Gender Gap in Intergenerational Educational Mobility: Evidence from India

Azhar Uddin*

Abstract

The lack of credible data without any missing information acts as a constraint in studying intergenerational educational mobility in countries like India. I use a household survey data which collects complete information on educational attainment of father-child pairs along with a nationally representative data to examine the existence of a gender gap in intergenerational educational mobility in India. My findings suggest that sons have higher intergenerational educational mobility than daughters. In other words, given father's educational attainment, a son is more likely than a daughter to move up in the educational attainment ladder. From the latent class analysis, I find insufficient evidence to conclude that the gender gap in educational mobility is higher for families who have more bias gender norms. I also find evidence of caste-wise differences in the gender gap in educational mobility and that the gap is smaller for more educated fathers.

*School of Economic Sciences, Washington State University, 2020 Wilson Rd, Hulbert Hall 105 C, Pullman, WA 99164, E-mail: azhar.uddin@wsu.edu

Key words: Gender gap, Intergenrational educational mobility, Educational attainment

1 Introduction

Persistence of economic status across generations which affects the distributional justice in the society, is also a measure of the society's success in providing equal opportunity to everyone. A strong intergenerational persistence of economic status signals that the advantage of birth plays an important role in children's later success in life while a weak association would imply the opposite. In other words, a weak association of intergenerational persistence of economic status is a desirable outcome as it would ensure that everyone in the society has a equal opportunity to move up in the ladder of economic status regardless of the family background. Educational attainment creates that equal opportunity and plays a key role in the process of intergenerational economic mobility. Along with the role of equal opportunity which helps mitigating socioeconomic inequality, education also helps tackling gender disparity (see, for example, Stiglitz (2012), Duflo (2012), Jayachandran (2015)). In this paper, I examine the existence of a gender gap in intergenerational educational mobility in India and how this gender gap varies across different socio-economic groups. It is also important to understand the underlying factors that might affect such gender gap. I also examine if the bias gender norms could explain the existence of a gender gap in intergenerational educational mobility.

Along many different dimensions including educational attainment, gender gaps favoring men are systematically larger in poorer countries (Jayachandran, 2015). India is no exception

to this and stands out for its unequal opportunities and outcomes for women (Dhar et al., 2019). India also has the dubious distinction of intergenerational educational mobility remaining constant after the economic liberalization initiated in 1991, but with significant variation across groups. These two aspects of overall gender disparity and almost constant intergenerational educational mobility makes India suitable for studying the existence of a gender gap in intergenerational educational mobility and how it varies across different socio-economic groups.

The findings of the paper are as follows. First, I find evidence of a gender gap in educational mobility from a large nationally representative survey as well as from a household survey that collects a complete set of father-child pairs in one district in India. Second, I find caste wise differences in gender gap in educational mobility from the large nationally representative survey data set. Third, from the smaller household survey conducted in one district in India, I find evidence that the gender gap in educational mobility exists for certain section of the population, namely, those who have "moderate bias" gender norms while such gender gap does not exist for people having "high bias" and "low bias" gender norms.

The gender disparity in intergenerational transmission of educational attainment hinges on distributional justice. This disparity is particularly important in Indian context because of son preference, caste system and inequality (Azam, 2016). It becomes necessary for policy makers to understand whether there exists a gender disparity in parent-child transmission of educational attainment. In case of gender disparity, targeted policy should be initiated to ensure that every child gets equal and fair opportunities in obtaining the desired level of education irrespective of the gender of the child. It is even more important to know which

section of the population has this gender disparity so that targeted affirmative action plan can be initiated. We can understand such disparity from the prevalent caste system in India.

India's caste system is one of the oldest surviving social hierarchy. Historically, Hindus are divided into 4 distinct hierarchical groups based on their work and duties in the society. In post-independence era, constitution of India recognized four groups namely forward caste (FC), schedule caste (SC), schedule tribe (ST) and other backward class (OBC). There is another significant category of people based on religion who are called religious minority which includes Muslims, Sikhs, Christians, Buddhists, Parsis and Jains. Among the religious minorities, Muslims are the largest group. Despite efforts from the government for social justice, caste-based disparities in educational attainment persist. Evidence suggests that lower caste people (SC, ST, OBC) have less educational attainment than forward caste (Anitha, 2000) people. The government's affirmative actions have helped reducing the disparities between the lower caste (SC, ST, OBC) and forward caste over time while Muslims have fallen behind the lower caste (Desai & Kulkarni, 2008).

Literature on intergenerational mobility is well developed in developed countries. For an extensive survey, please see Solon (1999), Black and Devereux (2011), and Blanden (2013) for an excellent survey of the literature in developed countries. The empirical literature on intergenerational mobility in both developed and developing countries alike, has primarily focused on father-son transmission. Only a few studies can be found which examined father-daughter linkage (see Chadwick & Solon 2002; DiPrete & Grusky 1990; Olivetti & Paserman 2015, Azam 2016, Torche 2015b). In developing country context, the lack of studies on father-daughter linkage can be attributed to the non-availability of women's father information

(Azam, 2016).

Along with intergenerational income mobility, intergenerational educational mobility is also extensively studied in literature. Using a sample of 42 countries, Hertz et al. (2007) find large regional differences in educational persistence, with Latin America displaying the highest intergenerational correlations, and the Nordic countries the lowest. In Indian context, Jalan and Murgai (2008) find mobility in years of education across generations for different social groups and classes. Azam and Bhatt (2015) analyzed how intergenerational educational mobility in India varies across castes and states. Asher et al. (2018) also analyzed how intergenerational educational mobility in India varies across castes using a different methodology called interval mobility. Thus, the literature on educational mobility primarily concentrates on the relationship between a father and a son's educational attainment, ignoring the potential relationship between a father and a daughter's educational attainment with a few exceptions.

