Correlates and Determinants of Intergenerational Mobility in India: An Econometric Analysis of Cross-Section and Panel Data

Thesis Submitted to the Jawaharlal Nehru University
in Fulfilment of the Requirements for the
Award of the Degree of

DOCTOR OF PHILOSOPHY

by

Swati Sharma

Centre for the Study of Regional Development

School of Social Sciences

Jawaharlal Nehru University

New Delhi-110067

2021

RECOMMENDATION FORM FOR EVALUATION BY THE EXAMINER/S

CERTIFICATE

This is to certify that the dissertation/thesis titled Correlates and Determinants of Intergenerational Mobility in India: An Econometric Analysis of Cross-Section and Panel Data ……………………………………………………………………………..submitted by Mr/Ms……..Swati Sharma…………………..in partial fulfillment of the requirements for award of degree of M.Phil/M.Tech/Ph.D of Jawaharlal Nehru University, New Delhi, has not been previously submitted in part or in full for any other degree of this university or any other university/institution.

We recommend this thesis/dissertation be placed before the examiners for evaluation for the award of the degree of M.Phil/M.Tech./Ph.D.

Signature of Supervisor

Date: 25-07-2021

Centre for the Study of Reg Dev.
School of Social Sciences
Jawaharlal Nehru University
New Delhi - 110067

Signature of Dean/Chairperson

Date: 29.7.2021

Chairperson
Centre for the Study of Reg Dev.
School of Social Sciences
Jawaharlal Nehru University
New Delhi - 110067

DEDICATED TO MY PARENTS

Acknowledgements

First, I would like to express my heartfelt gratitude to my supervisor Professor Amaresh Dubey for his unwavering perseverance, motivation, enthusiasm, and insightful comments. His guidance was invaluable for the research presented in this thesis.

Besides my supervisor, I would like to thank the staff of JNU Central Library, EXIM Library and Documentation Centre of CSRD for their full support and cooperation in allowing me to use their libraries. I consider myself lucky to have fellow batch mates in CSRD, who were always there to help me.

I express my heartfelt gratitude to Prateek Sharma, my husband, who with his loving and caring attitude has been a steady support especially throughout my periods of stress and strain and has always boosted my morale and confidence. I will always be indebted to my family: my beloved parents for showing confidence in my efforts, and for their patience and support. This journey would not have come to a fruitful end without the emotional support of my dear mother. The acknowledgement would be incomplete without mentioning my son, Vivaan (Duggu), who directly or indirectly made genuine efforts to make me become a more patient person.

My elder brother, Sandeep Sharma has been there as a source of love, concern, encouragement and strength in all these years. He has been my inspiration and motivation for continuing to expand my knowledge and move my career forward. He is my rock.

Most importantly, I thank God for the wisdom and perseverance that has been bestowed upon me during the last few years, and indeed, throughout my life.

Swati Sharma

Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

2SLS	Two-Stage Least Square
AP	Andhra Pradesh
AS	Assam
ATE	Average Treatment Effect
ATT	Average Treatment on Treated
BI	Bihar
BRIC	Brazil, Russia, India, China and South Africa
DL	Delhi
EUS	Employment and Unemployment Surveys
GDP	Gross Domestic Product
GMM	Generalized Method of Moments
GSDP	Gross Domestic State Product
GU	Gujarat
HR	Haryana
HP	Himachal Pradesh
HRD	Human Resource Development
ICC	Intergenerational correlation coefficient
IEM	Intergenerational Educational Mobility
IHDS-I	Indian Human Development Survey, 2005-05
IHDS-II	Indian Human Development Survey, 2011-12
IRC	Intergenerational regression coefficient
IV	Instrumental Variable
KA	Karnataka
KE	Kerala
LM	Langrange-Multiplier
MH	Maharashtra
MHSS	Matlab Health and Socioeconomic Survey
MP	Madhya Pradesh
NCAER	National Council of Applied Economic Research
NFHS-I	National Family Health Survey,1992-93
NFHS-II	National Family Health Survey,1998-99

NSS	National Sample Surveys
OBC	Other Backward Caste
OECD	Organization for Economic Cooperation and Development
OLS	Ordinary Least Square
OR	Orrisa
PSID	Panel Study of Income Dynamic
PSM	Propensity score matching
PU	Punjab
RBI	Reserve Bank of India
REDS	Rural Economic and Demographic Survey
RJ	Rajasthan
SC	Scheduled Caste
SGAP	Schooling Gap
SSA	Sarva Shiksha Abhiyan
ST	Scheduled Tribe
TN	Tamil Nadu
UP	Uttar Pradesh
UPSS	Usual Principal Subsidiary Status
UTs	Union Territories
WB	West Bengal

Chapter 1

Introduction

1.1 Background

India has achieved a remarkable improvement in overall literacy rate since its independence in 1947. The 1951 census of India found that only 27% Indian men and 9% Indian women were literate. Over the next six decades, the literacy rate in India has steadily improved to 74%. Lall (2005) notes that while enrolment at primary level has improved over time in India, nonetheless, he estimated that up to 60 million children aging between 6-14 years did not attend school. The study highlights the various problems of Indian education system such as poor and inadequate infrastructure, high teacher absenteeism, and insufficient funds, among others. Moreover, there is substantial educational inequality based on socioeconomic origins. For instance, the census of 2011 shows that the female literacy rate is 65% whereas the male literacy rate is 82%. Kingdon (2007) highlights the problem of high dropout rates in the Indian education system. The study observed that 93% of the children aged between 6–14 years were enrolled in primary schools. However, more than 20% of the children aged between 15 – 16 years were not attending school either because they never enrolled or because they dropped out after primary school. In terms of enrolment to secondary education, India fares the worst among the BRIC nations.

A plausible reason for low level of secondary education attainment could be the limitations of the labour market. If the labour market does not adequately reward incremental educational qualifications, there would be little incentive to progress to secondary or higher levels of education. Most of the prior studies on Indian labour market found positive and significant returns to education (Duraisamy, 2002; Kingdon & Unni, 2001; Tilak, 2007). Nonetheless, Indian labour market is characterized by widespread disparities in distribution of education and as well as the returns to education. For example, Kingdon (1998) found that female workers receive lower returns to education than male workers.

Duraisamy (2002) documents significant gender and location based differences in the returns to schooling. Several other studies relate education inequalities and labour market outcomes in India (Bhaumik & Chakrabarty, 2009; V. Borooah, 2010; Sundaram & Vanneman, 2008; Unni, 2010). The uneven distribution of education and its associated labour market outcomes contribute to the urban-rural gap in income. The peculiar nature of Indian demographics motivates researchers to examine the distribution of educational opportunities across different religion, caste, region, and gender-based groups. One of main challenges for the Indian policy makers is to formulate policies that can mitigate inequality of education opportunity based on socioeconomic origins.

A number of arguments can be adduced to explain intergenerational educational mobility (IEM). First, parents with high education have a tendency to earn more as compared to their low-educated counterparts, and therefore, they can provide better education to their children. Second, ceteris paribus, parents with high education have better unobserved abilities than their less educated counterparts. The inheritance of such unobserved abilities affects child's educational outcomes. Third, parental education also affects allocation of parental time and resources while raising their offspring (Craig, 2006; Guryan et al., 2008) which in turn determines educational outcome for their children. Fourth, education affects the bargaining power of an individual. For instance, educated mothers are more capable of channelling the household expenditure towards better development of their children (Baker & Stevenson, 1986; Currie & Moretti, 2003; Ware, 1984).

A substantial body of literature has examined IEM in different countries (Azam & Bhatt, 2015a; Emran & Shilpi, 2015a; Lillard & Willis, 1994; Mare, 1997). In the Indian context, most of the studies have estimated the intergenerational educational coefficient relying on a simple bivariate regression, where the child's education is regressed on the parents' education. However, this approach does not account for the temporal variations in the educational distributions across generations. In addition, the conventional regression scheme is prone to endogeneity bias. This is because several unobserved characteristics may be associated with both child's and parents' education. For instance, highly educated parents may have better skills and abilities that are pass on to their offspring, which affect their educational outcomes. Similarly, parental education may also be an important factor which can affect decision regarding how they choose to allocate their time and financial resources in raising their children. These parental choices influence child's attainment. Most of the previous studies do not account for the potential endogeneity problem due to the lack valid instrumental variables that satisfy the exclusion restriction. Additionally, although there is an extensive literature on the determinants of IEM, most of the analyses is limited to developed countries. This study examines the correlates and determinants of IEM in India,which still remains an under-researched topic . Thus, there are good reasons to study the determinants of IEM in India.

1.2 Rationale to Study Intergenerational Educational Mobility

Mobility refers to the changes in the social position of individuals or a group of individuals. The social position may be defined in terms of education, occupation, income, wealth or social class (Atkinson, 1981; Becker & Tomes, 1986; Beller, 2009; Beller & Hout, 2006; Blane et al., 1999; Platt, 2005; Solon, 1992a; Treiman & Ganzeboom, 1990; Zimmerman, 1992). IEM is a change in socioeconomic status across generations. These measures of mobility are generally estimated in the context of educational attainment, occupational status or earnings. IEM reflects the overall social welfare attained via allocation of resources across generations (Atkinson, 1981). It may be argued that our present actions represent the accumulated deeds of the past generations and our concern for future generations: past generations, as their legacy leave an impact on our current decision making, and upcoming generations, because our present actions influence the well-being of the future generations through various endowments offered to them as inheritance.

This study examines the IEM, which refers to the change in the level of educational attainment of an individual with respect to their parental education level. Examining social mobility in terms of education attainment rather than income is preferable as it obviates several estimation issues associated with the latter approach. First, any measure of income mobility is likely to be affected by life cycle bias, wherein different individuals realize their peak income at different ages across generations. Moreover, since people are likely to complete their education in their mid-twenties; therefore measuring educational mobility is likely to circumvent the measurement problems emanating from life-cycle bias (Black & Devereux, 2011). Second, unlike income mobility, educational mobility is less sensitive to measurement issues, as individuals generally tend to be more forthcoming in revealing their level of education as opposed to revealing their exact income. Due to these inherent advantages, even the studies on income mobility often use education to estimate imputed earnings (Björklund & Jäntti, 1997; Causa & Johansson, 2010; Dearden et al., 1997).

1.3 Research Objectives and Hypotheses

We examine the correlates and determinants of IEM in India. In this study, we have used different approaches to measure educational mobility across generations. This study relies on the information mainly from Indian Human Development Surveys (IHDS) household panel data which offers a unique opportunity to isolate the factors affecting IEM. However, we have also used data provided by the National Sample Surveys (NSS), to identify the macroeconomic correlates of IEM in India.

1.3.1 Research Objectives

Going through the brief background on IEM, we identify the following research objectives:

(i) To analyse trends in and growth of IEM.

(ii) To study the role of migration status on the degree of educational mobility across generations.

(iii) To investigate the impact of return migration in affecting IEM.

(iv) To evaluate the impact of aid programmes on schooling progression of the child and educational mobility across generations.

(v) To identify the possible reasons for interstate variations in IEM by investigating the association between different macroeconomic policy variables and IEM.

1.3.2 Hypotheses

This study is an empirical investigation of IEM in India. Therefore, we have the following null hypotheses which will be tested to evaluate our research objectives.

H1: There is no considerable interstate variations in IEM.

H2: Migration status does not affect the degree of IEM.

H3: Return migration has no impact on IEM.

H4: The impact of aid (financial or non-financial) on child's education does not vary by parental education.

H5: State level policies have no significant impact on interstate variations in IEM.

1.4 Uniqueness of the Present Study

There are many studies that focuses on the determinants of IEM. We try to offer robust estimates of IEM by applying different econometric strategies. The study uses the novel approach introduced by Lewbel (2012) to address the issue of endogeneity. We also employ semiparametric approach, like matching estimators to offer robust estimates of IEM. The study also aims to analyse interstate variations in IEM.

Different studies identify different factors that determine educational mobility across generations. Recent work on IEM have tried to investigate the impact of immigrants on IEM. We extend previous work by examining the role of internal migration on IEM. Further, we examine the underlying factors causing variations in IEM by exploiting information available from different rounds of NSS. In short, it is one of the earliest works to identify the factors affecting the IEM in India. The comprehensive dataset obtained from IHDS and NSS allows us to examine IEM over a period of five decades (1947–1996).

1.5 Organization of the Thesis

This thesis comprises of nine chapters. Chapter 1 presents the concept of IEM and discusses the importance of studying IEM, along with the research objectives and the uniqueness of the present study. Chapter 2 presents a literature review. Chapter 3 describes the data sources and empirical methodology. Chapter 4 discusses the trends in IEM in India, both at state level and all-India level. Inter-state disparities have also been analysed. Chapter 5 examines the impact of migration status on IEM. This chapter exploits the information available in IHDS data to identify families who migrated to other areas while their children were in school. Chapter 6 analyses the effect of return migration on the schooling progression of the children. Educational transition models reveal important dynamics of educational mobility. Therefore, in this chapter school progression has been used as an indicator of child’s educational progress. Chapter 7 examines the impact of aid/grants on IEM. In this chapter, child’s educational progress has been measured by estimating the gap in the schooling. Chapter 8 explores the determinants of IEM in India by using both parametric and nonparametric models. Chapter 9 presents the overall findings while discussing some policy implications.

Chapter 2

Survey of Literature

2.1 Introduction

This chapter provides a detailed review of literature on IEM. It will help in critically examining the theoretical as well as empirical studies investigating the various aspects of IEM. The literature on trends in IEM and its relationship with other economic variables has been reviewed in this chapter. This review begins by discussing the theoretical framework explaining the concepts regarding educational mobility.

It is well known that education is a key determinant of development and economic growth (for example, Becker et al.,1990 ; Lucas, 1988). Most of the decisions regarding children are made by their parents. Therefore, it is assumed that household structure holds a central position in affecting the investment in human capital (Becker, 1981; Becker et al., 1990). Family acts as a social institution which shapes the economic outcomes of its members by forging an intergenerational link between consecutive generations. The strength of these intergenerational links has been measured by many prior studies that have studied the intergenerational relation between education and income. This chapter surveys the literature that estimates the association between the various household characteristics and the educational outcome of the offspring, aiming to establish the intergenerational links.

The rest of the chapter is structured as follows. Section 2.2 starts with a detailed outline of the conceptual framework to provide a foundation for our empirical study. Section 2.3 discusses the relevant prior literature on IEM, with the aim of highlighting the key empirical findings. Next section discusses the literature on factors determining IEM. Hence, Section 2.4 comprises of two sub-sections. Sub- section 2.4.1 summarizes the literature on the relationship between migration background and IEM. In sub-section 2.4.2, we summarize the relevant literature analysing some other important determinants of IEM, except the migration background. We end this chapter by summarizing the arguments that shape our research questions of this thesis to which we seek answers in the next few chapters.

2.2 Concept of Intergenerational Educational Mobility

A low degree of IEM indicates that children's socioeconomic performance is highly dependent on his/her family background and a high level of IEM suggests the opposite. Our study focuses on the various dimensions of IEM. The term 'IEM' refers to the changes in the social position in terms of education attainment across successive generations within the same family. While there are many studies that have focused on income mobility across generations, but majority of these studies are limited to developed countries. In case of developing countries like India, lack of reliable long-term data on income makes it problematic to measure intergenerational income mobility. Although there is lack of studies measuring income mobility in developing countries, study on educational mobility can offer valuable information. Therefore, studying IEM could allow us to assess the effectiveness of educational policies to address mobility of income across generations.

2.3 Intergenerational Educational Mobility: A Theoretical Framework

It is well known to researchers that literature on human capital dates back to the late 1960s. There are many models that explain how human capital investments are done within families (Becker 1975, 1991; Behrman et al., 1982,1995; Mulligan 1997). The founding article by Becker (1962) on determinants of human capital investment proposes the idea that investment in human capital is done up to a point where the private marginal benefit received from the investment equals the private marginal expenditure of the investment.

Theoretical literature on the association between parental investment and children's outcomes suggests that there are many ways in which family background can influence the educational attainment of the children. Early theoretical work by Becker (1991) and Becker and Tomes (1976, 1979, 1984) assumed that parental investment on their children is guided by the altruistic behaviour of parents towards their children. In a perfectly competitive market, where individuals borrow and give money at the similar interest rate, the investment decision of the parents is assumed to be dependent on two conditions: (i) first, investing in child's education until a point where marginal return received from the investment equals to the interest rate which results in maximising the household wealth and (ii) to redistribute the maximized wealth among the members of the entire family in such a manner that household head's altruistic preferences get maximized. This redistribution can take the form of transfer across generations in either direction, from older generation to young generation or vice versa, depending on the family background. For example, children belonging to wealthy families may receive bequests or gifts from their parents, irrespective of the amount of investment done on their child's human capital. On the other hand, poor parents may demand partial or full repayment of the amount invested on their child's human capital, which the child is liable to pay in the future either as direct repayment (where parents may transfer the debt taken for investment on child's education) or indirect repayment (where child repays the debt in the form of old age transfer or gifts to their parents).

Irrespective of all the above-mentioned scenarios, in the presence of perfectly competitive capital market coupled with the altruistic behaviour of the parents, optimal investment is independent of the structure of family.

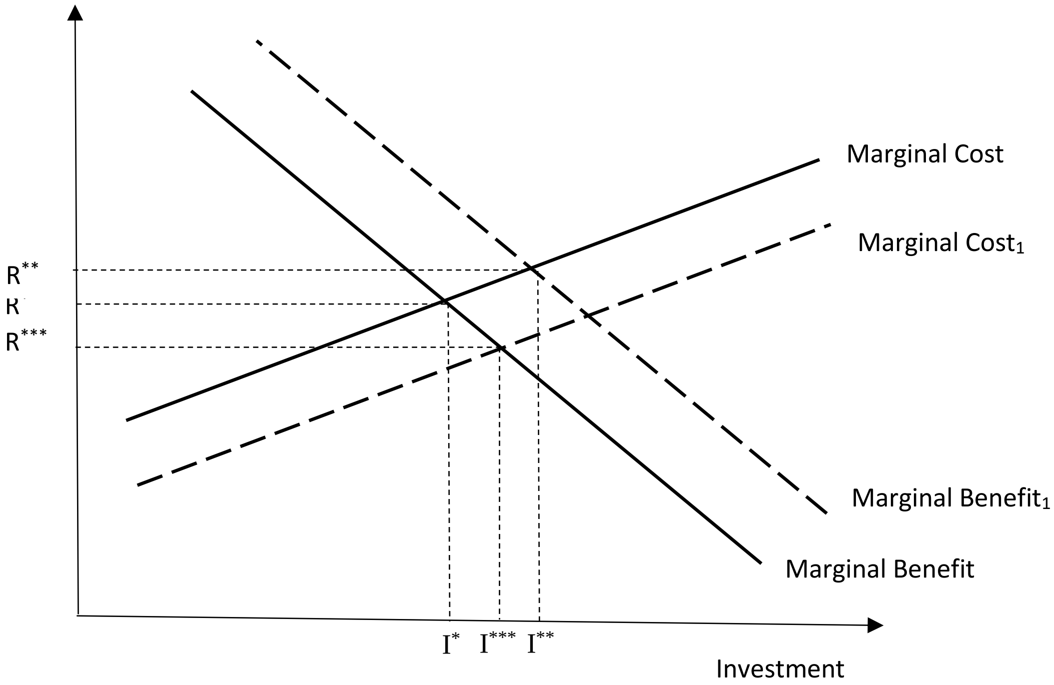


Figure 2.1: Private Marginal Cost (MC) and Marginal Benefit (MB) on Human Capital Investment

Figure 2.1 illustrates the optimal human capital investment as a function of private marginal cost and private marginal benefit of investment. The marginal benefit curve which shows the expected gain from investment in human capital (i.e., wages and salary in labour market) is downward sloping due to diminishing returns to investment. It is expected that the return to education is negatively correlated to the years of schooling. This means that private gains in the form of wage or salary for every additional year of education will be higher for low-educated individuals than for high-educated individuals. The reason could be partly attributed to fixed genetic endowments like, innate ability and motivation. Secondly, It is expected that higher schooling takes time and therefore delays one's entry into the labour market. This curbs the time in which an individual gains benefit, after the investment is being done. On the other hand, the private marginal cost curve is upward sloping showing the greater opportunity costs associated with higher investment in human capital and increasing private costs to finance higher education.

In a perfectly competitive market, an individual will invest education up to a level where the expected rate of return received from the investment in schooling equals the expected rate of return from the alternative investment option, regardless of family background. Therefore, I* represents the optimal investment level, where the association between schooling and family background is almost non-existent and schooling is considered only an investment with no consumption gain/loss attached with it. However, in real world, association between family background and child's schooling is highly plausible, even if schooling is considered solely an investment. Therefore, in the presence of market imperfections, the ideal investment level on schooling will rely on the family background of the individual.

One of the important reasons of variation in human capital investment on the child is the differences in returns to education. The latter can be attributed to individual-specific factors like ability of an individual or family-specific factors like genetics and neighbourhood effect. Secondly, changes in labour market may affect the investment behaviour of the parents. For example, a sudden increase in educated labour demand during the process of economic development may cause variations in the investment pattern of the parents. Labour market discrimination based on race, ethnicity and sex could be another reason that can affect family decision regarding education. Here, it is important to see that the direction of change in the investment due to labour market discrimination is uncertain. For example, if discrimination is higher in lower-skilled jobs when compared to higher-skilled jobs, the marginal return to education from higher education is greater for those who face discrimination. Therefore, this type of discrimination will result in increase in human capital investment.

To illustrate further how market imperfections render the socially optimal level of investment on human capital unattainable, we consider a higher -income household with highly educated parents. For this household, we could expect lower marginal private cost (dashed marginal cost curve below the solid marginal cost curve in Figure 2.1). Therefore, for these households I*** (see Figure 2.1) represents the optimal level of investment on schooling.

For lower-income households, we could expect higher marginal private cost of the investment on schooling. But, suppose the government offers financial aid to reduce the extra cost of education by exempting these households from payment of fees or by providing them with free uniforms, books, etc., then I** represents the optimal level of private investment on schooling. However, it is also possible, for marginal private benefits to be higher poor households with low educated parents. For example, many government policies are pro-poor which target lower-income households to reduce poverty and inequality by offering grants or fellowship or by allocating additional government funds on elementary education. Therefore, it becomes imperative to look at how the link between family background and schooling of the child is affected by the economic environment.

The initial theories of IEM were proposed in the humanistic sociology literature through the 1960s (Aldous & Hill, 1965; O'Donovan, 1962; Spady, 1967). These early models of social mobility were further refined by incorporating various psychological and socio psychological variables by Sewell and Hauser (1975) and later by Jencks et al. (1983). Their model, the Wisconsin model, identifies the various factors that determine educational attainment, earnings and occupational status of young adults. This "Wisconsin model" postulates that household characteristics affect the educational and occupational outcomes of the children, and these effects are mediated by psychological mechanisms, such as the expectations or occupational aspirations of their parents or spouses. Similarly, Becker and Tomes (1986) argue that the consumption patterns of parents affect the legacy they offer to their future generations. For example, high lifetime income of parents allows for better investment in the future of the upcoming generations. Further, the children born to rich parents are prone to have some innate as well as acquired characteristics that help to move up the social ladder, such as high motivation level and better allocation of material and time resources. If equal educational opportunities are available along with equal returns to a given education level, then education will be a sensible investment of one's time and resources. In such an environment, education should play a vital role in lowering persistence across generations.

2.4 Review of Literature

2.4.1 Measuring Mobility

The correlation between education level of parents and their descendants is found to be highly correlated in most of the countries (Hertz et al., 2007). Many household level characteristics like household income, parental wealth, parenting skills, genetics and attendance are found to be determinants of child's education level (Bowles & Gintis, 2002; Nielsen, 2006), but still education level of parent remains one of the most significant predictor of child's education level. Therefore, considerable research effort has been directed towards studying IEM. Using socioeconomic panel data from Germany, Heineck and Riphahn (2009) measure IEM in Germany for a period of five decades. They find that despite significant policy interventions and education reforms during their sample period, parental education continued to exert a strong effect on child's parental outcomes. For instance, they found that the probability of obtaining higher education qualifications conditional on parents who have completed only basic education is 12% and 9% for males and females, respectively. However, the study is restricted as it focused on the attainment of secondary education by the child. Using an international sample of 42 countries, Hertz et al. (2007) analyze the trends in IEM by ranking these countries in terms of IEM coefficients. They find large geographical differences in IEM, wherein the Latin American countries exhibit the lowest mobility whereas the Nordic countries display highest mobility. The global average of the correlation between parent and child schooling was estimated at 0.420 over a period of 50 years. They also show that the regression coefficient has declined noticeably over the period, while the correlation coefficient remained steady. They attribute this difference to the increase in standard deviation of parental education over time, which surpassed the increase in standard deviation of the education level of the child. Daude (2011) measures the IEM for 14 Latin American countries and found high persistence in education attainment across generations. Tverborgvik et al. (2013) use logistic regression to examine the impact of parental education level on the odds of receiving basic education by the child. They results show that children born to highly educated parents are three times more likely to get basic education as compared to the children born to low educated parents. Black and Devereux (2011) provide a comprehensive review of the latest works on IEM by examining the causal effect of parental education and family background on intergenerational education attainment.

While, most of the aforementioned studies report a high correlation between education level of children and that of their parents, various factors may affect the IEM. For instance, the gender of the parent and the child have considerable impact on the IEM. Farre and Vella (2013) found that child's education level is influenced by the education level of their same-gender parent. In other words, son's education level is likely to be correlated to the education level of their fathers, while education level of daughters are likely to be correlated to the education level of their mothers. This could be attributed to gender roles that are typically linked with the biological sex of an individual.

Unfortunately, most of these studies suffer from endogeneity problem that makes it difficult to distinguish between the nature and the nurture effects. This means, if children born to highly educated parents tend to be highly educated themselves, is it because of the genetic traits passed from parents to their children or is it because of a better learning environment provided by more educated parents. Some studies attempt to account for the endogeneity problem (see, for example, Checchi and Flabbi, 2007; Schnepf, 2002). These studies have used three different methods to dodge endogeneity related issues: adopted children, instrumental method and twins approach . The adoptee approach works well in isolating genetic transmission from parents to their biological offspring. However, this strategy has been criticized on the grounds of non-random assignment of adoptees to their parents. The second approach relying on identical twin parents differentiate the fixed effect of family as well as transmission of inherited abilities. However, this approach still has several limitations. First, even identical twins may suffer from unobserved heterogeneity that can affect the educational outcomes of their children (Bound & Solon, 1999; Griliches, 1979). Second, causal impact of twin parents on child outcomes may differ as they are married to different spouses, thereby not accounting for the impact of assortative mating (non-random matching of pairs based on observable characteristics). Finally, the instrumental variable approach relies on variables satisfying the exclusion restriction. It has been argued that the estimates obtained using this approach are generally not valid as they are likely to suffer from the problem of weak instruments (Holmlund et al., 2011b)

Studies on IEM in India has picked up pace in the last few years. Jalan and Murgai (2008) made one of the earliest attempts to investigate the degree of IEM in India. Using two rounds of the National Family Health Survey (NFHS): NFHS-I (1992-93) and NFHS-II (1998-99), their study examined IEM among individuals aged between 15 and 19 years. They found that IEM had increased over time, irrespective of social classes based on caste and wealth. They attempt to control for endogeneity of parental education level by using the extent of prenatal care received by cohort of mothers in 1992-93 as a proxy. However, there are some limitations associated with this study: First, their analysis is restricted to young children aged between 15-19 years, in order to dodge sample-selectivity issues caused by female children moving out of their birth family for marriage. Second, due to the limitations of their dataset, they are unable to find an exact measure of the prenatal care received by the mothers of the current 15-19 year old, when their mothers were pregnant with them.

Maitra and Sharma (2009) investigate the intergenerational changes in educational attainment of individuals aged 20 years and above. Using data from IHDS-I, they find that IEM has increased over time. This study also accounts for the potential endogeneity of parental educational attainment by using parental birth year and their original residing place as instrument variables for parental education level. Interestingly, their results show an insignificant causal impact of parental education on their child's education. In addition, they also analyze the impact of parental educational attainment on school progression by using a sample of children aging between 15 to 24 years. They find that father's education level has a significantly positive effect on the likelihood of continuing higher education. The main drawback of this study is that validity of instruments is checked only for overidentifying restrictions, i.e., they only test whether their instruments are valid. They do not test for the underidentifying restrictions, i.e., whether the instruments are relevant. Weak instruments can lead to large bias in the estimated coefficients. Using data from five rounds of the NSS, Hnatkovska et al. (2013a) investigate mobility across generations during the time period 1983 to 2005. They focus on exploring the variations in mobility patterns between the marginalized and non-marginalized sections of the society. Their findings show a significant increase in IEM among the disadvantaged section of the society, thereby moving towards their socially well-off counterparts. This analysis suffers from two issues. First, the intergenerational relationships are studied only for those parent-children pairs that are co-resident in the same household. Therefore, the estimated coefficients of IEM suffer from potential sample selection bias. Second, unlike the IHDS dataset, the NSS data does not offer longitudinal information. This means that long term intergenerational comparisons are problematic as the authors only have one point in time observation for each parent-child pair.

Recently, Emran and Shilpi (2015a) analyzed the trends and patterns in educational mobility using data provided by two rounds of NFHS: NFHS-I (1992-93) and NFHS-II (2005-06). They show that household background plays a vital role in determining the educational outcomes of the children. Comparing their findings with Latin American countries, they found that India has higher level of intergenerational correlation and sibling correlations.

Developing countries like, India face an issue of lack of datasets that can provide parent’s information without relying on co-residency condition, which pose a major challenge in measuring intergenerational educational persistence (Azam & Bhatt, 2015a). Most of the earlier works on IEM in India had relied on co-residency condition to measure IEM in India. However, Azam and Bhatt (2015) find that this limitation helps in identifying only one-third of the adult male population; and perhaps more importantly, this constraint creates a systemic bias as it over represents the young population. Azam and Bhatt (2015a) study IEM using the data from IHDS. The IHDS dataset has two advantages. First, unlike the NSS data, the panel nature of IHDS data offers many research opportunities, where same houses are re-interviewed in different rounds. This provides multiple observations for a parent-child pair. Second, IHDS data provides information for parental education even when the child is not residing in the same household as the parents. Therefore, it avoids the sample selection bias caused by the co-residency constraint imposed by most of the earlier studies. They found that IEM in terms of educational attainment has increased, irrespective of one’s association with a particular social group. In addition, the increase in mobility can be evidenced across all Indian states; however, differences still exist between states with some faring far better than others in terms of improvements in IEM.

Emran et al. (2016) finds that truncation due to co- residency condition can be a major challenge in measuring IEM. Using two rich datasets, the 1999 Rural Economic and Demographic Survey (REDS) of India and the 1999 Matlab Health and Socioeconomic Survey

(MHSS) of Bangladesh; they found that estimates of intergenerational regression coefficient (IRC) are severely downward biased. However, intergenerational correlation coefficient (ICC) estimates suffer from a lower bias. The authors recommend that the potential sample selection bias emanating from the co-residency condition could be resolved by measuring IEM using ICC rather than IRC. In their empirical analysis, they found that the average bias in ICC was less than 5%, whereas the average bias in IRC was about 30%.

This study makes contribution to different strands of literature which uses different indicators of child's educational status. It is well known that human capital accumulation is strongly associated with the level of economic growth (Barro, 2001). But, in the past few years increasing drop-out rates have become a cause of concern for policy makes, in both developed and developing countries. While this could be attributed to different reasons, but it is likely to cause intergenerational transmission of social disadvantage. Therefore, child's schooling progression can convey information about future social mobility. There has been vast work on the determinants of grade progression in developing countries. Different studies have used different indicators of schooling as proxy for child schooling, including completed years of education (Birdsall, 1985), ever attended school (Cochrane et al.,1986), completion of grades (Dreze and Kingdon, 2001) and late enrolment (Glewwe and Jacoby, 1995). The models are often estimated separately for girls and boys as well as for rural and urban sector. Most of these studies analyzing child schooling progression have relied on either Ordinary Least Square (OLS) or Two-Stage Least Square (2SLS) estimation method or multivariate ordered probit/logit.

Findings show that interaction of demand and supply side factors along with government policies influence the likelihood of educational transitions at different levels (Glewwe and Jacoby, 1994, 1995; Dreze and Kingdon, 2001). Studies on schooling progression have found that parental educational achievement along with their employment status increases the likelihood of their child moving up the education ladder (Lam and Schoeni, 1993; Knight and Sabot, 1990). However, Duraisamy (1998) finds mixed results with respect to the impact of parental education (both paternal and maternal education level) on school enrolment and child's completed years of education. Jayachandran (2002) finds that increased female work force participation could have a negative impact on schooling progression in girls. On the contrary, it is argued that higher level of work participation by women is likely to increase school enrolment (Alderman et al, 1996). The impact of employment status of parents on school enrolment is contingent on the economic condition of the household. Children belonging to financially constrained households are prone to drop out at secondary level of education.

Not only parental education level or employment status but also their socioeconomic position in the society determines the education attainment of their child. Shavit and Blossfeld (1993) and Gerber and Hout (1995) use the model of school continuation proposed by Mare (1980, 1981) and they find that the effect of social background erodes across educational transitions. However, their study does not control for the plausible biasness of estimates due to unobserved heterogeneity. Some of the recent studies address this problem to provide robust determinants of child schooling progression (Bernandi,2012; Breen and Jonsson, 2000; Buis,2011; Lucas, 2001; Lucas, Fucella and Berends,2011). Another strand of literature examines the inverse relationship between the in the household and the education attainment of the child (Downey,1995)

2.4.2 Migration as a determinant of Intergenerational Educational Mobility

The degree of IEM in terms of educational attainment among the non-migrants (natives) is likely to differ from their migrant counterparts. This issue has been touched upon in some of the recent studies that investigate the heterogeneity in IEM among natives and migrants. Borjas (1992) found that apart from the parental education level, education of immigrant children is influenced by the average education level of the ethnic community to which they belong. Furthermore, immigrants may often struggle to integrate in the local society. If the immigrant population faces difficulty in integrating with the local society, the parental education level generally plays a pivotal role in determining their offspring's education level. Lack of access to public resources offered to the migrants act as a barrier which prevents immigrant children from climbing up the social ladder; hence, they are likely to depend more on private investments, such as household assets or parental education level, rather than public investments (Ammermueller, 2007; Schneeweis, 2011). In the Indian context, a majority of migrants relocate from rural to urban areas in search of better employment prospects. Since, the education levels in the rural regions are much lower than those in urban regions, the average education level of the migrant population is prone to be on the lower side than that of their native counterparts. Therefore, the migrant children would be inherently disadvantaged in comparison to the children of the natives.

Studies show that internal migration in India has increased in the last few decades (Srivastava and Sasikumar, 2003). India has also seen a drastic spur in short-term migration in recent years. In developing countries like India, internal migration takes the form of temporary or circular movement between rural and urban This makes it imperative to explore the link between short-term migration and educational attainment. Overtime, an extensive literature has evolved on the causes of short-term/circular migration in India (Haberfeld et al., 1999; Nabi, 1984). Even though the concept of IEM has been gaining momentum lately, there is still lack of studies which try to look for possible link between short-term migration and IEM. Therefore, it will be interesting to explore whether short-term migrants play a role in transmission of education across generations. Therefore, we also study the role of short-term migration in IEM in India.

2.4.3 Other factors

To our knowledge, there is no existing study which explores the determinants of IEM in India. Some international research, however, has examined the role of certain socioeconomic factors, such as liquidity constraint, child's school entry age and parenting skills in affecting IEM.

Bauer & Riphahn (2006) used Swiss data to show that child's timing when entering into primary school plays an important factor in affecting the extent of IEM. They use data from Swiss population census, 2000 to find that early age at the time of stating school is a crucial determinant of educational mobility and diminishes the comparative gain of children having high educated parents. (Gaviria, 2002) analyzed the differences in IEM between rich and poor families. They showed that liquidity constraint is one of the important determinan of IEM. Their study used a sample of father and children from the Panel Study of Income Dynamic (PSID) sample to show that borrowing constraints inhibits poor parents from making optimal investment decisions regarding the education of their children. This causes low level of mobility among poor households. However, their study was based on IEM in terms of income. Nimubona & Vencatachellum (2007) found that blacks are more immobile than whites. They pointed out that factors such as access to financial market and availability and quality of schools are important factors that affect the variations in IEM.

This study tries to address some of the untouched issues related to IEM in India. First, we provide robust estimates of IEM by addressing the endogeneity problem using the novel two stage estimation procedure of Lewbel (2012). Second, whereas, Azam & Bhatt (2015) concentrate on IEM by analyzing only father–son pairs, we extend their analysis by including all possible parent-child pairs. This allows us to gain unique insight in the role of gender in influencing IEM. This analysis is of particular interest as gender inequality is pervasive in India, and Indian females are distinctly disadvantaged in comparison to their male counterparts over most socio-economic criteria (see, for example, Ackerson and Subramanian, 2008; Arora, 2012; Behrman, 1988; Bhattacharya, 2006; Borooah, 2004; Dunn, 1993; Jacobs, 1996; Kishor, 1993; Murthi et al., 1995). Third, the panel dataset of IHDS surveys allows us to examine the role of migration on IEM. In this study, we define migrants as individuals who relocated between the intervening period of the two rounds of the IHDS surveys, i.e., between 2004 and 2011, and who were enrolled at the time of migration. Fourth, this study also tries to identify the source of persistence in transmission of educational mobility among migrants and non-migrants by decomposition of ICC.

2.5 Summary and Conclusion

The review of literature suggests that there are many socio-economic variables that affect the extent of IEM. Most of the previous relevant literature has found that the overall educational attainment and degree of IEM has increased over time. But, still not all the sections of the society face equal educational opportunities. Even after decades of educational progress, children belonging to disadvantaged sections of the society are denied of equal educational opportunities. Therefore, the more recent literature took a step forward to identify the source of increasing educational mobility.

The conceptual framework and survey of past work reveals that there are several variables that influence the educational attainment of the child. Education level of parents is found to be the most significant factor determining the child's educational status. Being a complex phenomenon, the extent of mobility across generations is assumed to be driven by many social and economic variables. Some authors have found that migration, gender and other socioeconomic variables affects the degree of IEM. However, most of these studies are restricted to developed nations. In developing countries, like India there is dearth of literature which focuses on variables which mediate the association between the education level of parents and their children.

Chapter 3

Data Sources & Methodology

3.1 Introduction

The chapter describes dataset and empirical methodology used in the subsequent analyses. As pointed out in the previous chapters, our focus is on examining the correlates and determinants of IEM. Given the nature of the topic, we require use of different methodologies to assess the relationship between various socio-economic variables and IEM.

This chapter is organized as follows. In Section 3.2, we discuss the data sources used for the analysis. Section 3.3 outlines the various empirical methods of measuring IEM. Section 3.4 includes the methodology used to examine the relationship between parental education and child's education. Finally, Section 3.5 concludes.

3.2 Data Sources

This study provides new estimates of IEM in India. It also makes a novel attempt to examine the determinants of IEM. The estimation methods used in the study ensures that our estimates are robust. We have used two main datasets to measure IEM and identify their determinants.

3.2.1 The Indian Human Development Survey (IHDS)

The IHDS is a nationally representative household survey with data on education, health, employment, among other topics. The survey was jointly conducted by the University of Maryland and the National Council of Applied Economic Research (NCAER). We use both rounds of IHDS: IHDS-I and IHDS-II. IHDS-I was conducted in 2004-05 and it provides information regarding various socioeconomic topics by surveying 41,554 households. IHDS-2 was the follow-up survey that revisited more than 40,000 households which were also part of the IHDS-1 survey. This helped in creating the largest panel dataset and offers economic progress over the period of seven years. IHDS datasets has several benefits for estimating IEM compared to other Indian dataset like the NSS or the National Family Health Survey (NFHS). First, unlike NSS data, it offers information regarding the educational level of the father regardless of whether the parent co-resides with their children or not. This enables us to relax the co-residency restriction, and thereby avoid selection bias issues in IEM estimates. Second, IHDS provides a continuous variable on years of attainment rather than completed levels of education provided by the NSS dataset. Therefore, IHDS would be preferable dataset as compared to other Indian datasets like NSS, NFHS, etc.

3.2.2 The National Sample Surveys (NSS)

The National Sample Survey Organization (NSSO) conducts large scale surveys on socioeconomic indicators like employment, health, and education. The interviewing households are randomly selected by following a unique sampling design and it covers all Union Territories and states in India. We use six rounds of Employment and Unemployment Surveys (EUS) – a component of NSS – to identify the macroeconomics determinants of IEM. The EUS offers information on various aspects of labour market outcomes including wages and occupation. Various thick rounds of EUS helps us in performing a pooled regression. This allows us to identify the determinants of IEM in India.

3.3 Measuring IEM

3.3.1 Method 1: Intergenerational Regression Coefficient (IRC)

Following Hertz et al., we estimate IRC (IRC) is measured as follows

$$EDU_i^C = \propto + \beta^P EDU_i^P + \varepsilon_i \tag{3.1}$$

where EDU_i^C is the completed education (in years) for the i^{th}child and EDU_i^P is the parental education. β^P is the estimated IRC which shows the average change in the education variable of the child corresponding to a one - year change in the education level of the parent. The coefficient β^P shows intergenerational persistence of education.

3.3.2 Method 2: Intergenerational Correlation Coefficient (ICC)

The ICC, ρ, can be estimated as follows

$$\rho = \left(\frac{\sigma_i^P}{\sigma_i^C}\right)\beta^P \tag{3.2}$$

Here the superscript P refers to parents (fathers or mothers) and the superscript C refers to children. As noted earlier, the dataset is divided into ten-year birth cohorts. For a specific birth cohort, $\sigma_i^C\,(\sigma_i^P)$ is the standard deviation of child (parent) education. In the analyses of father–daughter and father–son pairs, β^P is the IRC for fathers. In the analyses of mother–son and mother–daughter pairs, β^P is the IRC for mothers.

The value of ICC ranges between -1 and +1. The main difference between IRC and ICC is that the former describes association between absolute attainment, whereas latter indicates similarity in positional ranking for both generations within the educational distribution of their cohort (Black and Devereux, 2011). Note that in the estimation of ICC, the differences in the educational distribution both generations are normalized by using a ratio of standard deviations of parent and child education, whereas the IRC is sensitive to differences in educational distributions of both generations. A decrease in ICC over time implies that the intergenerational persistence has declined. However, ICC measure could also fall over time due to education reforms, such as compulsory primary school education, that lower the variance in educational distribution of children relative to that of their parents. Hertz *et al.* (2007) show that both IRC and ICC can move in different directions over time, and for this reason, recommends reporting both measures of IEM.

3.3.3 Method 3: Transition Probabilities

Most of the empirical studies have relied on IRC, as a measure of the degree of IEM. The regression coefficient shows the average changes in the education level of parent and their child across generations. However, transition probability is yet another measure that can give a lot of insightful information regarding the probability of a child achieving certain levels of education conditional on parental education. Therefore, we compute two different indicators of IEM.

3.3.3.1 Bottom Upward Mobility (BUM)

The Bottom Upward Mobility (BUM) is computed as follows

$$BUM_{jk} = Prob\ (y_{ijk}^{c} \geq s | y_{ijk}^{p} < s) \tag{3.3}$$

where y_{ijk}^{c} is the education of the i^{th} child belonging to the j^{th} cohort of the k^{th} state. 's' is some minimum education level. In this analysis, 's' denotes completion of middle level of schooling. Therefore, in our case the BUM indicator shows the probability of children achieving at least secondary level of education given their parents have obtained less than secondary-level education.

3.3.3.2 Upper Class Persistence (UCP)

The Upper-Class Persistence (UCP) is measured as follows

$$UCP_{jk} = Prob\ (y_{ijk}^{c} \geq s | y_{ijk}^{p} \geq s) \tag{3.4}$$

where y_{ijk}^{c} is the education of the i^{th} child belonging to the j^{th} cohort of the k^{th} state. ‘s’ refers to the certain minimum level of education. In this analysis, ‘s’ denotes completion of secondary level of schooling. Therefore, in our case the UCP indicator shows the probability of children achieving at least secondary level of education given their parents have also obtained at least secondary-level education.

3.3.4 Method 4: Absolute and Directional Mobility

The BUM and UCP measures represent the absolute and relative attributes of IEM by measuring the changes in educational level of the parent and child within the distribution over time. However, they hide information regarding the size of these movements. Therefore, following Fields and Ok (1996) we measure two more indices which will allow us to estimate per capita movement in education attainment.

The absolute and directional mobility is measured as follows

$$M1_{jk} = \frac{1}{N_{jk}} \sum_{i=1}^{N_{jk}} |y_{ijk}^{c} - y_{ijk}^{p}| \tag{3.5}$$

$$M2_{jk} = \frac{1}{N_{jk}} \sum_{i=1}^{N_{jk}} (y_{ijk}^{c} - y_{ijk}^{p}) \tag{3.6}$$

M1 shows the average difference in the educational level of the parent and the child within the same families but does not show the direction of movement. It shows the net sum of upward and downward movement. The second measure M2 measures the average directional change in movement of education level between the two generations. Higher value of M2 can imply higher mobility due to the upward trend in educational attainment levels over time. Therefore, M1 and M2 together can give useful information on educational mobility. Smaller difference between these two measures implies lower degree of downward mobility.

3.4 Causal analysis

While there is plethora of papers concerning IEM, which have documented intergenerational persistence in educational outcomes, in the last few years there has been a surge of interest in identifying the underlying causes that drive this relationship across generations. In essence, these works try to disaggregate the observed parental education effect into two components– the component due to inherited genes and the component due to child's environment. This nature vs nurture debate has garnered a lot of attention in recent years because of its importance in shaping debates and policy initiatives.

Researchers have used different methodologies to identify the sources of correlation between parents and child education levels. This section describes different methods which we use to examine the causal effect of parental education on child education.

3.4.1 Decomposing ICC

From the perspective of policy makers, upward mobility may be viewed as a preferable outcome than downward mobility. However, irrespective of the type of mobility, high levels of mobility would reduce ICC. Therefore, better understanding of IEM can be gained by examining the source of ICC. The ICC measure can be decomposed as follows

$$\rho = \frac{\sigma^{CP}}{\sigma^{C}\sigma^{P}} = \sum_{C,F}(y^{C} - \mu^{C})(y^{P} - \mu^{P})\Pr(y^{C}|y^{P})\Pr(y^{P})/\sigma^{C}\sigma^{P} \tag{3.7}$$

Where, for a particular birth cohort, the superscript C indicates the set of all children (either sons or daughters) in the same birth cohort; and the superscript P indicates the set of all parents (either fathers or mothers) of the children in that cohort. μ^{C} and μ^{P} are the mean education level for the children and parents, respectively. σ^{C} and σ^{P} are standard deviations of the education distributions of children and parents, respectively. The covariance between parent and child education is given by σ^{CP}. y^{C} and y^{P} are the education level of a particular parent-child pair. $\Pr(y^{C}|y^{P})$ is the conditional probability that child achieves y^{C} education level, when parent's education level is y^{P}. $\Pr(y^{P})$ is the probability that the parent has the education level, y^{P}.

The decomposition in equation (3.7) implies that ρ relates to deviations from mean for both parent and child education, conditional likelihood that child will attain some education level given parent's education, and the probability that the parent can achieve that education level. When both generations parent and child achieve below (or above) mean education, the parent-child pair increases ρ , whereas if the parent is below (above) mean and the child is above (below) the mean, the parent-child pair reduces ρ.

3.4.2 Lewbel Instrumental Variable (IV) Approach

Heteroskedasticity can be exploited as exogenous shift similar to standard external instruments to solve the identification problem (Klein & Vella, 2009; Lewbel, 2012; Rigobon, 2003). We use the heteroscedasticity based identification strategy of Lewbel (2012), which involves a two-step estimation process. First, heteroskedastic covariance restriction is used to construct internal instruments. Then, these internal instruments are used in the standard two-stage instrumental variable estimation. Mishra & Smyth (2015) show that the Lewbel instrumental variable estimation works well in mitigating unobserved selection bias when conventional external instruments are not available. Recent applications of the Lewbel's twostage instrumental variable estimation include Acheampong *et al.* (2021), Churchill *et al.* (2020), Feeny *et al.* (2021),and Wang and Zhu (2021).

A major problem with study of IEM is that parental education variables are endogenous variable. The secondary data sources rarely provide information about valid instruments that satisfy the exclusion restriction. To overcome this limitation, we employ the novel identification strategy of Lewbel (2012). The strategy involves a standard two stage estimation process. In the stage, we generate synthetic instrumental variables for all endogenous regressors. These instruments are some linear combination of exogenous variables. The second stage involves estimating the model by substituting instrumental variables for the endogenous variables. This approach can be defined as follows.

We estimate the following three-equation model:

$$Edu_i^C = Edu_i^F \beta^F + Edu_i^M \beta^M + X_i \gamma + \varepsilon_i \tag{3.8}$$

$$Edu_i^F = X_i' \gamma^F + \varepsilon_i^F \tag{3.9}$$

$$Edu_i^M = X_i' \gamma^M + \varepsilon_i^M \tag{3.10}$$

Where Edu_i^M and Edu_i^F denote education of father and mother of ith child. X_i is a vector of controls. Although parental education variables are potentially endogenous, Lewbel (2012) notes that the regression coefficients, β^F and β^M, can be identified given some set of exogenous variables $Z_i \subseteq X_i$, that satisfy:

(i) $\mathrm{E}(X_i', \varepsilon_i) = 0$, $\mathrm{E}(X_i', \varepsilon_i^F) = 0$, $\mathrm{E}(X_i', \varepsilon_i^M) = 0$

(ii) $\mathrm{Cov}(Z_i, \varepsilon_i \varepsilon_i^F) = 0$,$\mathrm{Cov}(Z_i, \varepsilon_i \varepsilon_i^M) = 0$, and

(iii) $\mathrm{Cov}\ (Z_i, (\varepsilon_i^F)^2) \neq 0$ and $\mathrm{Cov}\ (Z_i, (\varepsilon_i^M)^2) \neq 0$

Under the above assumptions, $(Z_i - E(Z_i))\varepsilon_i^F$ and $(Z_i - E(Z_i))\varepsilon_i^M$ can instrument Edu_i^F and Edu_i^M. Since this identification strategy depends on higher moments, the Lewbel estimates tend to be less reliable than conventional instrumental variable based estimation (Lewbel, 2012). To examine whether this is problematic, one has to compare Lewbel IV estimates with standard IV estimates based on external instruments. Lewbel (2012) reports an empirical application which addresses measurement error in total expenditures and shows that the estimates obtained using heteroscedasticity-based identification are similar to those obtained using standard instrumental variable regression. Additionally, we examine the internal instruments utilized in Lewbel IV regressions using standard identification tests – Hansen J test for the test of overidentifying restrictions, and Kleibergen-Paap test for under-identification. We use the generalized method of moments (GMM) for model estimation, as GMM estimates tend to be more efficient than the standard two-stage least squares estimates (Baum et al., 2003).

3.4.3 Propensity score matching (PSM)

The first step in estimating causality requires random selection of subjects as well as random assignment of treatment to subjects. However, in observational studies the randomness of the selection of subjects who receive the treatment is not feasible. In case of observational studies, random assignment of treatment cannot be ensured. This nonrandom selection of subjects receiving the treatment makes it difficult to infer causal associations because difference in outcome may be due to heterogeneity in characteristics of treatment and control groups.

Rosenbaum and Rubin (1983) proposed propensity score matching (PSM) as a method to estimate treatment effect when random treatment assignment is not possible. In this method, propensity scores are used to match treatment and control units so that matched units have similar likelihood to receive treatment conditional on a set of observed covariates. Thus, PSM controls for selection on observed covariates while estimating the treatment effects. The primary concern when matching is done directly on the covariates is the problem of multidimensionality. This dimensionality problem is taken care of by the PSM technique, which reduces the dimensionality of matching to a single dimension.

3.4.3.1 Calculating Propensity Scores

We use logistic regression method for estimating the propensity scores. In this method, we estimate the probability of treatment assignment. The model can be written as

$$ln \frac{e(x_i)}{1-e(x_i)} = \ln \frac{\Pr(z_i=1|x_i)}{1-Pr(z_i=1|x_i)} = \propto +\beta x_i \tag{3.11}$$

where $e(x_i)$ represents the probability that the subject receives treatment $(z_i = 1)$, conditional on the observed values of covariates (x_i).

The model can also be expressed in terms of treatment probability $e(x_i)$ as follows,

$$e(X_i) = b_0 + b_1X_1 + b_2X_2 + \ldots \ldots \ldots + b_nX_n \tag{3.12}$$

3.4.3.2 Grouping the Data and Matching

After we have estimated the propensity scores, next step requires dividing the original dataset into two groups of those who received the treatment and the control group. The subjects from both groups are matched based on propensity scores. Different matching techniques use different methods to match propensity scores of the two groups. Finally, treatment effect is estimated as the mean of the difference between outcome of treatment units and matched control units.

3.4.3.3 Estimating Average Treatment Effect on Treated (ATT)

Matching is best suited to situations where there are large number of control units for each treated unit. Generally, in such scenarios, one can identify close matches for each treatment unit within the control group based on the estimated propensity scores. The subjects that are not able to get matched are removed. The primary benefit of matching is that the two groups get compared while controlling for all the observed covariates by matching on a single scalar value, named as propensity score. However, this technique is not devoid of limitations. For example, if the treatment and control groups do not have sufficient overlap in terms of matched characteristics, then treatment effect estimates obtained using matching may be spurious. It also relies on a strong assumption that all the relevant covariates have been included in the estimation of propensity scores. Further, there is a lack of consensus on whether matching of two groups with similar propensity scores should be based on ‘with replacement’ or ‘without replacement’. However, if matching is done with replacement where every control subject can be a match for more than one matching pair, increasing the number of matched pairs. But, the inclusion of same control subject in multiple matched pairs can lead to biased treatment effect estimates.

In chapters 4 and 5, we generate matching based estimates of IEM to check the robustness of traditional regression-based estimates of IEM. The parental education level is used as the treatment which equals 1 if parental education exceeds certain threshold and 0 otherwise. The former set of children comprise the treatment group, whereas the latter set of children comprise the control group. The outcome variable is the education attainment of their child measures as years of schooling.

While the outcome variable for each child is observed only under one of the two possible treatment states, the counterfactual framework suggests that each child has a potential outcome under both treatment states (Morgan & Winship, 2015). More formally, suppose the outcome variable for child i is $Y_i(T)$, where the treatment variable T equals 1 if the parental education exceeds certain threshold and 0 otherwise. Then, the treatment effect for the child, τ_i, is defined as follows

$$\tau_i = Y_i(1) - Y_i(0) \tag{3.13}$$

For each child, one of the two terms is always missing in our data, since for the treatment group we observe $Y_i(1)$ but we do not observe $Y_i(0)$, whereas for the control group we observe $Y_i(0)$ but we do not observe $Y_i(1)$. A naïve approach for solving this problem is to compare the average education of the treatment and control groups. This approach can be applied if the treatment assignment is random, and the children in treatment and control group are similar in all other characteristics that may potentially affect their educational outcomes, i.e., both groups have balanced covariate distributions. Since our study is not based on randomized trials but on observational data, the assumption of balanced covariate distributions is problematic due to the non-random assignment of treatment.

For example, suppose child's education is unrelated to parental education, however, individuals belonging to affluent households are more likely to have higher education. In this case, on an average, the treatment group would comprise of parent-child pairs from more affluent households as relative to parent-child pairs in the control group. It follows that the parent-child pairs in the treatment group would have higher average education that those in the control group. In this scenario, a naïve comparison of the average educational outcome of the treatment and control groups would suggest a spurious causal effect of parental education on child education, while the actual relation is driven by the difference in household wealth between the two groups.

We use PSM to match treatment and control groups over a large set of socioeconomic characteristics. The matching approach allows us to estimate quasi-experimental contrasts between by matching the individuals belonging to the treatment groups with comparable individuals from the control group (Morgan & Winship, 2015). The ATT is then estimated as

$$ATT = \frac{1}{N_T} \sum_{i,T_i=1} \left(Y_i(1) - \hat{Y}_i(0)\right) \tag{3.14}$$

where. T_i is the treatment indicator which equals 1 if the parental education exceeds certain threshold, 0 otherwise. N_T is the number of children for which $T_i = 1$. To estimate ATT we restrict the summation to children belonging to the treatment group ($T_i = 1$). $Y_i(1)$ is the education of child i in the treatment group, and $\hat{Y}_i(0)$ is the counterfactual control outcome for the matched counterpart of the child i.

We estimate this counterfactual outcome using PSM. We estimate a logistic regression model to determine each child's propensity score (P_i), i.e., the probability that child i receives the treatment as

$$\log \frac{P_i}{1-P_i} = \alpha + \beta_k X_i + \varepsilon_i \tag{3.15}$$

Where β_k is a vector of k coefficients corresponding to socioeconomic characteristics X_i for child i, and ε_i is the random error term that is logistically distributed.

We use the nearest neighbour method to identify closest control unit for each treatment unit. In addition, we employ several kernel based matching estimators (Heckman et al., 1998) that provides precise matching than the standard nearest neighbour or radius based matching (Frölich, 2004). With the kernel estimators, ATT is estimated as

$$ATT = \frac{1}{N_T} \sum_{i, T_i = 1} \left(Y_i(1) - \sum_{j, T_j = 0} w(i,j) Y_j(0) \right) \tag{3.16}$$

$w(i,j)$ is the weight of a control unit j matched to a treatment unit i. This weight is defined using a kernel function, $K(\theta)$

$$K(\theta), \theta = \frac{\frac{P_i}{1-P_i} - \frac{P_j}{1-P_j}}{h} \tag{3.17}$$

$\frac{P_i}{1-P_i}$ is the odds ratio of receiving treatment for the treatment unit i, and $\frac{P_j}{1-P_j}$ is odds ratio of receiving treatment for the control unit j. θ therefore represents the quality of the match. h is the optimal bandwidth parameter. The weight function is defined as $w(i,j) = \frac{K(\theta)}{\sum_j K(\theta)}$ so that summing kernel weights for all j would give a total weight of 1. Following Alcott (2017), we use four kernel based estimators, namely, uniform, normal, biweight and the Epanechnikov kernel to ensure that the estimates are robust to the choice of kernel weighting scheme.

We also generate IEM estimates based on Mahalanobis distance matching. The Mahalanobis distance matching measures the distance $(MD_{i,j})$ between a treatment unit i and a control unit j as follows

$$MD_{i,j} = \sqrt{(X_i - X_j)' S^{-1} (X_i - X_j)} \tag{3.18}$$

X_i and X_j are vectors of k individual and household characteristics for child i and j, respectively. S represents the pooled sample covariance matrix for these k individual and household characteristics. The control unit having the lowest Mahalanobis distance with the treatment unit is matched with that treatment unit.

We only estimate ATT, which is the average impact of treatment (parental education above a certain threshold) on those who received it, and do not attempt a broader estimate of the impact of treatment on those children who do not receive it, i.e., average treatment effect (ATE). Thus, one must exercise caution in extrapolating our results for those children who did not receive the treatment. Nonetheless, a positive ATT estimate suggests that for the group of the children receiving the treatment, their educational attainment would have reduced if they did not receive the treatment.

3.4.3.4 Sensitivity Analysis

The accuracy of results from matching techniques is based on the assumption of conditional independence or confoundedness. This means while matching one must control for all potential confounders that can affect both treatment assignment and the outcome. Obviosuly, this is a strong identifying assumption. In the presence of selection on unobservable confounders, matching estimators are less likely to be robust due to the 'hidden bias' (Rosenbaum, 2002). This makes it important to check for the sensitivity of our ATT estimates for different scenarios, considering the impact of deviation from the assumption of conditional independence. Therefore, to address this issue we compute the Rosenbaum Bounds (Rosenbaum 2002). This approach allows us to check the robustness of ATT estimates unobserved heterogeneity. Aakvik (2001) suggests to using Mantel and Haenzel (MH) test- statistic to estimate Rosenbaum bounds for ATT estimates.

Matching estimators simulate a quasi-randomized experiment controlling for observed variables which can affect the propensity of treatment assignment. Through matching, the distribution of these variables is made similar across both treatment and control groups, which in turn, makes the probability of treatment assignment similar for matched individuals. In studies based on observational data, the causal treatment effect estimates can be biased due to unobserved confounders that affect the likelihood of receiving treatment. The Rosenbaum sensitivity analysis defines employs a parameter Γ, which is the odds ratio of receiving treatment due to some unobserved confounder. Suppose two individuals i and j that are matched on a set of observed covariates, X. If there is no hidden selection bias, the probability of receiving treatment for individual i, π_i, should be same as probability of receiving treatment for individual j, π_j. However, in the presence of an unobserved confounder that affects the likelihood of treatment assignment $\pi_i \neq \pi_j$. Γ is the ratio of odds of treatment assignment to matched individuals i and j. $\Gamma = 1$ suggests that both the matched individuals would have same odds of treatment assignment. $\Gamma = 2$ indicates selection on unobserved confounders can double the odds of treatment assignment. We evaluating the ATT estimates at different levels of the sensitivity parameter Γ to ensure that they are robust to hidden selection bias. For instance, suppose the ATT estimate remains significant at $\Gamma = 6$, this suggests that the estimated causal effect can not be attributed to the unobserved confounder even when it causes a six-fold increase in the odds of treatment assignment.

3.5 Conclusion

In this chapter, we have discussed the data sources and methodological framework used in analysing the degree of IEM in India and identifying it's determinants. The first part of the chapter gives details regarding the sources of data used in the analysis. We use data from two round of IHDS to study the first four objectives of our study. To study the macroeconomic dynamics of IEM, we rely on the data provided by NSSO. We use two rounds of NSSO (55^{th} round and 68^{th} round) to identify the correlates of IEM in India. Subsequently, we discuss the different empirical strategies used to estimate IEM in this study. Finally, we describe the various approaches used to conduct causal analyses.

Chapter 4

Trends in Intergenerational Educational Mobility

4.1 Introduction

This chapter addresses the trends in IEM in India. IEM represents the extent to which child education depends on the education of their parents, and therefore it is an indicator of equality of education opportunity. Low level of IEM suggests that individual's educational outcome is strongly associated with parental education. In other words, low level of IEM implies that children of highly educated parents tend to have high education themselves, whereas higher education acts as a glass ceiling for children of less educated parents. Similar concepts such as intergenerational income and occupational mobility are used to represent persistence of economic status (Jerrim & Macmillan, 2015). Most studies on trends in educational level in developed countries show an upward trend in the general level of education as well as relative education level (Belanger, 2012). However, in the case of developing nations, extant evidence suggests high degree of educational persistence at both the ends of income and education distribution (Chusseau et al., 2013).

The literature on IEM in India has generally observed that parental education level significantly affects the level of education of their children. We extend this literature and examine the pattern of IEM across age cohorts using the latest available IHDS data. The study also employs a novel empirical strategies to examine the causal impact of parental education on child's educational level. Further, we decompose the ICC to identify the sources of persistence.

The remainder of this chapter is organized as follows. Section 4.2 describes the data and empirical methodology used for the analysis. Results are presented in Section 4.3 and Section 4.4 presents the decomposition of ICC. Section 4.5 concludes.

4.2 Data and Methodology

4.2.1 Data Source

We use IHDS-II data for this analysis. Education is measured as completed years of education. Inclusion of individuals who were still enrolled at the time of the survey may lead to downward bias in the IEM estimates. Hence, we restrict the analysis to individuals within the age-group 25 to 65 years, with the expectation that they would have completed their education.

Table 4.1 reports summary statistics for children who were enrolled at the time of IHDS-II. Column 1 of Table 4.1 shows the total descriptive statistics for two different age-groups: individuals who are aged between 20 to 24 years and those who are between 25 to 29 years. Column 2 and 3 reports the total number of individuals who are currently enrolled and their respective share in the total sample size, respectively, for both the age-groups. Our results show that the share of individuals who are currently enrolled is very large for those aged 20-24 when compared to those aged 25-29, irrespective of the parent-child pair. Column 4 and 5 of Table 4.1 presents the total number of individuals who are currently enrolled but have not achieved more than 14 years of education and their share in total sample size, for both the age-groups, respectively. This right-censoring in the data shows that the true value of education for these individuals would be just one or two year above the observed level of education for the individual. Results show that for those aged between 20-24, these share ranges between 20-30%, while for those aged between 25-29 share ranges between 2-5%. This shows that inclusion of individuals belonging to the latter age-group is unlikely to cause any biasness in the IEM estimates. Additionally, exclusion of individuals aged between 25-29 years will cause loss of valuable observations. We also observe that the percentage share of enrolled individuals is quite high for daughters than sons. This indicates the gender gap in delayed education.

The upper bound of 65 years of age is chosen as there were few respondents above the age of 65 for which data on paternal education was available. Additionally, inclusion of individuals aged 65 and above may also induce an upward bias in the mobility estimates due to a high correlation between life expectancy and education attainment (Olshansky et al., 2012). This sample size restriction is also consistent with previous analysis of IEM (Hout and Janus, 2011; Long et al., 2012).

The education variable ranges from 0 to 16 years, where 0 represents 'no education' and 16 represents highest educational qualification obtained by the child.

4.2.2 Matching Parent-Child Pair

Most household survey datasets suffer from co-residence restriction wherein the data on parent and their children is available only when both parent and their children co-reside in the same household. For creating parent-child pairs, we use the "*Household ID*" variable in the household roster and link individuals with their respective parents by using information under "*Relation to the Head"* and "*Gender*" variable. This allows us to identify parent-child pairs who are co-residing in the same household. IHDS data allows us to create matched parent-child pairs without imposing the co-residence restriction. The data truncation caused by the co-residence restriction can bias the mobility estimates, and therefore we attempt to quantify the magnitude of this bias by relaxing the co-residence restriction.

The IHDS-II dataset provide us some new information that is being used to create matched father-son pairs without imposing the co-residence restriction. IHDS provides information regarding education level of the father of the household head, even when the father does not reside in the same household. Using this variable, we are able to create father-son pairs without imposing the co-residence condition. In addition, a new questionnaire on "eligible women" allows us to generate father-daughter and mother-daughter pairs even when the parents and their daughters reside in different households. Therefore, we are able to relax the co-residence restriction for three out of the four parent-child pairs, namely, father-son, father-daughter and mother-daughter. