds positive effects of (M:F) sex ratio on violent crime only when controlling for family disruption. Another paper by Edlund, Li, Yi, and Zhang (2013) found that the elasticity of crime with respect to the youth sex ratio (age-group of 16 to 25 years) is 3.4, and also that sex ratios can account for one-seventh of the rise in crime.

Bose, Trent, and South (2013) highlighted that "...men will exercise extraordinarily strict control over women's behaviour when women's relationship options are plentiful and men's own options are limited" and also found that in Indian communities where sex ratio is biased in favor of male, women has higher likelihood to face distrust and intimate partner violence, after controlling for individual, household, and geographic characteristics. Other papers, such as Zhang (2010) and Hesketh et al. (2011) explored the relation between higher rape and sexual harassment incidences and scarcity of women in China and provided evidence in favor of the relationship.

Though there are no shortage of literature which show relation between sex ratios with overall crime rate, property crime and other violent crime, but works on the relationship between sex ratios and crime against women in particular are scarce. In case of India, the examination of the relationship and measuring the impact of biased sex ratio on crime against women become relevant with growing safety and security concerns for women in Indian society. Prakash and Vadlamannati (2014) examined the association between skewed sex ratio and shortage of girls with their illegal trafficking in India and found that 100 unit increase in child sex ratio is associated with a 0.635 per cent increase in illegal

trafficking of girls.

To fill-in the shortages in the existing literature, this paper attempts to explore the relationship between skewed sex ratios in favor of male and crime against women, particularly, in more details. After the submission of the first draft of this paper in 2015, another very recent paper by Amaral and Bhalotra (2017) shows that higher number of surplus males at the age-group of 20-24 years increases crime against women and finds that imbalanced sex ratio explains around 21 per cent increase in crime against women. My paper uses a broader definition of young males that is at the age-group of 15-24 years and examines the impact of surplus young males on crime against women in order to explain the puzzle in coexistence of scarcity of women and increasing violence on them. It gives an intuitive explanation of channels that has linked pro-male biased sex ratios with higher crime against women and also provides empirical evidence for this driving mechanism. In addition, it also explores whether the prevalence of preference for sons in the past that imbalanced the child sex ratio would have any impact on crime against women today. This presents a dynamic causal relationship that how past preference for sons can create a violent society for daughters in the years ahead. It also examines the difference between the impacts of biased sex ratios on crime against women and other crime. Further, the impact of sex ratios on women safety across regions and different crime-heads are also estimated to provide policy directions in areas of priority.

3.4 Data

In this paper, I examine the relation between the imbalanced sex ratios and the crime situation, especially crime against women in 395 districts of 35 states and union territories (UTs) of India during the period of 1995 to 2013. A district is an administrative division of a state or union territory. As of 2011 Census of India there are 640 districts and the previous two censuses - Census 1991 and Census 2001 record 466 and 593 districts respectively. Between 1971 and 2001, the number of districts increased from 356 to 593, a rise of about 67 per cent. This large increase in number of districts is caused by

continuous changes in formation of administrative boundaries of districts, many has been divided into multiple new districts, aggregation of two or small districts into a new one; and changes in geographical boundaries. Also, many districts have changed their name once or multiple times. A paper by H. Kumar and Somanathan (2009) finds that only 136 out of 356 districts in 1971 (38 per cent) were unaffected by boundary changes until 2001 census. Due to these multilevel changes in administrative boundaries and names, it becomes difficult to obtain consistent data on districts for long periods. In the presence of this issue, I could obtain reliable data on only 395 districts out of 640 districts reported in 2011. Among these 395 districts some districts have changed names over the years, I tried to map data from the old name to the new one (eg. Khandwa and Khargone districts were formerly known as the East and West Nimar districts respectively). In cases where multiple districts were formed by partitioning one single districts, data for the multiple districts are clubbed and kept as one mother district (eg. Imphal is kept as one district instead of Imphal East and West). In some cases the districts were divided under the same name but only divided as rural and urban locations, in such cases also these were kept as one district instead of two or more districts (eg. Trichy Rural, Trichy Urban and Trichy Railways are clubbed together as Trichy).

The dependent variable for this analysis is crime against women (CRW). The crime data are taken from National Crime Records Bureau of India (NCRB). NCRB collects crime data under different crime-heads as defined by the Indian Penal Code (IPC) and publishes this data at national, state and district levels every year. The crime data represents the total incidences reported formally as First Information Reports (FIR) at the police stations. For this analysis, the reported incidences of crime against women in 395 districts during 1995 to 2013 are considered. Crime against women primarily includes the incidences reported under different crime-heads, such as rape, kidnapping, molestation, sexual harassment, dowry deaths, cruelty by husband and husband's family members and importation of girls. The total number of incidences under CRW is the summation of incidences reported under these seven categories of crime committed against women. Total crime includes all types of crimes defined under IPC. And the other crime

includes all crimes except the crime incidences against women.

The main explanatory variable for this analysis is sex ratios, which is defined in the Census as the number of female per 1000 males. The Census data on population across different age groups are considered to calculate different sex ratios required for the analysis. The Census of India publishes the population data in every 10 years. For this analysis data from five Censuses (1971, 1981, 1991, 2001 and 2011) are used. The data for other variables, such as urban and literate population are taken from district level Censuses.

3.5 Empirical Strategy

The analysis considers district-year panel of 395 districts for 19 years period from 1995 to 2013. The relationship between youth sex ratio and crime against women can be examined using simple linear equation as:

$$CRW_{it} = \alpha_i + \theta_t + \beta * YSR_{it} + \gamma * X_{it} + \epsilon_{it} \tag{3.1}$$

The dependent variable is the log of crime incidences against women CRW_{it} in district i in year t. Due to large variation in crime incidences across districts, CRW is constructed by taking log of crime incidences which would help in data smoothening. However, taking log becomes problematic if the data has zero or negative values. In this data few districts have no crime incidences under a particular category in a year, and taking log will show them as missing values and thus will drop these observations from the analysis. To resolve this issue the value of crime incidences is scaled up by adding 10 units for all, before taking log. I believe that this linear altercation would not make any changes in the estimated results.

In equation (3.1), the main explanatory variable is youth sex ratio (YSR_{it}) in district i in year t. YSR is the ratio of female to male, more precisely number of females per 1000

males in the age-group and 15-24 years [5]. I consider this age-group to construct YSR as literature have shown that the single young males at this age-group are more prone to be involved in criminal activities (Graham and Bowling (1995); Edlund et al. (2007), among others). In India, the age-wise data on arrested people for committing different crime reveals that on average around 45 to 60 per cent people arrested on suspicion are below 30 years of age (Figure A3.1). To construct YSR, the number of males and females in age-group of 15-24 years are taken from census 1991, 2001 and 2011. The middle-year population such as from 1995-2000, 2002-2010 are linearly projected using growth rate in population between two consecutive census years and the data for 2012 and 2013 is linearly extrapolated using the same growth rate from census 2001 and 2011. For example, data from 2002 to 2010 is projected from the previous year population multiplied with the annual growth rate calculated from Censuses 2001 and 2011[6].

I employ a panel estimation method including both district fixed effects α_i and year fixed effects θ_t that control for district-specific and year-specific unobserved heterogeneity respectively. Precisely, the district fixed effects are meant to capture the local factors that may affect the district specific characteristics and can have an impact on the local crime rates whereas the year fixed effects are used to control for nationwide common shocks to crime rates in a year, such as election year, bad harvest year etc.

The control variables used in the estimations are urbanization rate, literacy rate and per capita income. These variables may have impacts on crime and/or on the unobserved factors of crime. Before including all the control variables together, I examine the presence of multicollinearity by computing the correlations between the variables and using Vari-

[5] $$YSR = \frac{No\ of\ Females}{No\ of\ Males} X1000 \ \ in\ the\ age-group\ of\ 15-24\ years.$$

[6] Annual growth rate is calculated by dividing the decadal growth with 10. And decadal growth rate (DG) let say between 2001 to 2011,

$$DG_{2001-2011} = [P_{2011} - P_{2001}]/P_{2001}$$

ance Inflation Factor (VIF). The correlation between per capita income and literacy rate is found 0.66 and testing with VIF (also tolerance indicator) confirms multicollinearity. It indicates that literacy rate and per capita income should not be used together in the same estimation. By estimating the equation with literacy rate and per capita income one at a time and comparing the R^2s, I decide to keep literacy rate instead of per capita income. The estimation also uses female population as a control variable to normalize crime against women with female population.

The estimation of equation (3.1) using Ordinary Least Squares (OLS) method will be an issue as YSR has potential endogeneity and can produce biased estimation results. Therefore I will use instrumental variables for YSR and estimate the equation using Two Stage Least square (TSLS) method. A potential instrument for YSR could be Child Sex Ratio with 15 years of lag (CSR_{it-15}). The first stage equation in the TSLS estimation is:

$$YSR_{it} = A_i + C_t + B * CSR_{it-15} + D * X_{it} + \eta_{it} \tag{3.2}$$

In the first stage equation, the dependent variable YSR, the main explanatory variable, which is the instrument, CSR_{it-15} and also other control variables (same as equation (3.1)) are used. The CSR_{it-15} is constructed from the census data considering the children at the age-group of 0-4 years in districts i with 15 years lag. For this, Census data for years 1971, 1981 and 1991 are considered and the lagged CSR for middle years are computed by linear projection and extrapolation for the period of 1980 to 1998.

The validity of the instrument requires two conditions to satisfy: (i) the instrument must be exogenous that is $Cov(CSR_{-15}, \epsilon)$ should be equal to zero, and (ii) the instrument must be correlated with the endogenous explanatory variable, that is $Cov(CSR_{-15}, YSR)$ shouldn't be zero.

The first condition can't be proved as ϵ is unobserved. In this case, we can assume that it is unlikely that CSR with 15 years of lag that is CSR in (t-15) years is correlated with unobservable factors of crime against women in year (t). It seems plausible to think that the current crime rate couldn't pollute the determinants of the juvenile sex ratio before 15 years. However it may happen that districts with prevalence of higher crime rates may

have male-biased child sex ratios. Parents in such districts may have higher preference for boy child as in one hand, sons are considered as tools of protection against violence and can provide higher safety to the family when grow up, and on the other hand, they may also fear that a girl child is not safe in such crime-prone localities. These perceptions due to higher contemporaneous crime rate in the districts may result in biased child sex ratio in favor of male. Therefore, to take into account of the externalities due to contemporaneous crime rates on child sex ratio, rape incidences with 15 years lag are included as control variable in the estimation. In addition, it may happen that districts where patriarchal mindset prevails from the past and has been aggravated further, and thus CRW in such districts are historically higher than CRW in others. I employ district specific linear time trend to control the time invariant unobserved factors in the districts that can influence CRW. Further, in all estimations standard errors are clustered at district levels, as the districts of a state can have some common characteristics that influence the crime of the locations. Also, in the estimations districts and year fixed effects are used to control for district and time specific characteristics. After employing the above treatments, it seems unlikely that people decided to have more boys compared to girls considering future crime directly, rather it seems possible that an area with higher patriarchal practices in the past may commit more crime today because of surplus young men it produced in the society.

The second condition can be proved by the non-zero correlation coefficient of CSR in the first stage equation. Child sex ratio is a primary determinant of future youth sex ratio and eventually adult sex ratio, therefore these sex ratios are supposed to be correlated. The correlation coefficient between YSR and CSR_{-15} is found 0.376.

The second stage equation is (3.1), where the explanatory endogenous variable YSR takes the estimated values from the first stage equation (3.2).

Another objective of the paper is to examine whether there is any difference between the effects of sex ratios on crime against women and other crimes. The effects of sex ratios on crimes can be compared between the estimated effects on crime against women, other crimes (that is excluding crime against women from total crime) and total crimes using the same estimation strategy. For this purpose, two more equations are estimated, such

as:

$$CRO_{it} = \alpha_i + \theta_t + \beta * YSR_{it} + \gamma * X_{it} + \epsilon_{it} \tag{3.3}$$

$$CRT_{it} = \alpha_i + \theta_t + \beta * YSR_{it} + \gamma * X_{it} + \epsilon_{it} \tag{3.4}$$

Both equations (3.3) and (3.4) are the second stage equations for other crime (CRO) and total crime (CRT) as dependent variables respectively. The main explanatory variable YSR, instrument CSR_{-15} and control variables (Xs) are same in these estimations, except I use total crime with 15 years lag instead of lagged (-15) rape incidences to take into account the contemporaneous effect of crime on CSR. Also, the first stage equation remain same as equation (3.2).

Other control variables used in both first and second stage estimations are urbanization rate, literacy rate and log of female population in the districts. Type of location, that is urban or rural can have some impact on the local crime rates. It can be expected that districts with higher percentage of educated (literate) people possibly face lower crime.

3.6 Results

3.6.1 Descriptive Statistics:

Table 3.1 shows the summary statistics of the variables used in the analysis. Table A3.1 in appendix shows the panel level descriptive statistics and indicates that the variabilities of all the variables used in the analysis are higher at district level compared to year levels. Average crime incidence against women in districts during the period of 1995-2013 is 329 and the highest crime incidences against women is as large as 12853, reported in Delhi in the year 2013. The second highest incidences of crime against women is 7363 and is reported in 24 Pargana district of West Bengal in 2011. Excluding these two districts from total, the mean crime incidences during 1995-2013 is 312 and the maximum reported incidence is 4614. The average total crime incidences in districts is 4014 in a year and the maximum incidence reported is as large as 80184 (again in Delhi in 2013).

Table 3.1: Descriptive Statistics:

Variables	Obs	Mean	Std. Dev.	Min	Max
District	7505	198	114.034	1	395
Year	7505	2004	5.478	1995	2013
Crime against women	7505	329.498	434.794	0	12853
Rape with 15 years lag	7358	22.300	25.742	0	544
Total crime	7439	4014.132	5026.613	0	80184
Total crime with 15 years lag	7336	3531.395	4087.959	10	64882
YSR	7505	910.561	91.081	215.148	1243.183
CSR wth 15 years lag	7505	958.300	41.941	628.735	1139.17
Urbanisation rate	7505	26.316	19.467	0	100
Literacy rate	7505	58.062	12.221	6.235	93.628
ln(Female Population)	7505	13.517	1.044	9.572	16.0314

The old data only takes into account the reported rape incidences under crime against women. For other crime against women the data doesn't provide gender dis-aggregated numbers, but includes overall crime incidences under a crime-head, such as provides total kidnapping data instead of kidnapping of females, males and children separately. Therefore for historical data I only consider rape incidences with 15 years lag that is data from 1980 to 1998. The total crime with 15 years lag includes data from 1980 to 1998 and on average it is 3531 incidences and highest incidences (64882) is reported in Delhi in 1998. The highest rape incidences (544) is reported in Delhi in 1997 and the second highest rape incidences (191) is reported in Moradabad district of Uttar Pradesh in the year 1987 (during the period of 1980-1998). Table 3.2 shows the top districts with highest reported crime against women and total crime.

Table 3.2: Highest crime incidences in districts during 1995-2013

	Highest Crime against Women			Highest total Crime		
Rank	District	Year	Incidences	District	Year	Incidences
1	Delhi	2013	12853	Delhi	2013	80184
2	24 Pargana	2011	7363	Delhi	1998	64882
3	24 Pargana	2010	6614	Delhi	1997	60883
4	24 Pargana	2009	6034	Delhi	1996	59871
5	24 Pargana	2008	6014	Delhi	1999	58701
6	Delhi	2012	5920	Delhi	2006	57963
7	24 Pargana	2013	5266	Delhi	2000	56249
8	Delhi	2011	5193	Delhi	2005	56065
9	24 Pargana	2012	5047	Delhi	2007	56065
10	Delhi	2007	4725	Delhi	2001	54384

- Top 10 districts with highest level of crime against women: Delhi, 24 Pargana (West Bengal), Murshidabad (West Bengal), Kolkata (West Bengal), Mumbai (Maharash-

tra), Burdwan (West Bengal),Nadia (West Bengal), Ahmedabad (Gujarat), Jaipur (Rajasthan) and Jalpaiguri (West Bengal).

- Top 10 districts with highest level of total crime: Delhi, Ernakulam (Kerala), Mumbai (Maharashtra), Bangalore (Karnataka), 24 Pargana (West Bengal), Jaipur (Rajasthan), Surat (Gujarat), Kolkata (West Bengal), Ahmedabad (Gujarat) and Pune (Maharashtra).

The NCRB data on crime against women in India reveals that Delhi has the highest incidence of crime against women, the second position is held by 24 Pargana (North and South aggregated data) of West Bengal during the period of 1995 to 2013. Not only that, five other districts of the state of West Bengal ranks among the top ten districts with very high incidences of crime against women. Maintaining consistent positions among the top 10 highest crime locations over the years, West Bengal also has the highest incidences of domestic violence (cruelty by husband or in-laws), around one-fifth of such cases in the country happens in this state. West Bengal presents a stark irony as this state is considered as a place with more intellectual and progressive mindset people in one hand, and with higher violence against women within household on other hand. The state also registers higher number of female trafficking cases. A newspaper editorial by Singh (2017) highlighted the issue of high crime incidences in the state and mentioned, "As the State shares a border with Bangladesh, Nepal and Bhutan, it has become a transit route in human trafficking. The distress-ridden tea gardens of north Bengal, the remote islands of Sunderbans and the districts of Malda and Murshidabad with poor human development indicators and high density of population serve as ideal source point for traffickers luring young girls on the pretext of jobs or marriage to other States... But despite the high levels of crimes against women, it has failed to garner adequate attention from the authorities". Singh also highlighted that the overall conviction rate for crime against women in India stands at 18.9 per cent, whereas in West Bengal, the conviction rate is the lowest in the country at 3.3 per cent.

The average YSR in districts during 1995-2013 is 910 females out of 1000 males and the lowest is 215 females in 1000 males in Daman and maximum is 1243 females out of 1000

males in Almora district of Uttarakhand. The CSR data with 15 years lag that is from 1980 to 1998, has average of 958 girls out of 1000 boys and the minimum and maximum is 628 and 1139 females respectively out of 1000 males in Darbhanga district of Bihar and lower Subansiri district of Arunachal Pradesh respectively.

In both census of 2001 and 2011, the two lowest YSRs (15-24 years age group) are registered in the two union territories; Daman and Dadra & Nagar Haveli. A report in a prime newspaper of India in 2011 has explained that the very biased sex ratio in Daman and Dadra & Nagar Haveli in favor of males are caused due to higher migration in the areas and most of the migrants are male laborers which has increased the population but disbalanced the sex ratio. Population has increased by 53 per cent and 55 per cent in Daman and Dadra & Nagar Haveli, respectively from 2001 to 2011. Apart from that, preference for boy child and rise in single child families have emphasized the biasedness in the population of these two union territories (Bhatt, 2011). According to 2011 Census, the bottom five states with lowest youth sex ratio are Daman, Dadra & Nagar Haveli, Chandigarh, Delhi and Haryana.

The primary factors that contribute to biased child sex ratio are preference for son, discriminatory behavior to the girl child, and practices like sex selective abortions, female foeticide and infanticide. Districts with very low child sex ratios are situated in Punjab, Haryana, Madhya Pradesh, Gujarat, Uttar Pradesh, Rajasthan and Maharashtra.

Table 3.3 shows the range of changes in the districts to give a brief idea on how the population has evolved and sex ratios have changed during the period of analysis. The CSR (0-4 years age) has declined in most of the districts, for 326 districts out of 395 districts (82.5 per cent). The highest decline in YSR from 2001 to 2011 is observed in Leh (Jammu and Kashmir) district. Till 2011, techniques for sex-detection of the fetus were neither well-known nor readily available in the Kashmir region. The not-very biased past CSR of Leh during 1981 to 1991 censuses also rules out the sex selective abortion as a factor to contribute to the biasedness in sex ratio of Leh. A reason behind sharp decline in YSR can be increased army postings in the area followed by the 1999 Kargil war. A Census official said, “The total population of the two districts of Ladakh (Leh

Table 3.3: Change in sex ratio over decades

Change in CSR (0-4 years of age)			Change in YSR (15-24 years of age)		
Change b/n 1981-1991	No of Districts	Cumulative Per cent	Change b/n 2001-2011	No of Districts	Cumulative Per cent
Decline:			Decline:		
More than 150 points	1	0.25	More than 200 points	1	0.25
			150 - 200 points	1	0.51
100 - 150 points	2	0.76	100 - 150 points	3	1.27
50 - 100 points	49	13.16	50 - 100 points	28	8.35
1 - 50 points	274	82.53	1 - 50 points	101	33.92
No Change	0	82.53	No change	0	33.92
Increase:			Increase:		
1 - 50 points	63	98.48	1 - 50 points	198	84.05
50 - 100 points	4	99.49	50 -100 points	61	99.49
100 - 150 points	2	100.00	100 - 150 points	2	100.00

and Kargil) is 2.9 lakh and they include several thousand troops. Even an increase of a few thousand (soldiers) can significantly alter the sex ratio" (Raina, 2011).

As per the last five census data on India, the CSR of India has continuously declining during 1971 - 2011, that is during this period the number of girl child in the age group of 0 to 4 years has diverting continuously from the number of boys in the same age-group. Studies have extensively discussed the causes of this divergence, such "as a kinship pattern (Agnihotri (1997); Dyson and Moore (1983); Miller (1981b); Miller (1981a)) or neglect of the girl child (Arokiasamy (2004); Bardhan (1974); Gupta (1987); Mayer (1999); Visaria (1969)) are mainly due to the social value attached to women because of their economic utility" (Saha & Paul, 2017).

3.6.2 Results: The relationship between the sex ratios and crime against women

To examine the relationship between sex ratio and crime against women, I use OLS estimation method. The dependent variable is the log of (reported) crime incidences against women (CRW) and the main explanatory variable is YSR in 395 districts of India during years 1995 to 2013. The OLS estimation results are given in table 3.4. In the first result column the estimation does not use any control variables and the second column presents results with control variables. The control variables used in this analysis are urbanization rate, literacy rate and log of female population. To capture the time

Table 3.4: OLS results: Impact of youth sex ratio on crime against women

Dependent Variable: ln(Crime against Women)	Without Controls	With Controls	District Trend and w Controls
YSR	-0.00153***	-0.00165***	-0.00166***
	(0.00042)	(0.00045)	(0.00044)
Urbanization Rate		0.00106	0.00102
		(0.00193)	(0.00190)
Literacy Rate		-0.00298	-0.00289
		(0.00237)	(0.00235)
ln(Female Population)		0.474***	0.453***
		(0.154)	(0.153)
Constant	7.342***	0.933	0.521
	(0.413)	(2.127)	(2.122)
N	7505	7505	7505
R-sqr	0.917	0.919	0.910
Year FE:	Yes	Yes	Yes
District FE:	Yes	Yes	No
District Trend:	No	No	Yes

Standard errors clustered at district levels are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

specific and location specific factors, the year fixed and district fixed effects are used in the analysis in the first and second columns' estimations. There can be cases that CRW in some districts are higher than CRW in others and also districts may have some time trend which can have effects on the crimes of the location. It would be difficult to examine the time series data of 395 districts and also because nature and trend of crime rate against women can be similar in some of the districts. Accounting for that, I also performed estimation including district specific linear time trends and results are shown in the third column of table 3.4. The standard errors are clustered at district levels for all the estimations. The estimated coefficients between CRW and YSR from the three models are similar in magnitude and significance (statistically significant at 1 per cent level). The coefficients in table 3.4 can be interpreted as increase in one female per 1000 males that is one unit increase in youth sex ratio in favor of female will significantly reduce crime against women by around 0.16 per cent, keeping all other control variables constant. From the coefficients of other control variables, it can be said that urban places may have higher crime against women compared to rural locations and improvement in literacy rate can reduce crime against women in a location, but the coefficients for these control variables remain insignificant.

However, the OLS results in table 3.4 are suspected to provide biased estimator due to potential endogeneity in the main explanatory variable YSR. Therefore, YSR is instrumented by CSR with 15 years of lag and TSLS method is used for estimations. YSR data used in the analysis is from 1995 to 2013 and its instrument, CSR with 15 years lag (CSR_{-15}) is considered from 1980 to 1998. The TSLS results are presented in table 3.5. In TSLS method, the first stage regression considers YSR as dependent variable and (CSR_{-15}) as the independent variable. From the first stage results, the estimated correlation coefficient between YSR and (CSR_{-15}) is 0.376. The estimated correlation coefficients between YSR and (CSR_{-15}) using different estimation strategies, different control variables and fixed effects are shown in the main tables of estimations, and these coefficients of (CSR_{-15}) are positive and highly significant at 1 per cent level in all cases. In the second stage the dependent variable is CRW and the main explanatory variable

Table 3.5: Impact of youth sex ratio on crime against women

Dependent Variable: ln(Crime against Women)	Without Controls	With Controls	District Trend and w Controls
TSLS results:			
YSR	-0.00492**	-0.00535**	-0.00536**
	(0.00226)	(0.00219)	(0.00220)
Urbanization Rate		-0.000107	-0.000187
		(0.00221)	(0.00219)
Literacy Rate		-0.00548*	-0.00537*
		(0.00327)	(0.00324)
ln(Female Population)		0.520***	0.501***
		(0.173)	(0.172)
Rape with 15 years lag		-0.00111	-0.00115
		(0.00119)	(0.00120)
Constant	10.67***	4.056	2.554
	(2.219)	(3.645)	(2.730)
N	7358	7358	7358
R-sqr	0.912	0.913	0.913
Year FE:	Yes	Yes	Yes
District FE:	Yes	Yes	No
District Trend:	No	No	Yes
First Stage Regressions - Dependent Variable: YSR			
CSR with 15 years lag	0.1723***	0.1750***	0.1726***
	(0.0350)	(0.0133)	(0.0134)

Standard errors clustered at district levels are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

is the estimated YSR from the first stage. The control variables and fixed effects remain same in both the first and second stage regressions.

Table 3.5 shows TSLS results from three different estimation strategies, similarly as table 3.4. The first result column of coefficients refer to the TSLS estimates by using only YSR instrumented with (CSR_{-15}) and both time and district fixed effects but without any control variables; the second result column shows regression after adding control variables. Here, apart from the control variables used in table 3.4, rape with 15 years of lag is included as another covariate. It may happen that districts with historically higher (lower) crime rates against women are more (less) son biased and thus have more (less) skewed sex ratios in favor of male. To take into consideration of this possibility, the estimation includes the reported rape incidences in the districts with 15 years lags ($Rape_{-15}$) as a control variable. The ($Rape_{-15}$) data therefore includes information for districts from 1980 to 1998. Due to unavailability of old rape data for some districts in some years, 147 observations are dropped from estimations to equate the number of observations for comparison of the results in three models. I also add the district specific linear time trend in the analysis to take into account of the heterogeneity at the district levels. The third column gives results adding district trend, control variables and with year fixed effect but without the district fixed effects.

The results in table 3.5 indicate that YSR has a significant and negative relation with CRW in all three estimations. The coefficients can be interpreted as increase in one female per 1000 males that is 1 unit increase in youth sex ratio in favor of female will reduce crime rate against women by around 0.49 per cent when estimated with year and district fixed effects but without any control variables. After adding control variables in the estimation, I find that 1 unit increase in YSR in favor of female can decrease CRW by 0.53 per cent. Both these coefficients are statistically significant at the 5 per cent level. Adding district trend the results remain robust in size and significance. Among the control variables, urbanization rate, literacy rate and past rape incidences have negative relations with CRW, whereas the log of female population has a positive relation with CRW.

Table 3.6 shows the reduced form results. As similar as the earlier estimations shown in table 3.5, three strategies are used, except here, the main explanatory variable is the

instrument itself, (CSR_{-15}) instead of YSR. The first, second and third result columns show the estimations without controls, with control variables, and with district trend (without district fixed effects) respectively. The (CSR_{-15}) coefficients are negative and statistically significant in all cases. The results can be interpreted as increase in one girl child in 1000 boys in the past that is 1 unit increase in child sex ratio in favor of girls 15 years before could significantly reduce current crime rate against women by around 0.093 per cent, keeping all other things same.

Table 3.6: Impact of child sex ratio on crime against women

Dependent Variable: ln(Crime against Women)	Without Controls	With Controls	District Trend and w Controls
Reduced form results:			
CSR with 15 years lag	-0.000847**	-0.000936**	-0.000931**
	(0.000399)	(0.000375)	(0.000376)
Urbanization Rate		0.00166	0.00164
		(0.00204)	(0.00201)
Literacy Rate		-0.00264	-0.00256
		(0.00246)	(0.00244)
ln(Female Population)		0.500***	0.476***
		(0.145)	(0.144)
Rape with 15 years lag		-0.00188*	-0.00193*
		(0.00107)	(0.00108)
Constant	6.705***	-0.101	-101.5***
	(0.408)	(2.131)	(7.303)
N	7358	7358	7358
R-sqr	0.915	0.917	0.917
Year FE:	Yes	Yes	Yes
District FE:	Yes	Yes	No
District Trend:	No	No	Yes

Standard errors clustered at district levels are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In the descriptive statistics section, I have discussed that Delhi and 24 Paragana (West Bengal) have historically very high incidences of CRW compared to other districts, and these two districts are outliers in the sample of 395 districts. Therefore, I drop these two districts and re-estimate the earlier two tables 3.5 and 3.6 (detail results are given in tables A3.2 and A3.3 in Appendix). Even after dropping these two outliers, the results remain robust in significance, but the size of the impact increases by a small amount. An increase in YSR by 1 unit will reduce CRW by 0.54 per cent, *ceteris paribus*. The reduced form regression also remains robust, an increase in the past CSR by 1 unit causes reduction in current CRW by 0.095 per cent on average, keeping other things same.

3.6.3 Comparison of relationships of crime against women, total crime and other crime with sex ratio

Table 3.7: Relationship of sex ratios with different crime categories

Dependent Variable:	Crime against women (CRW)	Other crime except CRW	Total crime reported
Panel A:			
TSLS regression: YSR	-0.00514**	-0.00451**	-0.00414**
	(0.00229)	(0.00194)	(0.00185)
Reduced form: CSR_{-15}	-0.000869**	-0.000762**	-0.000700**
	(0.000398)	(0.000299)	(0.000291)
N	7232	7232	7232
Controls	No	No	No
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
District Trend	No	No	No
Panel B:			
TSLS regression: YSR	-0.00565**	-0.00497**	-0.00458**
	(0.00221)	(0.00205)	(0.00194)
Reduced form: CSR_{-15}	-0.000965***	-0.000837***	-0.000772***
	(0.000372)	(0.000301)	(0.000290)
N	7232	7232	7232
Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
District Trend	No	No	No
Panel C:			
TSLS regression: YSR	-0.00511*	-0.00545**	-0.00511**
	(0.00295)	(0.00234)	(0.00222)
Reduced form: CSR_{-15}	-0.000959**	-0.000819***	-0.000756***
	(0.000374)	(0.000299)	(0.000288)
N	7232	7232	7232
Controls	Yes	Yes	Yes
District FE	No	No	No
Year FE	Yes	Yes	Yes
District Trend	Yes	Yes	Yes

SEs clustered at district levels are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Another objective of this chapter is to compare the relationships of sex ratios with CRW, other crime (CRO) and total crime (CRT) in order to examine how differently sex ratios can impact different types of crime. Other crime represents the overall crime except the CRW. Table 3.7 shows the comparison of the results between CRW, CRO and CRT, using different estimation strategies as earlier. In these estimations, 147 and 126 observations were dropped due to missing values in $Rape_{-15}$ and total $crime_{-15}$ respectively.

Without any control variables but with year and district fixed effects, the relationship of YSR with CRW is higher in size compared to the same with CRO and CRT, and statistically significant at 5 per cent level in all three cases of crimes (Panel A). After adding control variables in the estimation, the impact of YSR on CRW is also higher than the impact of YSR on CRO and CRT. In case of estimations with dependent variables CRO and CRT, I use log of total reported crime with 15 years lag instead of $Rape_{-15}$. The

results in the Panel B with control variables can be interpreted as increase in 1 female in 1000 males that is 1 unit increase in YSR causes 0.53 per cent reduction in CRW, 0.50 per cent reduction in CRO, and 0.46 per cent reduction in CRT on average, *ceteris paribus*. In Panel C, after adding district specific linear time trend in the estimations, I find that the impacts of YSR on CRW and CRT become same in magnitude and impact on CRO becomes larger. More specifically, 1 unit increase in YSR, on average reduces CRW and CRT by 0.51 per cent and CRO by 0.54 per cent, keeping other variables constant.

The reduced form regression results with control variables and fixed effects indicate that 1 unit increase in past CSR can cause reduction in current crime rates, such as CRW by 0.096 per cent, CRO by 0.084 per cent and CRT by 0.077 per cent. Therefore, the impact of past CSR on current crime against women is the highest and the impact is found lowest on total crime. Adding district trend, the impact of 1 unit increase in past CSR can reduce CRW, CRO and CRT by 0.096 per cent, 0.082 per cent and 0.076 per cent respectively. From all these estimation results, it can be said that impact of biased sex ratios on crime against women is higher compared to the impact of the same on other crimes apart from CRW.

3.6.4 YSR Quintile-wise analysis of the relationship between sex ratios and Crime against women

In this section, I divide the data as per the YSR quintiles in order to examine the impact of sex ratios across their intensity of biasedness on crime against women. The quintiles are calculated over the youth sex ratio values of districts over time, and a district can be in different quintiles in different year. The first quintile represents the most biased sex ratios, ranges from 215 to 835 female per 1000 males. The second quintile ranges from 836 to 891 females, third quintile is from 891 to 938 females; fourth quintile ranges from 938 to 989 females; and fifth quintile is from 989 to 1243 females per 1000 males.

Table 3.8 shows the TSLS estimations in quintiles including control variables and district and year fixed effects. All standard errors are clustered at the district level. The results indicate that in districts with the most biased YSR (YSR is in the lowest quintile), the

Table 3.8: Impact of youth sex ratio on crime against women across quintiles

Dependent Variable: ln(Crime against Women)	YSR Q1	YSR Q2	YSR Q3	YSR Q4	YSR Q5
YSR	-0.00437**	-0.0108	0.00592	0.0180	0.00376
	(0.00207)	(0.00826)	(0.0103)	(0.0240)	(0.00742)
Urbanization Rate	-0.00422**	0.00432	-0.00556	0.00127	0.0104***
	(0.00207)	(0.00442)	(0.00580)	(0.00542)	(0.00197)
Literacy Rate	-0.00372	-0.00211	0.00530	-0.00651	-0.00318
	(0.00362)	(0.00383)	(0.00823)	(0.00748)	(0.0124)
ln(Female Population)	0.488***	0.594***	0.669***	0.874***	0.779**
	(0.120)	(0.155)	(0.134)	(0.221)	(0.396)
Rape with 15 years lag	0.000747	-0.00403***	-0.00535***	-0.00627***	-0.00284***
	(0.000480)	(0.000908)	(0.00117)	(0.00228)	(0.00104)
Constant	2.224	6.299	-9.619	-23.88	-9.802***
	(2.442)	(5.903)	(10.75)	(21.40)	(2.890)
N	1482	1467	1473	1481	1455
First Stage Regressions - Dependent Variable: YSR					
CSR with 15 years lag	0.164***	0.0894***	0.0755***	0.0538	0.128***
	(0.0191)	(0.0232)	(0.0223)	(0.0334)	(0.0430)

Standard errors clustered at district levels are presented in parentheses.
Both Districts and Year FEs are used. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
YSR quintile ranges are as follows - Q1: 215.15 - 835.56; Q2: 835.65 - 890.61; Q3: 890.63 - 938.43; Q4: 938.45 - 989.34; and Q5: 989.38 - 1243.18.

impact of YSR on crime rate is highly significant but lower in size compared to districts where YSR is moderately biased (around 835 to 890 females compared to 1000 males in the second lowest quintile). In the lowest YSR districts, an increase in YSR by 1 unit causes a significant reduction in CRW by 0.4 per cent. The size of the impact of YSR on CRW in the second lowest quintile increases, but becomes statistically insignificant. In the districts, where YSR lies in the second lowest quintile, 1 unit increase in YSR can reduce crime against women by 1 per cent (insignificant).

In districts, where YSR lies in the third to fifth quintiles, that is where YSR is above 890 females per 1000 males, the impact of biased YSR on crime becomes positive and remains statistically insignificant. Therefore, I can infer that in districts with very biased YSR (lower than 890 females per 1000 males), increase in women can reduce crime against women. As the availability of women increases as partners, number of single men will reduce and that can reduce CRW. In other words, in already very biased YSR districts, people will value women if number of women increases, and therefore violence towards them may reduce. But, starting from the third quintile (above 890 females per 1000

Table 3.9: Impact of child sex ratio on crime against women across quintiles

Dependent Variable: ln(Crime against Women)	YSR Q1	YSR Q2	YSR Q3	YSR Q4	YSR Q5
Reduced form results:					
CSR with 15 years lag	-0.000714**	-0.000962	0.000447	0.000966	0.000483
	(0.000328)	(0.000711)	(0.000755)	(0.00114)	(0.000924)
Urbanization Rate	-0.00122	0.00650	-0.00380	-0.00230	0.00983***
	(0.00139)	(0.00395)	(0.00431)	(0.00261)	(0.00171)
Literacy Rate	0.000485	-0.00347	0.000824	-0.0115***	0.00256
	(0.00265)	(0.00359)	(0.00279)	(0.00296)	(0.00405)
ln(Female Population)	0.490***	0.451***	0.626***	0.971***	0.963***
	(0.116)	(0.107)	(0.112)	(0.131)	(0.133)
Rape with 15 years lag	0.000282	-0.00429***	-0.00489***	-0.00468***	-0.00256***
	(0.000416)	(0.000841)	(0.000827)	(0.000801)	(0.000921)
Constant	-0.792	-0.129	-3.906**	-8.501***	-9.078***
	(1.630)	(1.669)	(1.651)	(1.842)	(1.969)
N	1482	1467	1473	1481	1455

Standard errors clustered at district levels are presented in parentheses.
Both Districts and Year FEs are used. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
YSR quintile ranges are as follows - Q1: 215.15 - 835.56; Q2: 835.65 - 890.61; Q3: 890.63 - 938.43; Q4: 938.45 - 989.34; and Q5: 989.38 - 1243.18.

males) the impact of YSR on crime becomes positive, that is increase in number of female may increase the crime against women.

Therefore, the overall inference from the quintile-wise analysis is that in places where females are very low in numbers compared to males, crime against women can be reduced by reducing the number of single men and/or increasing the number of women. However, in districts where YSR is more than 890 females in 1000 males, increase in females may not have any significant impact on the crime against women.

In table 3.9, the reduced form results using quintiles of YSR indicate that in lowest quintile of YSR, past CSR has significant but lower reduction impact on CRW, compared to the districts at the second lowest quintiles. In the second lowest quintile the relation between past CSR and CRW remains negative, but larger than lowest quintile and becomes insignificant. Starting from third quintiles onward, the relation between past CSR and CRW becomes positive and insignificant. The quintile-wise analysis also performed adding the district-wise time trend and year fixed effects (see table A3.4 in appendix), and the results remain robust as tables 3.8 and 3.9.

3.6.5 The region wise analysis of the relationship between sex ratios and crime against women

The states of India can be categorized into six regions in terms of geographical location in the country. The regions are as follows:

- Northern Region,
- Southern Region,
- Central Region,
- Western Region,
- Eastern Region and
- North Eastern Region.

Table 3.10: Impact of sex ratios on crime against women across regions

Dependent Variable:	Explanatory Variable	Northern	Southern	Western	Eastern	Central	North Eastern
IV with controls							
CRW	YSR	-0.00278	-0.0159	-0.0133***	0.00198	0.0151	0.0155
		(0.00517)	(0.0281)	(0.00484)	(0.00421)	(0.0259)	(0.0463)
N		2493	1350	903	1042	810	760
CRO	YSR	-0.0210	-0.00716	-0.00324	0.00285	0.0221	0.0197
		(0.0130)	(0.0145)	(0.00379)	(0.00316)	(0.0152)	(0.0295)
N		2477	1330	885	1027	795	735
CRT	YSR	-0.0187	-0.00814	-0.00392	0.00331	0.0208	0.0173
		(0.0115)	(0.0153)	(0.00365)	(0.00337)	(0.0146)	(0.0263)
N		2477	1330	885	1027	795	735
Reduced Form							
CRW	CSR_{-15}	-0.000464	-0.00236	-0.00479***	0.000251	0.000588	-0.00117
		(0.000862)	(0.00217)	(0.00169)	(0.000518)	(0.00135)	(0.00145)
N		2493	1350	903	1042	810	760
CRO	CSR_{-15}	-0.00358***	-0.000984	-0.00108	0.000343	0.00133	-0.00240**
		(0.000862)	(0.00156)	(0.00127)	(0.000341)	(0.000903)	(0.000945)
N		2477	1330	885	1027	795	735
CRT	CSR_{-15}	-0.00318***	-0.00112	-0.00130	0.000397	0.00125	-0.00210**
		(0.000812)	(0.00154)	(0.00122)	(0.000362)	(0.000905)	(0.000980)
N		2477	1330	885	1027	795	735

Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3.5 in appendix shows the detail of states in the regions and the number of districts included in the analysis from each of the states. The results of region-wise analysis in table 3.10 reveal that in the Western region biased sex ratio in favor of males may impact crime against women substantially and significantly. Improvement in YSR

by one unit can cause 1.3 per cent reduction in crime against women on average, but such improvement in sex ratios would not bring any significant change in other and total crime in western part of the country. In the Northern and Southern regions of India, the biased sex ratio has negative but insignificant impact on CRW, with larger impact in Southern part compared to Northern. The past CSR with 15 years lag has a negative and highly significant impact on both CRO and CRT by around 3 to 3.5 per cent in the Northern region, keeping other things same. It means that in the Northern region patriarchal values and prevalence of son preference from the past are significantly detrimental to society, by increasing other crimes or crimes in general. However, the past prevalence of son preference does not significantly cause increase in today's crime against women in Northern India. In other regions, such as Eastern and Central India the biased sex ratios in favor of male has positive but insignificant relation with violence against women, other and total crimes. In the North Eastern region, the past CSR with 15 years lag has significant and negative relation with both CRO and CRT, keeping other things same.

The comparison of impacts of sex ratios on crime between six regions indicates that biased sex ratio has most significant effect on women's life in Western part than in any other regions of India, and improvement in sex ratios in favor of female may significantly reduce crime against women in the Western region. In the Northern states, biased sex ratios have the most significant relation with other crime and total crime rather than crime against women, and policies towards improving sex ratios in favor of female may significantly reduce violence rate, in general, in this region. Both Western and Northern regions include states and districts with more pro-male biased sex ratios and also crime rates are higher in the districts of these two regions. Using district-wise trend, the estimation results across regions are shown in table A3.6 in appendix, and these results remain robust as the results in table 3.10.

3.6.6 Crime-head wise analysis:

This section examines the impact of imbalances in sex ratios on different types of crime under the category of crime against women. Table A3.7 presents the descriptive statistics

of different crime heads under crime against women. Among 395 districts during the period of 1995-2013, average annual rape incidences is 37, kidnapping is 43, molestation is 78 and sexual harassment incidences on average is 21. Among these crimes, annual average of domestic violence incidences is 141, it is as high as four times of average rape incidences, twice of molestation and seven times of sexual harassment incidences per year. While Delhi has the highest rape and kidnapping incidences, 24 Parganas has the highest domestic violence occurrences. Delhi has been always in the position among the top three districts in all crime-heads of CRW, followed by 24 Parganas.

Table 3.11 shows the relationship of sex ratios with different crime-heads under CRW. Domestic violence has the highest share in CRW and its growth rate is much higher than other crime incidences against women (see Figure 3.4). The estimated results show that YSR has the highest impact on domestic violence compared to other crime-heads. Improvement in YSR by 1 unit in favor of female can significantly reduce the rate of domestic violence by 0.8 per cent on average, *ceteris paribus*. And, increase of 1 female per 1000 young males may cause significant reduction in kidnapping of girls (women) by 0.74 per cent on average, keeping other things same. Not only that, even if the past child sex ratio has increased by 1 unit in favor of girls, it would cause significant decline in both domestic violence and kidnapping by 0.12 per cent. For other crime-heads under crime against women the estimated coefficients are not statistically significant. Skewed sex ratio may have a negative impact on dowry deaths, more specifically, one unit improvement in YSR due to increase in female can cause 0.2 per cent decrease in deaths due to unmet dowry demands, keeping other things same. The past CSR may also have a negative impact on dowry deaths, but such effects are small in size and statistically insignificant. The biased sex ratios, both YSR and CSR_{-15} do not have negative and statistically significant impact on rape incidences. The sexual harassment incidences may also reduce due to improvement in sex ratio, however molestation may not show a decline due to improvement in sex ratios in favor of females. Careful examination of data reveal that there are some outliers in the incidences of crime. For example, the maximum rape incidences are registered as 1636, whereas only 4 and only 12 districts have reported rape

Table 3.11: Crime-head wise analysis: Impact of sex ratios

Dependent Variable:	TSLS:			Reduced Form:		
Crime-head:	Main explanatory variable -			Main explanatory variable -		
	YSR	YSR	YSR	CSR_{-15}	CSR_{-15}	CSR_{-15}
Rape	0.000362	0.000139	0.000214	0.0000588	0.0000240	0.0000365
	(0.00199)	(0.00188)	(0.00188)	(0.000326)	(0.000324)	(0.000322)
N	7409	7279	7279	7409	7279	7279
Kidnapping	-0.00881***	-0.00736***	-0.00738***	-0.00142***	-0.00125***	-0.00124**
	(0.00298)	(0.00271)	(0.00276)	(0.000491)	(0.000480)	(0.000485)
N	7339	7208	7208	7339	7208	7208
Domestic Violence	-0.00845**	-0.00785**	-0.00778**	-0.00130**	-0.00128**	-0.00125**
	(0.00388)	(0.00354)	(0.00357)	(0.000512)	(0.000525)	(0.000527)
N	7282	7151	7151	7282	7151	7151
Molestation	0.000524	0.000115	0.000107	0.0000780	0.0000183	0.0000164
	(0.00388)	(0.00354)	(0.00357)	(0.000512)	(0.000525)	(0.000527)
N	7303	7172	7172	7303	7172	7172
Sexual Harassment	-0.000118	-0.000797	-0.000879	-0.0000166	-0.000121	-0.000132
	(0.00322)	(0.00305)	(0.00309)	(0.000455)	(0.000464)	(0.000464)
N	7042	6913	6913	7042	6913	6913
Dowry Death	-0.00187	-0.00207	-0.00205	-0.000288	-0.000338	-0.000331
	(0.00230)	(0.00204)	(0.00207)	(0.000374)	(0.000350)	(0.000351)
N	7197	7067	7067	7197	7067	7067
Controls	No	Yes	Yes	No	Yes	Yes
District FE	Yes	Yes	No	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District Trend	No	No	Yes	No	No	Yes

Standard errors clustered at distrct levels are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

more than 600 and 500 respectively in a year. Similarly domestic violence has extreme value at 5640, whereas only 15, 81 and 332 observations have reported incidences of domestic violence in a year above 2000, 1000 and 500 respectively. It may happen that such extreme values can skew the results upto some extent. To examine the robustness of the crime-head wise estimations, I re-estimate the results after dropping these extreme (outliers) observations from the analysis (see table A3.8 in appendix). The results remain robust, though the size of the impact increases in some cases. For domestic violence, exclusion of 81 observations that is observations with values above 1000, increases the impact of YSR: an increase in YSR by 1 unit reduces domestic violence by 0.9 per cent on average (not shown in the table). And after exclusion of 332 observations that have reported domestic violence cases above 500, an improvement in YSR by one unit may significantly reduce violence against women within household by 1 per cent, *ceteris paribus*.

Overall, the crime-head analysis provides evidence towards significant and higher crime-

reducing impacts of sex ratios on particularly domestic violence and kidnapping of girls: that is of around 0.8 per cent reduction in these two crime-heads due to one unit improvement in YSR in favor of female. But the impact of pro-male biased sex ratios on the rest of the crime-heads, especially sexual crimes, such as rape, molestation and sexual harassment remains lower in magnitude and statistically insignificant. Therefore, this results provide significant evidence towards the channel of 'scarcity of women and increasing vulnerability' and not statistically significant evidence for the channel of 'surplus single men and increasing violence due to sexual frustration'. As the results support the channel of 'scarcity of women and increasing vulnerability' then it can be expected that acute shortage of women as partners in the districts with very skewed sex ratios in favor of male may lead to more non-traditional marriages, that is increasing acceptance towards inter-caste, inter-religious and also inter-state marriages. In such locations it can also be expected that increasing number of males are getting married with younger brides than earlier due to shortage in marriage cohorts. Though at this point, I do not have district-wise data on marriage statistics over the years, but the national level or nationally representative survey data can give us some idea about the marriage trends over the years. Table 3.12 shows marriage statistics from Census 1991, 2001 and 2011 in Panel

Table 3.12: Some Marriage Statistics of India:

Panel A - Census:	1991		2001		2011	
Married individuals (in %):						
Age Group	Female	Male	Female	Male	Female	Male
10-14	4.5	2.1	2.4	0.9	2.7	1.5
15-19	35.3	9.4	24.4	5.2	19.5	4.7
20-24	81.8	39.6	75.7	34.3	68.5	30.4

Source: Census of India (1991, 2001, and 2011)

Panel B - NFHS:	2005-06	2015-16
Inter-caste marriage (%)	9.9	12.6
Inter-religious marriage (%)	2.1	2.6
Women get married (in %):		
Before age 10	1.0	1.6
Before age 15	12.8	12.5
Before age 18	39.5	38.4
Couples with age gap: (%)		
Below 0 years	5.1	3.04
0 Years	3.85	3.16
1 years	5.88	7.62
2 years	8.70	14.84
3-5 years	31.24	39.21
6-10 years	34.38	25.79
More than 10 years	10.85	6.34

Source: NFHS-3 (2005-06) and NFHS-4 (2015-16).

A and also some related data from nationally representative survey, NFHS-3 (2005-06) and NFHS-4 (2015-16) in Panel B. The census data shows the percentage of individuals who are already married in total number of individuals across age-groups. Even at the early age as 10-14 years around 4.5 per cent girls in 1991 and around 2.5 per cent girls in 2001 and 2011 are already married. In 1991, at the age-group of 15-19 years, around 35 per cent girls were already married, whereas only around 10 per cent boys at the same age were married. Though the percentage of married girls has declined over the last two censuses, in 2001 and 2011, but still the percentage is very high, such as around 20 per cent girls get married within 15 to 19 years of age whereas only 5 per cent boys among total boys at the same age got married according to Census 2011. According to UNICEF, even in 2018, around 47 per cent girls get married before their 18^{th} birthday and around 18 per cent girls get married before 15 years of age. In states of Bihar and Rajasthan, the prevalence of child marriage among girls are as high as above 60 per cent (UNICEF et al., 2018). The 2017 UNICEF report estimated the prevalence of child marriage in India as 27 per cent (UNICEF et al., 2017). In table 3.12 panel B, the NFHS data also provide evidence that around 13 per cent girls get married before age 15 and around 38-40 per cent girls get married before the age of 18 years. It is also important to mention that in India the legal age of marriage for girls is 18 years (21 years of age for boys), by the Prohibition of Child Marriage Act. The information from three different sources support the wide prevalence of child marriages among Indian girls, even today.

In the third (2005-06) and fourth (2015-16) rounds of NFHS data, the information on inter-caste and inter-religious marriages show that both these types of marriages have increased from NFHS-3 to NFHS-4. The percentage of inter-caste marriages increased from around 10 per cent in 2005-06 to 12.6 per cent in 2015-16. Though the percentage of inter-religious marriages are low compared to inter-caste marriages, but it has also increased over 10 years. In addition, a report on human trafficking in 2013 by UN Office of Drugs and Crime (UNODC) has shown concern about increasing trafficking of young girls to meet the shortages of bride in states with acute male-biased sex ratios, such as Haryana, Punjab and Uttar Pradesh. The report cites a survey of 92 villages in Haryana

which shows that around 90 per cent married women had been bought from poor villages in other states (UNODC & Union, 2013).

The NFHS-3 and NFHS-4 have information on the age gap (husband's age - wife's age) between couples, and show that the percentage of marriages within age gap of 1, 2 and 3-5 years have increased from 2005-06 to 2015-16, but for age gap of 6-10 and more than 10 years has decreased from NFHS-3 to NFHS-4. Though percentage of marriages in higher age-gap has not increased, but it can be the case that in more progressive states with increasing female labor-force participation, women in such states get married late and have partners of similar age, thus, the age-gap widening is not abundant in these states. But, in states where acute pro-male biased sex ratios are prevalent, higher percentage of males may get married with younger girls due to shortage. In national data the progressive and patriarchal states statistics may balance each other. To examine the age-gap trend among couples at district levels, more detail data would be required.

Overall, the descriptive statistics on Indian marriages from different sources provide evidence to the argument that shortages of female cohort to marry have impacts on marriage market of India, in forms of increasing inter-caste, inter-religious and inter-state marriages, and also leads to child marriage in various parts of the country. Female, who marry at early ages and even before adulthood, tend to remain less educated, less involved in labor market, begin having children earlier, will have less bargaining power in the family and will be more vulnerable to abuse from the family members. In addition, the cross marriages bring the brides to not-so-familiar cultural set-up, therefore these girls need more efforts and time to adjust with culture, customs and location and are more likely to face higher discrimination and violence due to lack of support from in-laws family and from the unknown neighborhood.

3.7 Discussion and Policy Implications

As a result of widely adopted sex-selective family-planning in the prevalence of preference for sons along with patriarchal norms, India has an increasingly biased sex ratios in favor

of males. With this, the coexistence of increasing criminal activities, especially larger increase in crime against women compared to other crimes generates a puzzle that in spite of scarcity of women their value in the society does not increase, but creates an increasingly unsafe society for women. In this paper, I examine this coincidence of biased sex ratio in favor of male and higher incidences of violence against women. Using panel data on population from Census and crime data from NCRB for 395 districts of India from 1995 to 2013, I estimate the relationship between sex ratios and crime against women.

The results from the analysis indicate that there is a significant negative impact of pro-male biased sex ratios (YSR and lagged CSR) on crime against women, that is while sex ratios increase in favor of females the crime rate against women may decline. However, it is important to mention that higher crime rate can be attributed to various factors and it is not justified to say that only gender imbalances in population have caused it, but better to say that it is partly responsible for increasing crime, especially incidences against women. In this analysis, I have controlled for urbanization, literacy rate, past crimes and female population, but there can be some other factors that contribute to lower or higher crime, such as poor law and order in some districts than others. Due to data unavailability for districts over the years, other relevant variables remain omitted from this analysis.

The main results reveal a negative and significant relation between sex ratios and crime against women. An increase in YSR by 1 unit (or 1 female per 1000 males) can reduce crime against women by 0.53 per cent on average, *ceteris paribus*. This estimated impact is statistically significant at 5 per cent level. Further, it is also estimated that an improvement in the past CSR (with 15 years lag) by 1 unit would have significantly reduce recent crime against women by 0.093 per cent. India, being the second highest populated country in the world, the size of the impact is not trivial. It demands for a proactive role of authorities to correct the imbalances in demographic numbers and create a safer society for all, particularly for women.

Comparing the effects of imbalanced sex ratios on crime against women, any other crime except CRW and crime in total, I found that the impact on crime against women is higher

than the impact on other crimes. Therefore, it indicates that the surplus male can cause a more violent society in general, but the situation will be worse for women compared to men. However, using district specific linear time trend, the impact of youth sex ratio was highest on other crime.

In this paper, I also examine the impact of sex ratios on crime against women across the regions and also across quintiles of YSR. The results indicate that the locations with most biased sex ratio have the most significant impact of sex ratios on CRW. For districts with moderately biased to balanced sex ratios the impact becomes insignificant. In the most biased YSR districts, the impact of improvement in YSR in favor of female has the most significant crime reducing impact until the YSR values lie in the second lowest quintile range, and starting from the third quintile onwards (YSR values above 890 females per 1000 males) YSR has increasingly positive relation with CRW. The quintile-wise analysis gives a direction towards the locations that need actions to be taken on priority basis, as correcting the demographic imbalances would be most successful in reducing crime in the districts with the most biased sex ratios.

The region-wise analysis shows that the impact of sex ratios on CRW in the western part is highest among the six regions of India. Crime against women can be significantly (at 5 per cent level) reduced by 1.3 per cent in the western region only through correcting YSR by 1 unit. In addition, if past CSR (with 15 years lag) would increase by only 1 unit, it could cause 0.5 per cent reduction in current crime rate against women on average in the western part of the country. The most orthodox states of India belong to the northern part. The result shows negative impacts of both YSR and CSR_{-15} on CRW in the northern and southern regions, but these results are statistically insignificant. However, in the northern region the biased past child sex ratio has significant impact on both other and total crime.

It is surprising that the eastern region does not show a negative relation between sex ratios and CRW. Though this region shows very high rates in all types of crime incidences against women, the estimated result indicates that this high crime rate is not significantly related with the sex ratios in the region. The sex ratios in the eastern region is not as

biased as the sex ratios in the northern and western regions of India. Among the eastern states, especially in West Bengal many districts show very high incidences of crime against women in the analysis years, however these districts do not record very imbalanced sex ratios. It is ambiguous in the case of West Bengal, as in one hand the state is considered as a more progressive state in terms of women empowerment and probably that causes higher reporting of violence against women, whereas on the other hand the statistics show higher rates of domestic violence which indicate the presence of patriarchal power relations within the household setup in the state.

In addition, a detail crime-head-wise analysis provides evidence that it is not rape incidences which have the highest growth but domestic violence has the highest increase over the analysis period. It is also found that the impact of sex ratios is the highest on domestic violence, followed by kidnapping compared to other CRW. An increase in YSR and past CSR_{-15} by 1 unit would reduce domestic violence incidences by 0.78 and 0.13 per cent on average, respectively. These results are statistically significant at 5 per cent level. Similarly kidnapping of girls can be reduced by 0.74 and 0.12 per cent due to improvements in YSR and CSR_{-15} by 1 unit, respectively.

It can be explained that imbalances in the ratio of females to males can distort peace and even safety within household due to the power relations of males with female family members. The lower number of female on average compared to males in the household may emphasize the dominance by male members and can result in more physical and mental violence towards females in the family. The negative relation and non-trivial size of the impact of sex ratio on dowry death also reinforces the belief that more males can create a more patriarchal society and the norms of patriarchy will be adopted vigorously in such society. Without the fulfillment of such norms, like dowry system can cause extreme violence towards women, even death.

Therefore, the crime-head-wise findings provide evidences towards the channel that lower ratio of female to male leads to increase women's vulnerability within household due to increasing inter- cultural, inter-caste, inter-religion and inter-region marriages. However, neither rape nor molestation holds a negative and significant relation with both current

YSR and/or past CSR. The impact of imbalanced sex ratio on sexual harassment is much smaller in size and insignificant compared to the same on domestic violence, kidnapping and dowry death, though holds a negative relation. Therefore, we can say that the surplus men can cause a more violent society in general, but the frustrations of surplus men do not channel into sexual violence towards women and extreme crime like rape.

The evidence of this significant negative impact of sex ratios on crime against women directs a distinctive area for policy implication that women safety and security situations can improve partly by correcting the sex ratio imbalances. Though sex selective abortions have been controlled upto some extent by the amendment of PCPNDT act, it has not stopped fully and an underground market still is in operations to provide such illegal services. In addition, sex selective family planning depending on the gender of the earlier child(ren) still continues in varying extent across India, at higher rate in northern and western parts of India. The use of past CSR in the relation with CRW also implies that the preference for son from the past results into increasingly unsafe society for women today. If the practices of fulfilling the desire for sons continue, it will make the situation worse in future, give us a more dangerous society for women.

The imbalances in sex ratio against female have many negative consequences, the household level decisions of having boy(s) instead of girl(s) affect the whole nation in an aggregate level with an increasingly violent society. It becomes important for government and authorities to adopt a more proactive role to balance the number of females with males in the coming generations. This task becomes difficult in case of India due to its size of population as the second highest in the world. Government initiatives towards restricting population growth and promoting smaller families may actually go against the initiatives towards the balancing of sex ratios. Population control policies can influence the couples to become strictly adherent towards preference for son, and thus if only two children they can have, they would likely to have at least a son. Therefore, it is utmost necessary to implement the PCPNDT Act vigorously and strict regulation should be implemented on sex-determination procedures and thus to stop sex-selective abortions. In addition, civil society, non-governmental organisations and media can play a crucial role through

advocacy, mass campaign and awareness-raising programmes on importance of girls and the drawbacks of having surplus men. It is required to change people's mindsets and attitudes towards girls and highlighting women's success stories though campaigns may have positive influences on the views that only see sons being worthy in carrying forward family name and reputation.

Other initiatives that can raise the importance of girl child in society are equally important. The equal right of inheritance in parents property irrespective of gender of the offspring, strict dowry prohibition laws and better policies towards old-age security, improvement in pension schemes will be useful to reduce parents dependence particularly on sons. Apart from these, women empowerment through better access to education and job opportunities may increase women's financial situation, and they can have more say in marriage, child-bearing and other important decisions of life, and also taking care of own parents even after being married. The government authorities should undertake a very strict policy and legal framework to stop child marriage.

Government bodies may also provide incentives to parents for birth of girl child, providing child-care fund to poor parents at the birth of girl child, scholarship for girls' education, etc. These can be helpful to share the so-called burden of bringing up a girl child and also can be helpful in empowerment of them. The reduction in preference for sons and thus discrimination against girls need continuous and multiple initiatives in the long term to see the fruitful results in terms of sex ratio.

In countries like India where population pressure is a problem in itself, a very high share of male in comparison to female can only add to the problem by resulting an increasingly unsafe society for all. If policies would not improve sex ratios in favor of females; it not only challenges the safety of women in the society but it may also enhance gender gap in areas like education, income, labor market; reduce labor force participation; narrow down choices and opportunities of jobs for women; reduce earning and bargaining power of women in family and in society.

Appendix 3

Table A3.1: Panel level Descriptive Statistics

Variables		Mean	Std. Dev.	Min	Max	Observations
Crime against women	overall	329.498	434.794	0	12853	N = 7505
	between		361.124	0.421	3942.526	n = 395
	within		242.793	-1913.028	9239.972	T = 19
$Rape_{-15}$	overall	22.300	25.742	0	544	N = 7358
	between		20.938	0	205.526	n = 395
	within		14.991	-183.226	360.774	T-bar = 18.628
Total crime	overall	4014.132	5026.613	0	80184	N = 7439
	between		4818.783	32.737	55875.26	n = 395
	within		1411.632	-13806.13	33293.82	T-bar = 18.833
Total $Crime_{-15}$	overall	3531.395	4087.959	10	64882	N = 7336
	between		3833.88	28.842	39193.26	n = 394
	within		1443.23	-29289.92	29220.13 —	T-bar = 18.619
YSR	overall	910.561	91.081	215.148	1243.183	N = 7505
	between		88.596	333.531	1179.932	n =
	within		21.568	739.409	1049.585	T = 19
CSR_{-15}	overall	958.300	41.941	628.735	1139.17	N = 7505
	between		35.533	854.047	1059.381	n = 395
	within		22.348	704.202	1112.102	T = 19
Urbanisation	overall	26.316	19.467	0	100	N = 7505
	between		18.881	0	100	n = 395
	within		4.830	-11.221	93.789	T = 19
Literacy	overall	58.062	12.221	6.235	93.628	N = 7505
	between		10.602	31.071	86.703	n = 395
	within		6.101	12.373	100.700	T = 19
ln(Female Population)	overall	13.517	1.044	9.572	16.031	N = 7505
	between		1.040	9.601	15.894	n = 395
	within		0.109	12.835	14.505	T = 19

Table A3.2: Impact of YSR on CRW after dropping Delhi and 24 Pargana

Dependent Variable: ln(Crime against Women)	Without Controls	With Controls	District Trend and w Controls
TSLS results:			
YSR	-0.00491**	-0.00546**	-0.00547**
	(0.00226)	(0.00219)	(0.00221)
Urbanization Rate		-0.000223	-0.000292
		(0.00224)	(0.00222)
Literacy Rate		-0.00558*	-0.00546*
		(0.00332)	(0.00329)
ln(Female Population)		0.513***	0.494***
		(0.174)	(0.172)
Rape with 15 years lag		-0.00271***	-0.00277***
		(0.00078)	(0.00078)
Constant	10.66***	4.305	-110.8***
	(2.217)	(3.671)	(7.763)
N	7320	7320	7320
R-sqr	0.910	0.911	0.911

Standard errors clustered at district levels are given in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A3.3: Impact of CSR_{-15} on CRW after dropping Delhi and 24 Pargana

Dependent Variable: ln(Crime against Women)	Without Controls	With Controls	District Trend and w Controls
Reduced form results:			
CSR with 15 years lag	-0.000849**	-0.000952**	-0.000947**
	(0.000400)	(0.000374)	(0.000376)
Urbanization Rate		0.00160	0.00158
		(0.00207)	(0.00204)
Literacy Rate		-0.00271	-0.00261
		(0.00249)	(0.00247)
ln(Female Population)		0.494***	0.470***
		(0.145)	(0.143)
Rape with 15 years lag		-0.00338***	-0.00344***
		(0.0007)	(0.0007)
Constant	6.708***	0.0376	-103.8***
	(0.409	(2.124)	(7.225)
N	7320	7320	7320
R-sqr	0.913	0.915	0.916

Standard errors are clustered at district levels and are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A3.4: Impact of sex ratios on crime against women across quintiles

Dependent Variable: ln(Crime against Women)	YSR Q1	YSR Q2	YSR Q3	YSR Q4	YSR Q5
TSLS regressions: Adding district time trend					
YSR	-0.00731**	-0.0135	0.00615	0.282	0.00720
	(0.00366)	(0.0161)	(0.0220)	(5.615)	(0.0110)
Urbanization Rate	-0.00560*	0.00262	-0.00517	0.0529	0.0106*
	(0.00334)	(0.00881)	(0.0120)	(1.114)	(0.00570)
Literacy Rate	-0.00967	-0.00195	0.00565	0.0696	-0.00931
	(0.00831)	(0.00859)	(0.0172)	(1.623)	(0.0216)
ln(Female Population)	0.108	0.656**	0.649**	-1.549	0.564
	(0.453)	(0.327)	(0.290)	(48.87)	(0.705)
Rape with 15 years lag	0.000962	-0.00397***	-0.00539***	-0.0295	-0.00314
	(0.000837)	(0.00133)	(0.00206)	(0.493)	(0.00313)
Reduced form regressions: Adding district time trend					
CSR with 15 years lag	-0.000706	-0.000941	0.000463	0.00101	0.000529
	(0.000577)	(0.00146)	(0.00163)	(0.00248)	(0.00194)
Urbanization Rate	-0.00119	0.00618	-0.00348	-0.00230	0.00972**
	(0.00227)	(0.00735)	(0.0107)	(0.00543)	(0.00400)
Literacy Rate	0.000392	-0.00337	0.000984	-0.0115**	0.00217
	(0.00514)	(0.00630)	(0.00439)	(0.00514)	(0.00956)
ln(Female Population)	0.461**	0.437	0.606**	0.942*	0.931**
	(0.229)	(0.270)	(0.235)	(0.523)	(0.402)
Rape with 15 years lag	0.000247	-0.00432***	-0.00492***	-0.00472***	-0.00260
	(0.000609)	(0.00121)	(0.00118)	(0.00163)	(0.00277)
N	1482	1467	1473	1481	1455

YSR quintile ranges are as follows:
Q1: 215.15 - 835.56; Q2: 835.65 - 890.61; Q3: 890.63 - 938.43; Q4: 938.45 - 989.34; and Q5: 989.38 - 1243.18.
Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A3.5: Region-wise States/UTs and number of districts

Regions:	States	Number of Districts
Northern Region	Chandigarh	1
	Delhi	1
	Haryana	13
	Himachal Pradesh	11
	Jammu and Kashmir	14
	Punjab	12
	Rajasthan	26
	Uttar Pradesh	47
	Uttarakhand	8
Southern Region	Andaman and Nicober Islands	2
	Andhra Pradesh	23
	Karnataka	19
	Kerala	12
	Puducherry	2
	Lakshadweep	1
	Tamil Nadu	15
Central Region	Chattisgarh	5
	Madhya Pradesh	38
Western Region	Dadra and Nagar Haveli	1
	Daman and Diu	2
	Goa	1
	Gujarat	18
	Maharashtra	26
Eastern Region	Bihar	22
	Jharkhand	5
	Odisha	12
	West Bengal	17
North Eastern Region	Arunachal Pradesh	8
	Assam	8
	Manipur	6
	Meghalaya	2
	Mizoram	3
	Nagaland	7
	Sikkim	4
	Tripura	3

Table A3.6: Impact of sex ratios on crime against women across regions

Dependent Variable:	Explanatory Variable	Northern	Southern	Western	Eastern	Central	North Eastern
TSLS results: Adding district time trend							
CRW	YSR	-0.00425	-0.0160	-0.0132***	0.00200	0.0153	0.0144
		(0.00585)	(0.0287)	(0.00480)	(0.00428)	(0.0256)	(0.0433)
CRO	YSR	-0.0489	-0.00708	-0.00314	0.000969	0.0222	0.0189
		(0.0724)	(0.0145)	(0.00374)	(0.00102)	(0.0152)	(0.0281)
CRT	YSR	-0.0420	-0.00809	-0.00383	0.00131	0.0209	0.0165
		(0.0619)	(0.0155)	(0.00360)	(0.00109)	(0.0145)	(0.0249)
Reduced Form: Adding district time trend							
CRW	CSR_{-15}	-0.000417	-0.00236	-0.00478***	0.000249	0.000598	-0.00110
		(0.000865)	(0.00218)	(0.00169)	(0.000518)	(0.00134)	(0.00143)
CRO	CSR_{-15}	-0.00353***	-0.000962	-0.00105	0.000345	0.00134	-0.00234**
		(0.000861)	(0.00156)	(0.00126)	(0.000342)	(0.000888)	(0.000937)
CRT	CSR_{-15}	-0.00313***	-0.00110	-0.00128	0.000399	0.00126	-0.00204**
		(0.000811)	(0.00154)	(0.00121)	(0.000363)	(0.000890)	(0.000971)

Standard errors in parentheses. Both Districts and Year FEs are used. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A3.7: Descriptive Statistics: Different crime-head under CRW (1995-2013)

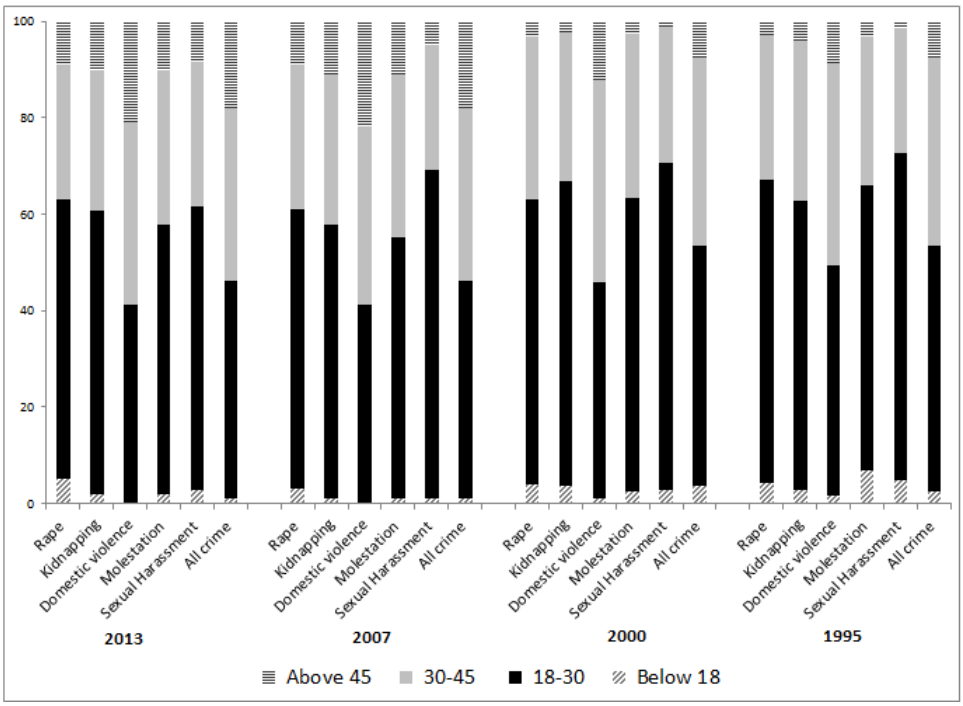
Variables	Obs	Mean	Std. Dev.	Min	Max
Rape	7409	37.5168	48.5052	0	1636
Kidnapping	7339	42.8451	89.9438	0	3609
Domestic Violence	7282	140.9179	251.7133	0	5640
Molestation	7303	77.8584	97.0633	0	3515
Sexual Harassment	7042	21.6531	59.1259	0	2154
Dowry Death	7197	14.9573	18.4562	0	168

Table A3.8: Crime-head Wise Analysis: Excluding extreme observations (outliers)

Dependent Variable:	TSLS:			Reduced Form:		
Crime-head:	Main explanatory variable -			Main explanatory variable -		
	YSR	YSR	YSR	CSR_{-15}	CSR_{-15}	CSR_{-15}
	Excluding observations (12) that have rape above 500					
Rape	0.000418	0.000177	0.000252	0.0000684	0.0000304	0.0000431
	(0.00195)	(0.00184)	(0.00184)	(0.000322)	(0.000318)	(0.000315)
	Excluding observations (21) that have kidnapping above 500					
Kidnapping	-0.00867***	-0.00739***	-0.00741***	-0.00140***	-0.00126***	-0.00125**
	(0.00298)	(0.00275)	(0.00280)	(0.000496)	(0.000489)	(0.000490)
	Excluding observations (332) that have domestic violence cases above 500					
Domestic Violence	-0.0108***	-0.0104***	-0.0104***	-0.00166***	-0.00171***	-0.00168***
	(0.00400)	(0.00352)	(0.00354)	(0.000474)	(0.000470)	(0.000471)
	Excluding observations (41) that have molestation cases above 500					
Molestation	0.000127	-0.000293	-0.000308	0.0000191	-0.0000466	-0.0000482
	(0.00279)	(0.00249)	(0.00252)	(0.000418)	(0.000396)	(0.000396)
	Excluding observations (13) that have sexual harassment cases above 500					
Sexual Harassment	0.000408	-0.000329	-0.000404	0.0000578	-0.0000500	-0.0000606
	(0.00323)	(0.00305)	(0.00309)	(0.000457)	(0.000464)	(0.000465)
	Excluding observations (8) that have dowry deaths above 140					
Dowry Death	-0.00180	-0.00201	-0.00199	-0.000280	-0.000328	-0.000321
	(0.00229)	(0.00204)	(0.00207)	(0.000375)	(0.000350)	(0.000351)
Controls	No	Yes	Yes	No	Yes	Yes
District FE	Yes	Yes	No	Yes	Yes	No
State FE	Yes	Yes	Yes	Yes	Yes	Yes
District Trend	No	No	Yes	No	No	Yes

Standard errors in parentheses. Both Districts and Year FEs are used. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure A3.1: Age group-wise percentage of persons arrested for different IPC Crime-head under CRW



Source: Author's calculation using NCRB data.
Note: Age group for years 1995 and 2000 are below 18, 18-30, 30-50 and above 50.

References

Agal, R. (2006). India's 'bride buying'country. *BBC News, South Asia. April*, *5*.

Agarwal, B. (1997). "bargaining"and gender relations: Within and beyond the household. *Feminist economics*, *3*(1), 1–51.

Agnihotri, S. B. (1997). Workforce participation, kinship and sex ratio variations in india. *Gender, Technology and Development*, *1*(1), 75–112.

Ahlawat, N. (2009). Missing brides in rural haryana: A study of adverse sex ratio, poverty and addiction. *Social change*, *39*(1), 46–63.

Ahmed, S. M. (2005). Intimate partner violence against women: experiences from a woman-focused development programme in matlab, bangladesh. *Journal of Health, Population and Nutrition*, 95–101.

Alderman, H., & King, E. M. (1998). Gender differences in parental investment in education. *Structural Change and Economic Dynamics*, *9*(4), 453–468.

Ali, T. S., & Bustamante Gavino, I. (2007). Prevalence of and reasons for domestic violence among women from low socioeconomic communities of karachi. *World Health Organization*.

Amaral, S., & Bhalotra, S. (2017). *Population sex ratios and violence against women: The long-run effects of sex selection in india* (Tech. Rep.).

Andersson, G., & Woldemicael, G. (2001). Sex composition of children as a determinant

of marriage disruption and marriage formation: Evidence from swedish register data. *Journal of Population Research*, *18*(2), 143–153.

Angrist, J. (2002). How do sex ratios affect marriage and labor markets? evidence from america's second generation. *The Quarterly Journal of Economics*, *117*(3), 997–1038.

Arokiasamy, P. (2004). Les variations régionales de la discrimination et de la surmortalité des petites filles en inde. *Population*, *59*(6), 947–982.

Attanasio, O., & Kaufmann, K. (2009). *Educational choices, subjective expectations, and credit constraints* (Tech. Rep.).

Attanasio, O., Kugler, A., & Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, *3*(3), 188–220.

Attanasio, O. P., & Kaufmann, K. M. (2014). Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics*, *109*, 203–216.

Barber, N. (2000). The sex ratio as a predictor of cross-national variation in violent crime. *Cross-Cultural Research*, *34*(3), 264–282.

Barber, N. (2003). The sex ratio and female marital opportunity as historical predictors of violent crime in england, scotland, and the united states. *Cross-Cultural Research*, *37*(4), 373–392.

Bardhan, P. K. (1974). On life and death questions. *Economic and Political Weekly*, 1293–1304.

Beegle, K., Frankenberg, E., & Thomas, D. (2001). Bargaining power within couples and use of prenatal and delivery care in indonesia. *Studies in family planning*, *32*(2), 130–146.

Bhatt, H. (2011). Migrant population blamed for low sex ratio in daman and diu. *The Times of India*, *March 31*.

Bittman, M., England, P., Sayer, L., Folbre, N., & Matheson, G. (2003). When does gender trump money? bargaining and time in household work. *American Journal*

of sociology, *109*(1), 186–214.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, *120*(2), 669–700.

Blanchet, T. (2005). Bangladeshi girls sold as wives in north india. *Indian Journal of Gender Studies*, *12*(2-3), 305–334.

Bommier, A., & Lambert, S. (2000). Education demand and age at school enrollment in tanzania. *Journal of Human Resources*, 177–203.

Bose, S., Trent, K., & South, S. J. (2013). The effect of a male surplus on intimate partner violence in india. *Economic and political weekly*, *48*(35).

Brown, P. H. (2009). Dowry and intrahousehold bargaining evidence from china. *Journal of Human Resources*, *44*(1), 25–46.

Bureau, N. C. R. (1995-2013). *National crime records bureau (ncrb), india.* New Delhi: Ministry of Home Affairs, Government of India.

Census. (1951-2011). *Registrar general of india.* Registrar General and Census Commissioner.

Census. (2011). *Registrar general of india.* Registrar General and Census Commissioner.

Chakraborty, L. S., & Sinha, D. (2006). Determinants of declining child sex ratio in india: An empirical investigation. *MPRA paper*.

Chaudhry, S., & Mohan, T. D. (2011). Of marriage and migration: Bengali and bihari brides in a up village. *Indian Journal of Gender Studies*, *18*(3), 311–340.

Cochrane, S. H., Mehra, K., & Osheba, I. T. (1986). *The educational participation of egyptian children.* World Bank Washington, DC.

Davin, D. (2005). Marriage migration in china: The enlargement of marriage markets in the era of market reforms. *Indian Journal of Gender Studies*, *12*(2-3), 173–188.

Davin, D. (2007). Marriage migration in china and east asia. *Journal of Contemporary China*, *16*(50), 83–95.

Deere, C. D., & Doss, C. R. (2006). The gender asset gap: What do we know and why does it matter? *Feminist economics*, *12*(1-2), 1–50.

Ding, Q. J., & Hesketh, T. (2006). *Family size, fertility preferences, and sex ratio in china in the era of the one child family policy: results from national family planning and reproductive health survey.* British Medical Journal Publishing Group.

Doss, C. (2013). *Intrahousehold bargaining and resource ocation in developing countries.* The World Bank.

Doss, C. R., et al. (1996). *Women's bargaining power in household economic decisions: Evidence from ghana* (Tech. Rep.).

Dreze, J., & Khera, R. (2000). Crime, gender, and society in india: Insights from homicide data. *Population and development review*, *26*(2), 335–352.

Dyson, T., & Moore, M. (1983). On kinship structure, female autonomy, and demographic behavior in india. *Population and development review*, 35–60.

Edlund, L., Li, H., Yi, J., & Zhang, J. (2007). More men, more crime: Evidence from china's one-child policy. *IZA Working Paper*(3214).

Edlund, L., Li, H., Yi, J., & Zhang, J. (2013). Sex ratios and crime: Evidence from china. *Review of Economics and Statistics*, *95*(5), 1520–1534.

Fafchamps, M., & Quisumbing, A. R. (2002). Control and ownership of assets within rural ethiopian households. *Journal of development Studies*, *38*(6), 47–82.

Frempong, R. B., Stadelmann, D., et al. (2017). *Does female education have a bargaining effect on household welfare? evidence from ghana and uganda* (Tech. Rep.).

Friedberg, L., & Webb, A. (2006). *Determinants and consequences of bargaining power in households* (Tech. Rep.).

Gandhi Kingdon, G. (2002). The gender gap in educational attainment in india: How much can be explained? *Journal of Development Studies*, *39*(2), 25–53.

Gertler, P., & Glewwe, P. (1992). The willingness to pay for education for daughters in contrast to sons: Evidence from rural peru. *The World Bank Economic Review*, *6*(1), 171–188.

Ghouri, A. M., & Abrar, N. (2010). *The women violence in pakistan: evidence from rural and urban areas.*

Glick, P., & Sahn, D. E. (2000). Schooling of girls and boys in a west african country:

the effects of parental education, income, and household structure. *Economics of education review*, *19*(1), 63–87.

Gomes, M. (1984). Family size and educational attainment in kenya. *Population and Development Review*, 647–660.

Graham, J., & Bowling, B. (1995). *Young people and crime.* Home Office Research and Statistic Department, London.

Gray, J. S. (1998). Divorce-law changes, household bargaining, and married women's labor supply. *The American Economic Review*, *88*(3), 628–642.

Greenhalgh, S. (1985). Sexual stratification: The other side of" growth with equity" in east asia. *Population and development review*, 265–314.

Guilmoto, C. (2012). *Sex imbalances at birth: current trends, consequences and policy implications*. UNFPA.

Guilmoto, C. Z. (2012). Skewed sex ratios at birth and future marriage squeeze in china and india, 2005–2100. *Demography*, *49*(1), 77–100.

Guilmoto, C. Z., & Attané, I. (2007). The geography of deteriorating child sex ratio in china and india. *Watering the neighbour's garden: The growing demographic female deficit in Asia*.

Gupta, M. D. (1987). Selective discrimination against female children in rural punjab, india. *Population and development review*, 77–100.

Guttentag, M., & Secord, P. F. (1983). *Too many women? the sex ratio question*. Beverly Hills Calif./London England Sage Publications.

Hesketh, T., Lu, L., & Xing, Z. W. (2011). The consequences of son preference and sex-selective abortion in china and other asian countries. *Canadian Medical Association Journal*, *183*(12), 1374–1377.

Hesketh, T., & Xing, Z. W. (2006). Abnormal sex ratios in human populations: causes and consequences. *Proceedings of the National Academy of Sciences*, *103*(36), 13271–13275.

Hill, M. A., & King, E. (1995). Women's education and economic well-being. *Feminist Economics*, *1*(2), 21–46.

Hoddinott, J., & Haddad, L. (1995). Does female income share influence household expenditures? evidence from côte d'ivoire. *oxford Bulletin of Economics and Statistics*, *57*(1), 77–96.

Hudson, V. M., & Boer, A. D. (2002). A surplus of men, a deficit of peace: Security and sex ratios in asia's largest states. *International Security*, *26*(4), 5–38.

Hudson, V. M., & Den Boer, A. (2004). *Bare branches: The security implications of asia's surplus male population*. MIT Press.

Imai, K. S., Annim, S. K., Gaiha, R., Kulkarni, V. S., et al. (2012). Does women's empowerment reduce prevalence of stunted and underweight children in rural india? *DP2012-11, Kobe University*.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, *125*(2), 515–548.

Jensen, R., & Thornton, R. (2003). Early female marriage in the developing world. *Gender & Development*, *11*(2), 9–19.

Jha, P., Kesler, M. A., Kumar, R., Ram, F., Ram, U., Aleksandrowicz, L., ... Banthia, J. K. (2011). Trends in selective abortions of girls in india: analysis of nationally representative birth histories from 1990 to 2005 and census data from 1991 to 2011. *The Lancet*, *377*(9781), 1921–1928.

Jha, P., Kumar, R., Vasa, P., Dhingra, N., Thiruchelvam, D., & Moineddin, R. (2006). Low male-to-female sex ratio of children born in india: national survey of 1· 1 million households. *The Lancet*, *367*(9506), 211–218.

Jiang, Q., Feldman, M. W., & Li, S. (2014). Marriage squeeze, never-married proportion, and mean age at first marriage in china. *Population Research and Policy Review*, *33*(2), 189–204.

Johansson, S., & Nygren, O. (1991). The missing girls of china: A new demographic account. *The Population and Development Review*, 35–51.

Kaur, R. (2004). Across-region marriages: Poverty, female migration and the sex ratio. *Economic and political weekly*, 2595–2603.

Kaur, R. (2008). Dispensable daughters and bachelor sons: Sex discrimination in north

india. *Economic and Political Weekly*, 109–114.

Kaur, R. (2010). Bengali bridal diaspora: Marriage as a livelihood strategy. *Economic and Political Weekly*, 16–18.

Kaur, R. (2012). Marriage and migration: Citizenship and marital experience in cross-border marriages between uttar pradesh, west bengal and bangladesh. *Economic and Political Weekly*, *47*(43), 79.

Kaur, R. (2013). Mapping the adverse consequences of sex selection and gender imbalance in india and china. *Economic and Political Weekly*, *48*(35), 37–44.

Keane, M. P., & Wolpin, K. I. (1997). The career decisions of young men. *Journal of political Economy*, *105*(3), 473–522.

Klasen, S., & Wink, C. (2003). " missing women": Revisiting the debate. *Feminist Economics*, *9*(2-3), 263–299.

Knodel, J., Havanon, N., & Sittitrai, W. (1990). Family size and the education of children in the context of rapid fertility decline. *Population and Development Review*, 31–62.

Knodel, J., & Wongsith, M. (1991). Family size and children's education in thailand: Evidence from a national sample. *Demography*, *28*(1), 119–131.

Koenig, M. A., Ahmed, S., Hossain, M. B., & Mozumder, A. K. A. (2003). Women's status and domestic violence in rural bangladesh: individual-and community-level effects. *Demography*, *40*(2), 269–288.

Kukreja, R., & Kumar, P. (2013). Tied in a knot: Cross-region marriages in haryana and rajasthan: Implications for gender rights and gender relations. *Tamarind Tree Films, Kingston and New Delhi*.

Kumar, D. (2012). *India's long-distance brides.*

Kumar, H., & Somanathan, R. (2009). Mapping indian districts across census years, 1971-2001. *Economic and Political Weekly*, 69–73.

Lamichhane, P., Puri, M., Tamang, J., & Dulal, B. (2011). Women's status and violence against young married women in rural nepal. *BMC women's health*, *11*(1), 19.

Li, J., & Lavely, W. (2003). Village context, women's status, and son preference among

rural chinese women. *Rural Sociology*, *68*(1), 87–106.

Li, L., & Wu, X. (2011). Gender of children, bargaining power, and intrahousehold resource allocation in china. *Journal of Human Resources*, *46*(2), 295–316.

Li, S., Zhang, Q., Yang, X., & Attané, I. (2010). Male singlehood, poverty and sexuality in rural china: An exploratory survey. *Population*, *65*(4), 679–693.

Lillard, L. A., & Willis, R. J. (1994). Intergenerational educational mobility: Effects of family and state in malaysia. *Journal of Human Resources*, 1126–1166.

Lise, J., & Seitz, S. (2011). Consumption inequality and intra-household allocations. *The Review of Economic Studies*, *78*(1), 328–355.

Mason, K. O. (1984). *The status of women: A review of its relationships to fertility and mortality*. New York NY Rockefeller Foundation Population Sciences Division 1984.

Mauldin, T., Mimura, Y., & Lino, M. (2001). Parental expenditures on children's education. *Journal of Family and Economic Issues*, *22*(3), 221–241.

Mayer, P. (1999). India's falling sex ratios. *Population and development review*, *25*(2), 323–343.

Messner, S. F., & Sampson, R. J. (1991). The sex ratio, family disruption, and rates of violent crime: The paradox of demographic structure. *Social Forces*, *69*(3), 693–713.

MHRD. (2012). *Education statistics at a glance.* Ministry of Human Resource Development, Government of India.

Miller, B. D. (1981a). Changing patterns of juvenile sex ratios in rural india, 1961 to 1971. *Economic and Political Weekly*, *24(22)*, 1229–1236.

Miller, B. D. (1981b). *The endangered sex: Neglect of female children in rural north india.* Ithaca, New York and London: Cornell University Press.

Morgan, S. P., Lye, D. N., & Condran, G. A. (1988). Sons, daughters, and the risk of marital disruption. *American journal of sociology*, *94*(1), 110–129.

NFHS-3. (2009). *India national family health survey, 2005-06.* International Institute for Population Sciences (IIPS) and Macro International.

NFHS-4. (2015). *India national family health survey, 2012-14.* International Institute for Population Sciences (IIPS) and Macro International.

Oldenburg, P. (1992). Sex ratio, son preference and violence in india: A research note. *Economic and Political Weekly*, 2657–2662.

Pal, S. (2004). How much of the gender difference in child school enrolment can be explained? evidence from rural india. *Bulletin of Economic Research*, *56*(2), 133–158.

Panda, P., & Agarwal, B. (2005). Marital violence, human development and women's property status in india. *World development*, *33*(5), 823–850.

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

Pandey, G. (2011). *Where are india's millions of missing girls?* BBC News, South Asia.

Parish, W. L., & Willis, R. J. (1993). Daughters, education, and family budgets taiwan experiences. *Journal of Human Resources*, 863–898.

Park, C. B., & 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*, 59–84.

Pollak, R. A. (2005). *Bargaining power in marriage: Earnings, wage rates and household production* (Tech. Rep.).

Pong, S.-L. (1997). Sibship size and educational attainment in peninsular malaysia: Do policies matter? *Sociological Perspectives*, *40*(2), 227–242.

Prakash, N., & Vadlamannati, K. C. (2014). Girls for sale? child sex ratio and girls trafficking in india. *IZA Discussion Paper*, *June*(8293).

Premi, M. K. (2001). The missing girl child. *Economic and Political Weekly*, 1875–1880.

Raina, M. (2011). Leh springs sex ratio puzzle. *The Telegraph*, *April 8*.

Raley, S., & Bianchi, S. (2006). Sons, daughters, and family processes: Does gender of children matter? *Annu. Rev. Sociol.*, *32*, 401–421.

Reuben, E., Wiswall, M., & Zafar, B. (2017). Preferences and biases in educational

choices and labour market expectations: Shrinking the black box of gender. *The Economic Journal*, *127*(604), 2153–2186.

Saha, S., & Paul, K. (2017). Evolving patterns of child sex ratio in india. *Social Change*, *47*(2), 243–263.

Sami, N., & Ali, T. S. (2012). Domestic violence against infertile women in karachi, pakistan. *Asian Review of Social Sciences*, *1*(1), 15.

Schmidt, E. M. (2012). The effect of women's intrahousehold bargaining power on child health outcomes in bangladesh. *Undergraduate Economic Review*, *9*(1), 4.

Sen, A. (1972). The crisis in indian education. In A. Sen & P. Chaudhuri (Eds.), *Aspects of indian education* (chap. Lal Bahadur Shastri Memorial Lectures, 10–11 March 1970). Allen and Unwin.

Sen, A. (1990). More than 100 million women are missing. *New York*, 61–66.

Sen, A. (1992). Missing women. *BMJ: British Medical Journal*, *304*(6827), 587.

Shreeniwas, S. (1993). Family size, structure, and children's education: ethnic differentials over rime in peninsular malaysia. *Chapter in Lloyd (ed.)*.

Singariya, M. (2012). Determinants of declining child sex ratio in rajasthan. *Journal of Economics and Sustainable Development*, *3*(1), 9–19.

Singh, S. S. (2017). In bengal, crime against women at alarming levels. *The Hindu*.

South, S. J., & Trent, K. (1988). Sex ratios and women's roles: A cross-national analysis. *American Journal of sociology*, *93*(5), 1096–1115.

Srinivasan, S. (2017). Cross-region migration of brides and gender relations in a daughter deficit context. *Migration and Development*, *6*(1), 123–143.

Steffensmeier, D., & Allan, E. (1996). Gender and crime: Toward a gendered theory of female offending. *Annual review of sociology*, *22*(1), 459–487.

Teitelbaum, M. S. (1972). *Factors associated with the sex ratio in human populations.*

Tolhurst, R., Amekudzi, Y. P., Nyonator, F. K., Squire, S. B., & Theobald, S. (2008). "he will ask why the child gets sick so often": The gendered dynamics of intra-household bargaining over healthcare for children with fever in the volta region of ghana. *Social Science & Medicine*, *66*(5), 1106–1117.

Tucker, C., & Van Hook, J. (2013). Surplus chinese men: Demographic determinants of the sex ratio at marriageable ages in china. *Population and Development Review*, *39*(2), 209–229.

Tucker, J. D., Henderson, G. E., Wang, T. F., Huang, Y. Y., Parish, W., Pan, S. M., ... Cohen, M. S. (2005). Surplus men, sex work, and the spread of hiv in china. *Aids*, *19*(6), 539–547.

UNESCO. (2014). *Database.* UNESCO Institute for Statistics, Paris, France.

UNICEF, et al. (2017). *The state of the world's children 2012: children in an urban world* (Tech. Rep.).

UNICEF, et al. (2018). *The state of the world's children 2012: children in an urban world* (Tech. Rep.).

UNODC, & Union, E. (2013). *India country assessment report on human trafficking* (Tech. Rep.).

Visaria, P. M. (1969). *The sex ratio of the population of india.* Office of the Registrar General, India, Ministry of Home Affairs.

Waldron, I. (1983). Sex differences in human mortality: the role of genetic factors. *Social science & medicine*, *17*(6), 321–333.

Waldron, I. (1993). Recent trends in sex mortality ratios for adults in developed countries. *Social science & medicine*, *36*(4), 451–462.

Warner, R. L. (1991). Does the sex of your children matter? support for feminism among women and men in the united states and canada. *Journal of Marriage and the Family*, 1051–1056.

Washington, E. L. (2008). Female socialization: how daughters affect their legislator fathers. *American Economic Review*, *98*(1), 311–32.

Yang, W.-S., & Lu, M. C.-W. (2010). *Asian cross-border marriage migration: demographic patterns and social issues* (Vol. 2). Amsterdam University Press.

Zhang, J. (2010). *Imbalanced sex ratio at birth and women's human rights in china: A rights analysis and comparative implications.*

10.14754/CEU.2019.02