Credible data without non-random missing data is essential in analyzing the gender gap in intergenerational educational mobility. Majority of the existing household survey rely on co-residency to identify fathers' information (see Maitra and Sharma 2009, Hnatkovskay et al. 2013, Emran and Shilpi 2015). This leads to missing father-child pairs in a non-random fashion resulting truncation bias in the estimate of intergenerational regression coefficient to measure intergenerational mobility. This truncation bias is stronger for the girls as co-residency rates are lower and hence the gender gap would look smaller than the actual (Emran et al. (2018). To avoid this truncation bias Azam and Bhat (2015) and Azam (2016) used India Human Development Survey-II, 2011-12 (IHDS-II hereafter) data. IHDS-

II has a separate questionnaire for women and hence father's information of these women was also collected. So, matched father-daughter pairs were identified by Azam (2016). IHDS-II doesn't rely on co-residency to identify father's information as we have father's information for women. However, we still have missing father-child pairs. In IHDS-II, if the women respondent is the wife in the household, we do not have information on educational qualification of all her siblings. This way, we have missing children (and hence father-child pairs) even in this data set. This is similar to the truncation bias mentioned above. So, depending on whether we have missing information on sons or daughters (which is unknown), we will have high or low truncation bias resulting in unreliable estimate for the gender gap in educational mobility. To avoid this missing father-child pairs, I used the data from a household survey (EAS hereafter) I conducted in the state of Assam, India during the summer of 2019. In this survey, I collected information on educational attainment from 201 married couples. I also recorded the educational attainment of all siblings of both the husband and the wife. In this way, I ensured no possible missing father-child pairs in the sample I collected.

I use IHDS-II for matched father-son and father-daughter pairs to study the extent of intergenerational educational mobility in India since 1940s and if there is any gender-wise differences in this mobility. It is expected that the gender gap will exist in educational mobility. I am particularly interested if this gender gap in educational mobility is concentrated on particular sections of the population. So, I examine caste-wise gender disparity. As mentioned before, given the missing father-child pairs, I do the similar analysis with EAS data set which does not suffer from missing observations. However, this data set is not a nationally representative data and has much smaller sample compared to IHDS-II. In EAS

data set, I also have the information on gender norms of the parents. It is important to understand the process through which these gender norms are formed and if they are passed on to the next generation. The gender norms which decide the gender preferences could be influenced by religious doctrine (Psacharopoulos & Tzannatos, 1989; Seguino, 2011) among other factors. This factor is particularly important in Indian context given many prevalent religious beliefs. Dhar et al. (2019) finds strong positive correlation between parent and child attitudes, with mothers having greater influence than fathers. This leads us to think that gender norms might explain the gender gap in intergenerational educational mobility. Utilizing the gender norms information on EAS data set, I examine whether gender gap in educational mobility differs for different sections of the population having different gender norms.

The rest of this paper is organized as follows. Section 2 provides a discussion on the two data sets used for the empirical analysis, how matched father-child pairs were identified and the descriptive statistics of the two data sets used in the analysis. Section 3 outlines the empirical framework used for empirical analysis. The next section discusses the estimation results. The following section discusses the robustness check, and the final section concludes.

2 Data

I used two data sets for my analysis. The first data set is known as Indian Human Development Survey-II (IHDS-II) 2011-12, a nationally representative, multi-topic survey of 42,152 households. It covers 1,420 villages and 1,042 urban neighborhoods including all states and

union territories of India[1]. This survey of households is jointly organized by the National Council of Applied Economic Research (NCAER) and the University of Maryland. The survey was conducted between November 2011 and October 2012 and collected information on topics concerning health, education, employment, economic status, marriage, fertility, gender relations, social capital, village infrastructure, wage levels, and panchayat composition by face-to-face two one hour interviews in each household.

The second data set I used for analysis is from a household survey I conducted in the state of Assam, India in 2019 (EAS). A sample of 201 married couples were interviewed to gather information on educational attainment, gender norms and socio-demographic characteristics. Households were identified based on the condition that they had at least one child of at least 18 years old. The reason for this condition is to make sure that there is at least one pair of father-child where the child has almost completed his/her education. A face to face interview of about an hour was conducted with the head of the household (husband) as well as with the wife simultaneously. A male interviewer interviewed the husband while a female interviewer interviewed the wife.

The unique advantage of this data set is that, it collected educational attainment of siblings for both the husband and the wife. This information is not available for IHDS-II data set. I asked both the husband and the wife about the educational attainment of their siblings and in case they are not sure about educational attainment of their siblings, they were asked to confirm the same through a phone call to their siblings. In case of deceased siblings, I could not gather information on educational attainment as that information could

[1]The survey covers 33 states of union territories with the exception of Andaman and Nicobar and Lakshadweep that have small population

not be verified.

2.1 Identification of father-son and father-daughter pairs

I outline here how I matched father-son and father-daughter pairs and their educational attainment from both the datasets. To match unique father-child pairs, I do not have to rely on co-residency requirement as in many literature mentioned in the previous section. To help understand the data creation process, I have drawn a family tree diagram as shown below.

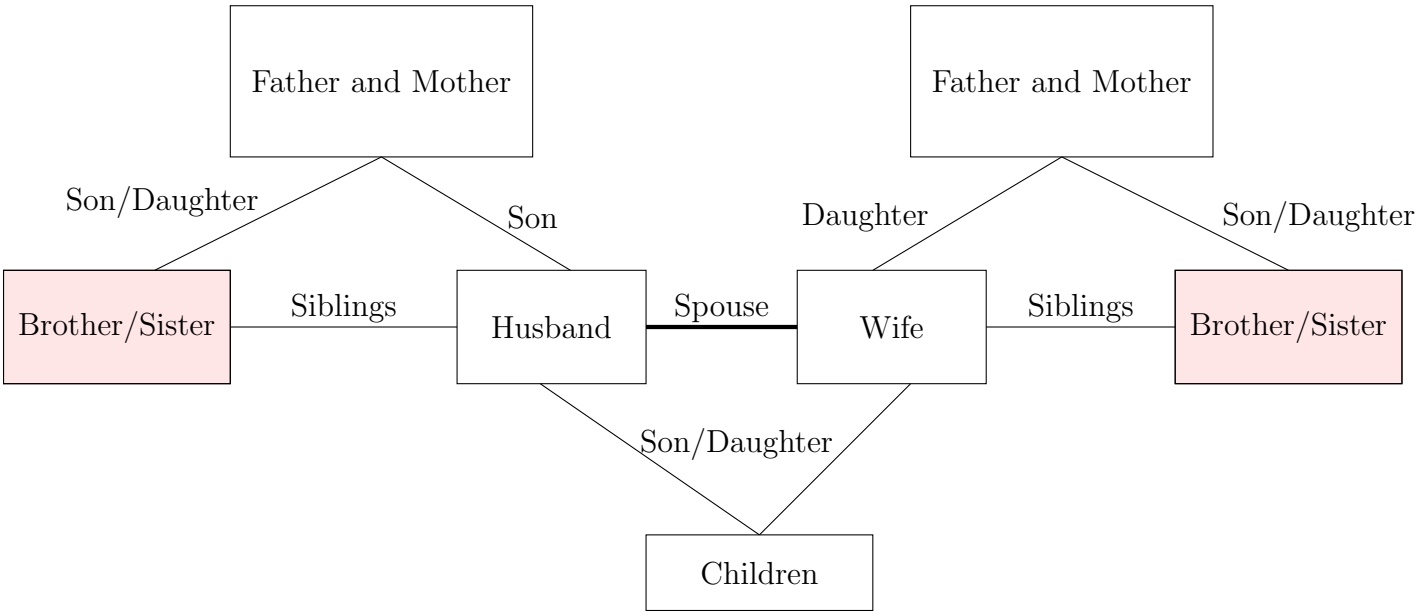

Figure 1: A family tree diagram

In IHDS-II, to locate father-son and father-daughter pairs, I only considered those individuals who are aged above 23. In Indian context, majority of people complete their college and university education around the age of 23. From the household roster, I identified the educational attainment of both the resident and non-resident father-son and father-daughter

pairs. In the figure 1 below, I collected educational attainment of "Husband", "Wife" and their "Children" (children can be resident as well as non-resident). For example, if there are 2 children above the age of 23 having educational attainment of 12 and 15, I will have two father-child pairs with the same level of education for father. I also have the information on educational attainment of household head's father. In other words, I have the educational attainment of "Husband"'s father. In this way, I have the information of educational attainment of 3 generations (Father-Husband-Children). For majority of the cases, household head is a male (husband).[2] The survey also has a women questionnaire, where an ever married women aged 15-49 were interviewed. If the woman was married to that household, I collected information on educational attainment of that woman's father. This is shown in the right pane of the diagram above. So, I collected father's educational attainment of wife/daughter-in-law's/sister-in-law/mother's. Note that if the woman is the daughter, then I already have that information from the household roster.

For the EAS data set, the process of identifying the matched father-son and father-daughter is similar. Here, I considered children aged above 21 given the smaller sample size compared to IHDS-II. The advantage of this data set is that it does not have any missing father-son and father-daughter pairs from the survey. In this data set, I also have the information on educational attainment of wife's siblings which was missing in IHDS-II data set. In terms of the diagram above, I did not have educational attainment of "Brother/Sister" for both the "Husband" and the "Wife" in in IHDS-II data set. Also in EAS data set, I

[2]In case the household head is female(wife), I have the educational attainment of her husband, which helps me to find matched father-son and father-daughters pairs from a household even when the head of household is female. Usually, a household has a female head when her husband is dead but the educational attainment of the dead husband was also collected.

collected the information on educational attainment of "Husband" and "Wife"'s mother which is not available for IHDS-II data set.

2.2 Descriptive Statistics

2.2.1 For Data set IHDS-II

After removing the missing values, I have educational attainment data of unique father-son and father-daughter pairs respectively. In the table below I summarize the educational attainment of all pairs.

Table 1: Years of education for father-son and father-daughter pairs

Father-son Pairs	N	Mean	SD	Min	Max
Father	51,508	3.36	4.35	0	16
Son	51,508	7.19	5.02	0	16
Father-daughter Pairs					
Father	37,226	3.34	4.41	0	16
Daughter	37,226	4.89	4.89	0	16

As you can see, the number of observations for father-daughter pair is less than the father-son pairs. I have 51,508 father-son pairs while I have 37,226 father-daughter pairs. The average years of education of a son is 7.19 years while for a daughter it is 4.89 years. The years of education of father is 3.36 and 3.34 years in fathers-son and father-daughter pairs respectively.

2.2.2 For Data set EAS

In the table below I summarize the educational attainment of all father-son and father-daughter pairs.

Table 2: Years of education for father-son and father-daughter pairs

Father-son Pairs	N	Mean	SD	Min	Max
Father	1,331	3.59	4.03	0	18
Son	1,331	7.75	4.87	0	18
Father-daughter Pairs					
Father	1,209	4.23	4.37	0	18
Daughter	1,209	7.40	4.46	0	18

As you can see, the number of observations for father-daughter pairs is less than the father-son pairs. I have 1,331 father-son pairs while I have 1,209 father-daughter pairs. The average years of education of a son is 7.75 years while for a daughter it is 7.40 years. The years of education of a fathers is 3.59 and 4.23 years in father-son and father-daughter pairs respectively.

We can see similar difference of educational attainment in both father-son and father-daughter pairs. On average, sons have 4.16 years of more education than fathers in father-son pairs while daughters have 3.17 years of more education than fathers in father-daughter pairs. These figures are 3.83 and 1.55 years for father-son and father-daughter pair respectively in IHDS-II data set. This suggests that gender-wise average years of education is more comparable in EAS data set compared to IHDS-II. From the summary itself of IHDS-II data set, we can have a fair idea that the daughters are at a disadvantageous positions compared

to sons in terms of educational attainment vis-a-vis their fathers' educational attainment. This makes the results of EAS data more reliable than IHDS-II data set.

3 Empirical Framework

3.1 For Data set IHDS-II

In this section, I outline the empirical framework of the study. I am interested in studying the gender gap in educational attainment of child vis-a-vis father's education. We can look at the existence of a gender gap in educational attainment from the upward mobility perspective. I define upward mobility in education when the child's educational attainment is greater than father's educational attainment in terms of the number of years of education. For example, if father attained 10 years of education, I am interested to know whether a son or a daughter has more chance of attaining more than 10 years of education. In other words, I am interested to know, given father's educational attainment, if a son is more likely to move up in the educational attainment ladder compared to a daughter. To measure this gender gap in educational mobility, I specify the empirical model as shown below,

$$Y_i^* = \beta_0 + \beta_1 MaleChild_i + \beta_2 AgeChild_i + U_i \tag{1}$$

where the dependent variable Y_i^* is a binary indicator equal to 1 if the child's educational attainment exceed that of the father's, and 0 otherwise. The independent variable *MaleChild* takes a value of 1 or 0 if the child is a son or a daughter respectively. The variable *AgeChild*

specifies the age of the child and U_i is the error term. Given the set up of the model, I am interested in the sign of β_1 (coefficient of the variable $MaleChild$). A positive sign of β_1 would indicate that a son has a higher probability than a daughter of attaining more education than the father. This implies that daughters do not have equal chance as sons in terms of educational attainment. In other words, there exists a gender gap in educational attainment for daughters vis-a-vis sons. This is my baseline regression. Once I find the evidence of a gender gap in educational mobility, I would be interested to know the characteristics of this gap. In particular, I am interested to know whether this gap is more or less for more educated fathers and how this gender gap is spread across different section of the population.

To understand whether the gap is more for less educated fathers, I add two variables *FathersEducation* and the interaction term *FathersEducation*Malechild* to the above baseline specification as shown below.

$$Y_i^* = \beta_0 + \beta_1 MaleChild_i + \beta_2 FathersEducation_i + \beta_3 MaleChild_i * FathersEducation + \beta_4 AgeChild_i + U_i \quad (2)$$

In this specification, a negative sign of β_3 would indicate that the gender gap in educational mobility is smaller for more educated fathers.

It is expected that the gender gap in educational mobility will not be similar across all section of the population. Given the caste dynamics in India, lower castes people are likely to have higher discrimination against acquiring higher education, more so for a women of a lower caste. That is why, it is very important to understand which section of the population

has lower educational mobility. To know which section of the population has more gender gap in educational mobility, I add the caste dummies and their interaction with the gender of child as shown below,

$$Y_i^* = \beta_0 + \beta_1 MaleChild_i + \beta_2 FathersEducation_i + \beta_3 MaleChild_i * FathersEducation + \beta_4 AgeChild_i + \beta_5 \sum_{i=1}^{6} Caste_i + \beta_6 \sum_{i=1}^{6} MaleChild_i * Caste_i + U_i \quad (3)$$

I will be interested in the caste interaction terms with the gender of child to know which section of the population has more gender gap compared to "Brahmin caste" (comparison group). I expect different behavior in this gender gap for different region and so I add district level fixed effect.

Now, I describe the two standard methods used in literature to measure educational mobility: intergenerational educational regression (IER) coefficient and intergenerational educational correlation (IEC) coefficient. These methods are particularly helpful when comparing mobility across country and time.

I show here IER coefficient to measure the gender gap in educational mobility. I estimate the following regression,

$$ChildsEducation_i = \beta_0 + \beta_1 MaleChild_i + \beta_2 FathersEducation_i + \beta_3 MaleChild_i * FathersEducation_i + \beta_4 ChildAge_i + U_i \quad (4)$$

where *ChildsEducation* and *FathersEducation* denote the years of education completed by

the child and the father respectively. *MaleChild* is a dummy variable (Son=1, daughter=0). I have controlled for child's age as educational attainment will be different depending on the age of the child. A negative sign on age coefficient will indicate that educational attainment is decreasing over time. I can find the marginal effect of father's education on child's education. So, I find the marginal effect of father's education on son's education as,

when $Gender_i = 1$,

$$\frac{\partial(ChildsEducation)}{\partial(FathersEducation)} = \beta_1 + \beta_3 \tag{5}$$

Similarly, the marginal effect of father's education on daughter's education can be found as,

when $Gender_i = 0$,

$$\frac{\partial(ChildEducation}{\partial(FathersEducation)} = \beta_1 \tag{6}$$

So, I can estimate the gender gap of educational attainment of a son and a daughter vis-a-vis father's educational attainment as the difference between equation (5) & (6) (i.e., β_3). A positive sign of β_3 would imply that a son's educational attainment is more persistent than a daughter's educational attainment vis-a-vis father's educational attainment. In other words, daughters' education has more intergenerational mobility than sons' education. So, a negative β_3 would mean that a son's educational mobility is higher than a daughter's mobility. This implies that there exists gender gap in educational mobility between a son and a daughter. If we want to see the caste wise gender gap in intergenerational educational mobility, we have to add two-way and three-way interaction dummies, which will be difficul to interpret. Hence, I am not showing caste wise gender gap in educational mobility in this

set up.

3.2 For data set EAS

My baseline regression is same as for IHDS-II data set as shown below,

$$Y_i^* = \beta_0 + \beta_1 MaleChild_i + \beta_2 AgeChild_i + U_i \tag{7}$$

where the dependent variable Y_i^* is defined as,

$$Y_i^* = \begin{cases} 1, & \text{if } ChildsEducation > FathersEducation \\ 0, & \text{otherwise} \end{cases}$$

Again, as in the case of IHDS-II data set, I am interested in the coefficient of β_1 and the interpretation is similar as described in the previous section. Given the small sample size and unavailability of all castes in this sample, I am unable to examine the other two questions, namely, whether the gender gap in educational mobility is lower for educated fathers and caste-wise differences in gender gap in educational mobility. However, with this data set, I will examine whether gender norms could affect the the gender gap in educational mobility. I specify the following regression,

$$Y_i^* = \beta_0 + \beta_1 MaleChild_i + \beta_2 AgeChild_i + \beta_3 \sum_{i=1}^{7} Bias_i + \beta_4 \sum_{i=1}^{7} Bias_i * MaleChild + \beta_5 TotalBias_i + \beta_6 TotalBias_i * MaleChild + U_i \tag{8}$$

With this specification I am interested to know whether the gender gap in educational mobility is higher for people with bias gender norms. I run the above regression separately for each bias (and total bias) gender norm as the gender norm biases are likely to be correlated. I will explain these bias gender norms in the result section.

I am also interested to know whether the gender gap in educational mobility is same across all section of the population. One way to classify the population would be based on gender norms they have. A latent class analysis would be used to determine the groups in the population based on the gender norms.

A latent class model is used when we believe that there are groups in a population and individuals in different groups behave differently. But we do not have a variable that can identify these groups. Latent class model allows us to identify such unobserved groups and who is likely to be in a particular group. We can fit a latent class model to determine which individual belongs to which group from the other variables we have. These other variables are the gender norms perception which will be used to determine which individual is likely to belong to which group, known as class. To fit latent class model, we need to provide how many classes are there in the population. To identify the number of classes, measures of goodness of fit such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are commonly used; the best-fitting model is the one with the smaller AIC and BIC.

4 Results

4.1 Results from IHDS-II Data Set

In this section, I present the result of my estimation. The table 3 below shows the result of linear probability model. As described before, I am interested in the sign of the coefficient of *MaleChild* and the coefficients of caste variables interacted with the *MaleChild*.

From the table 3 below, model 1 gives the results of the baseline regression. In model 1,we can see that the coefficient of *MaleChild* is positive and significant. This means that given father's education, a son is more likely to have higher education than a daughter compared to father's education. In other words, there is a gender gap in upward educational mobility, a son has higher probability than a daughter in upward educational mobility. This implies that intergenerational educational mobility is higher for a son compared to a daughter and hence there exists a gender gap in educational mobility. Also, we can see that the coefficient of child's age is negative and significant in model 1. This means that educational mobility (the likelihood that a child exceeds their father) is decreasing over time.

Now, we move to other critical aspect of this result. In model 2, the coefficient of the interaction term (FathersEducation*Malechild) is negative and significant meaning gender gap in educational mobility is smaller for more educated fathers. This implies that educational mobility is lower for more educated fathers.

In model 3, I added caste dummies and their interaction terms with *MaleChild*. The interaction terms of caste dummies and *MaleChild* will help us understanding how gender gap in educational mobility is spread across different castes. As you can see from model

Table 3: Gender gap in intergenerational educational mobility
Dependent Variable: Binary: Child's education > Father's education = 1, 0 otherwise

Variables	Model 1	Model 2	Model 3
Intercept	0.65^{***} (0.010)	0.74^{***} (0.012)	0.92^{***} (0.016)
Male Child	0.25^{***} (0.009)	0.28^{***} (0.009)	0.35^{***} (0.021)
Age of Child	-0.01^{***} (0.0002)	-0.01^{***} (0.0003)	-0.01^{***} (0.0003)
Father's years of education		-0.01^{***} (0.001)	-0.02^{***} (0.001)
Father's years of education*Male Child		-0.01^{***} (0.001)	-0.01^{***} (0.001)
Caste Forward			-0.01 (0.017)
Caste OBC			-0.12^{***} (0.017)
Caste Dalit			-0.21^{***} (0.017)
Caste Adivasi			-0.26^{***} (0.039)
Caste Muslim			-0.18^{***} (0.027)
Caste Sikh, Christian and Jain			0.14^{***} (0.004)
Male Child*Forward			-0.06^{***} (0.021)
Male Child*OBC			-0.04^{***} (0.018)
Male Child*Dalit			-0.05^{***} (0.018)
Male Child*Adivasi			-0.08^{***} (0.027)
Male Child*Muslim			-0.11^{***} (0.022)
Male Child*SikhChristianJain			-0.20^{***} (0.024)
Observations	88,690	88,690	88650
R-square	0.075	0.098	0.1311

Notes: Robust Std. error in parenthesis
***Significant at 1%
**Significant at 5%
*Significant at 10%

3, all the caste and *MaleChild* interaction terms are negative and significant. This means that compared to Brahmin, all other castes have smaller gender gap in educational mobility. Brahmin and Forward castes are considered upper caste in socio-economic status and gender gap in educational mobility is expected to be similar. However, I find that forward castes have a smaller gender gap in educational mobility (coefficient is negative and significant) compared to Brahmin. Now, let's look at the coefficient of the lower castes. The lower castes are OBC, Dalit and Adivasi. It is to be noted here that Indian constitution recognizes Dalit and Adivasi's as Schedule caste (SC) and Scheduled tribe (ST) respectively. Dalit and Adivasis are at the bottom end of the social status and it is expected that gender gap will be higher for this section of the population as they face many discrimination in social life. If we look at the coefficient (interaction of caste and *MaleChild*) of these castes, they are negative and significant meaning that the gender gap in educational mobility is smaller for these castes compared to Brahmin caste. Now let's look at the Muslim population. The educational attainment of Muslim in India is poor and are similar to those of the lower caste population. The coefficient of Muslim is also negative and significant. The Sikh, Christian and Jain have higher educational mobility in general and significantly smaller gender gap.(in terms of magnitude of the interaction coefficient). These people in general are at good position in socio-economic status, and the result suggests that the gender gap in this section of the population is the smallest across all castes.

4.2 Results from EAS Data Set

As mentioned before, the IHDS-II data set suffers from the potential missing observations. So, the above results from IHDS-II data must be backed up by some other information. The data set EAS does not suffer from any missing observations and hence it is used to verify the existence of a gender gap in intergenerational educational mobility.

First, I present the results of the baseline regression from EAT data set below to determine the existence of intergenerational educational mobility.

Table 4: Gender gap in intergenerational educational mobility
Dependent Variable: Binary: Child's education > Father's education = 1, 0 otherwise

Variables	Coefficient
Intercept	1.04^{***} (0.0644)
Male Child	0.05^{***} (0.0205)
Age of Child	-0.01^{***} (0.0015)
Observations	2,539
R-square	0.5463

Notes: Robust Std. error in parenthesis.
***Significant at 1%, **Significant at 5%,

The findings are similar to IHDS-II data set. Again, I have the coefficient of *MaleChild* positive and significant reiterating my previous finding of the existence of a gender gap in intergenerational educational mobility. Here, I have not added the caste effect here given this is a small sample and I do not have all the castes present in this data set. I also do not verify if the gender gap is less for more educated fathers given the small sample. I have added father fixed effect as different fathers might behave differently and clustered the standard error at each father.

Now, I discuss how gender norms affect the gender gap. I asked both the husband and the wife a series of questions regarding gender norms. I asked questions such as, who should earn money, who should have higher level of education, who should be responsible for washing, cleaning and cooking, who should fetch water in case of no water pump or tap, who should be responsible for feeding and bathing children, who should help the children in their studies at home, who should be responsible for looking after ill person in the family. Respondents were asked to choose either male, female or both option for each question.

I created 7 dummy male bias gender norm variables and and an overall male bias index. If the respondent answered "male" for question 1 and 2, I considered that person (and so the household) has male bias gender norms. For question 3 to 7, if the respondent answered "female", I considered that person (and so the household) has male bias gender norms. Finally, an index of total male bias gender norms was created which is the summation of all the 7 male bias gender norm dummies.

I am interested to know how these gender norms affect the gender gap in intergenerational educational mobility. For these I have to make a couple of assumptions. First, I consider data for the husband and the wife's siblings and their parents only, not the children of the husband and the wife. Second, I am assuming that the gender norms questions answered by the husband and the wife represents the gender norms of their respective household. In other words, husband's response represent the gender norms of the husband's family (husband's siblings and parents excluding the wife). Similarly, wife's response represent the gender norms of the wife's family (wife's siblings and parents excluding the husband). A descriptive statistics of the gender norm questions is given below.

Table 5: Summary Statistics of Gender Norms

Variables	N	Mean	SD	Min	Max
Earn money bias	2,231	0.49	0.5000	0	1
Higher Education bias	2,231	0.20	0.3983	0	1
Cooking bias	2,231	0.84	0.3700	0	1
Fetch water bias	2,231	0.26	0.4401	0	1
Feeding children bias	2,231	0.82	0.3848	0	1
Children studies	2,231	0.48	0.4997	0	1
Ill person bias	2,231	0.37	0.4829	0	1
Total bias	2,231	3.46	1.7369	0	7

From the table above, 49% of the respondents indicated that only men should earn money in the family. Similarly, I can interpret the other male bias gender norm questions. The highest male bias is for the question "Who should be responsible for washing, cleaning and cooking?", 84% responded that only women should be responsible for these work. Total male bias is the summation of all the 7 individual male bias questions.

Now, I am interested to know how these gender norm questions affect the gender gap in intergenerational educational mobility. I summarize the effects of gender norms in the following table.

The results are generated by adding the bias dummies and the interaction of these dummies with *MaleChild* as shown in equation (8) . In table 6, I am particularly interested in the column "Bias" vis-a-vis "Non-bias". These two column compare the gender gap in male bias families compared to non-bias families. For the first three male bias norms, namely, "earn money", "higher education" and "feeding children" biases, the magnitude of the gender gap is higher for families with these three biases. For the other four biases, namely, "cooking,

Table 6: Gender gap with non-biased and biased population

Variables	Non-bias	Bias	Difference
Bias:Earn money	0.047	0.082***	0.035
	(0.0298)	(0.0292)	(0.0417)
No. of observations	1,135	1,094	
Bias:Higher education	.055**	0.106*	0.051
	(0.0222)	(0.0551)	(0.0594)
No. of observations	1,789	440	
Bias:Cooking	0.112**	0.055**	-0.057
	(.0548)	(.02263)	(.0592)
No. of observations	365	1,864	
Bias: Fetch water	0.081***	0.021	-.059
	(.0239)	(.0432)	(.0493)
No. of observations	1,644	585	
Bias:Feeding children	0.061	0.066***	0.004
	(.0488)	(.0232)	(.0538)
No. of observations	403	1,826	
Bias:Children studies	0.076**	0.053*	-.024
	(.0305)	(.0287)	(.0418)
No. of observations	1,157	1,072	
Bias:Ill person	0.073***	0.051	-.023
	(.0268)	(.0337)	(.0430)
No. of observations	1,405	824	
Total Bias	0.080	0.076**	-.004
	(.0490)	(.0378)	(.0128)

Notes: Std. error in parenthesis
***,**,* Significant at 1%, 5%, 10% respectively

"fetch water", "children studies" and "ill person", the magnitude of the gender gap is lower for families with these three biases. This indicates that the gender gap in intergenerational mobility is higher in families having selected type of male bias gender norms while it is lower for other families having another set of male bias gender norms. In conclusion, we have insufficient evidence to conclude that the gender gap is higher for families having male bias gender norms.We can also examine the potential relationship between the gender gap in

educational mobility and gender norms from latent class analysis which is discussed below.

4.3 Latent Class Results

I found the evidence of a gender gap in educational mobility. This gender gap might exists only for specific section of the population, namely those who have bias gender norms. I use latent class model to identify the potential groups in the population based on the gender norm responses.

The latent class presented below has identified three distinct groups (class) in the population. Respondents in group 1 have "high bias" gender norms who represents 32% of the population. The second group who represents 50% of the population have "moderate bias" gender norms while the third group consisting of 16% of population have "low bias" gender norms. The results given in the following table is generated using equation (7) running separately for each group.

Table 7: Gender gap in intergenerational educational mobility
Dependent Variable: Binary: Child's education > Father's education = 1, 0 otherwise

Variables	Class 1 (High bias)	Class 2 (Moderate bias)	Class 3 (Low bias)
Intercept	1.09***	0.95***	1.07***
	(0.1059)	(0.0906)	(0.1348)
Male Child	0.02	0.06**	0.06
	(0.0360)	(0.0294)	(0.0486)
Age of Child	−0.01***	−0.01***	−0.01***
	(0.0024)	(0.0020)	(0.0032)
Observations	823	1,288	428
R-square	0.6014	0.5426	0.4834

Notes: Robust Std. error in parenthesis
***Significant at 1%
**Significant at 5%
*Significant at 10%

As before, I am interested in the coefficient of *MaleChild* which will determine the ex-

istence of a gender gap in educational mobility. From the table above, only for group 2 of the population who have moderate bias gender norms, the coefficient is positive and significant implying the existence of a gender gap in educational mobility. However, we do not find evidence of a gender gap in educational mobility for those who have "high bias" (group 1) or "low bias" (group 3) gender norms. Hence, gender gap in educational mobility exists only for a certain section of the population, namely, those who have "moderate bias" gender norms. In conclusion, there is not enough evidence to conclude that gender gap in educational mobility is higher for families who have more bias gender norms.

5 Robustness

In this section I describe the robustness check of my results. First, I discuss the standard IER results for IHDS-II data set. This result is presented at table 8 in the appendix. As discussed in the empirical framework, I am interested in the coefficient of the interaction between *MaleChild* and *FathersEducation*. A negative sign on the coefficient of this interaction term will indicate the existence of a gender gap in educational mobility. As you can see from the table 8, the coefficient of the interaction term MaleChild*FathersEducation (β_3) is negative and significant. This establishes and reiterates my previous finding of the existence of a gender gap in educational mobility.

I used linear probability model for my analysis. Since the dependent variable is a binary variable, I also estimate the probit model (did not use caste dummies here) and the results are very similar to the linear probability model. As in the linear probability model, the coefficient

of *MaleChild* is positive and significant, reiterating my earlier finding of the existence of a gender gap in intergenerational educational mobility. Also, the coefficient of the interaction term *MaleChild*FathersEducation* is negative and significant as before impying that the gender gap is less for more educated fathers. Hence, the conclusions I derived in the previous section using linear probability model remain valid. This robustness check is conducted for IHDS-II data set. This result is shown at table 9 in the appendix.

6 Conclusion

This paper aims to understand whether there exists a gender gap in intergenerational educational mobility in India. For the analysis, I used two data sets namely IHDS-II and EAS. Using the India Human Development Survey-II 2011-12, a nationally representative household survey, I find evidence of a gender gap in intergenerational educational mobility. Educational mobility is higher for a son compared to a daughter. This means that given father's education, a son is more likely to obtain higher education than a daughter. I also find that the gender gap is smaller for all castes in comparison to Brahmin caste. The smallest gender gap in educational mobility exists for the religious minorities, Christian, Sikh and Jain (in terms of magnitude of gender gap). My findings also confirm that the gender gap is lower for more educated fathers.

The above findings are based IHDS-II data set which suffers from potential missing observations in terms educational attainment of siblings of both the husband and the wife. To rectify that I used EAS data set, which does not suffer from missing observations to verify

the existence of a gender gap in intergenerational educational mobility. However, EAS data set is not a nationally representative data set like IHDS-II. I collected data from a specific region with a much smaller sample compared to IHDS-II. Based on EAS data set also, I find evidence of a gender gap in intergenerational educational mobility. However, such gap exists only for certain section of the population, namely, those who have "moderate bias" gender norms.

My finding of the existence of a gender gap in intergenerational educational mobility suggest that girls do not have equal opportunities in terms of access to education in comparison to boys. The policy makers have a significant role in ensuring that everyone irrespective of gender or caste has equal access to education which will go a long way in achieving a just society. Government policy should be initiated which will incentivize girls for higher education and thereby reducing the gender gap in terms of intergenrational educational mobility. Also, policy makers need to undertake affirmative actions for specific section of the population to reduce the caste-wise gender gap in educational mobility. For future research, a larger representative sample which collects the educational attainment of siblings of both the husband and the wife would enhance our understanding on the gender gap in intergenerational educational mobility and the underlying determinants of such gender gap.

References

Anitha, B.K. 2000. Village, caste and education. New Delhi: Rawat Publications.

Asher, S, Novasad, P and Rafkin, C, (2020), "Intergenerational mobility in India: New Methods and Estimates Across Time, Space and Communities." World Bank working paper.

Azam, M. (2016). Intergenerational Educational Persistence among Daughters: Evidence from India. The B.E. Journal of Economic Analysis & Policy, 16(4):1–16.

Azam, Mehtabul and Vipul Bhatt, " Father, Like Son? Intergenerational Educational Mobility in India," Demography, 2015, 52 (6).

Black, S., and P. Devereux. (2011). "Recent Developments in Intergenerational Mobility." Handbook of Labor Economics, 4(B), Ch-16, 1487- 1541.

Blanden, J. 2013. "Cross-Country Rankings in Intergenerational Mobility: A Comparison of Approaches from Economics and Sociology." Journal of Economic Surveys 27 (1): 38–73.

Chadwick, L., and G. Solon. 2002. "Intergenerational Income Mobility Among Daughters." American Economic Review 92 (1): 335–344.

Desai, S., & Kulkarni, V. (2008). Changing educational inequalities in India in the context of affirmative action. Demography, 45(2), 245-270.

Dhar D, Jain T, Jayachandran S. 2019. The intergenerational transmission of gender attitudes: evidence from India. The Journal of Development Studies. Vol. 55, No. 12, 2572–2592, https://doi.org/10.1080/00220388.2018.1520214

DiPrete, T. A., and D. B. Grusky. 1990. "Recent Trends in the Process of Stratification." Demography 27 (4): 617–637.

Duflo, E. (2012). Women Empowerment and Economic Development. Journal of Eco-

nomic Literature, 50(4):1051–1079.

Emran, M. S. and Shilpi, F. (2015). Gender, Geography, and Generations: Intergenerational Educational Mobility in Post-Reform India. World Development, 72:362–380.

Emran, M. S., Greene, W., and Shilpi, F. (2018). When Measure Matters: Coresidency, Truncation Bias, and Intergenerational Mobility in Developing Countries. Journal of Human Resources, 53(3):589–607.

Hertz, Tom, Jayasunderay, Tamara, Pirainoz, Patrizio, Selcuk, Sibel, Smithyy, Nicole, and Verashchagina, Alina. 2007. "The Inheritance of Educational Inequality: Interna- tional Comparisons and Fifty-Year Trends." The B.E. Journal of Economic Analysis and Policy (Advances), 7 (2), 1-46.

Hnatkovskay, V., A. Lahiri, and S. B. Paul. 2013. "Breaking the Caste Barrier: Intergenerational Mobility in India." Journal of Human Resources 48 (2): 435–473.

Jalan, J. and Murgai, R. (2008), "Intergenerational Mobility in Education in India," Paper Presented at the Indian Statistical Institute, Delhi.

Jayachandran, S. (2015). The Roots of Gender Inequality in Developing Countries. Annual Review of Economics, 7(1):63–88.

Maitra,P.,andA.Sharma.(2009)."ParentsandChildren:EducationAcrossGenerationsinIndia.Paperpresent on Economic Growth and Development. Delhi: Indian Statistical Institute.

Olivetti, C., and M. D. Paserman. 2015. "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850–1940." American Economic Review 105 (8): 2695–724.

Psacharopoulos, G., & Tzannatos, Z. (1989). Female labor force participation: An inter-

national perspective. World Bank Research Observer, 4(2), 187–201.

Seguino, S. (2011). Help or hindrance? Religion's impact on gender inequality in attitudes and outcomes. World Development, 39(8), 1308–1321

Solon, G. (1999). Intergenerational mobility in the labor market. In Ashenfelter, O. and Card, D., editors, Handbook of Labor Economics, volume 3 of Handbook of Labor Economics, pages 1761–1800. Elsevier.

Stiglitz, J. (2012). The Price of Inequality. W. W. Norton Company.

Torche, F. (2015b). Intergenerational Mobility and Gender in Mexico. Social Forces, 94:5633–587.

Appendix

Table 8: Gender gap in intergeneratonal educational mobility (Data set: IHDS-II)
Dependent Variable: Child's education

Variables	Coefficient
Intercept	12.68^{***} (0.2652)
Male Child	2.30^{***} (0.2653)
Father's years of Education	-0.21^{***} (0.0360)
Father's years of Education*Male Child	-0.14^{***} (0.0340)
Age of Child	-0.17^{***} (0.0037)
Observations	88,734
R-square	0.9269

Notes: Robust Std. error in parenthesis.
***Significant at 1%

Table 9: Probit Model(Data set: IHDS-II)
Dependent Variable: Binary: Child's education > Father's education = 1, 0 otherwise

Variables	Coefficient
Intercept	0.64^{***} (0.0162)
Male Child	0.74^{***} (0.0115)
Father's years of Education	-0.03^{***} (0.0015)
Father's years of Education*Male Child	-0.02^{***} (0.0020)
Age of Child	-0.02^{***} (0.0004)
Observations	88,734

Notes: Std. error in parenthesis. ***Significant at 1%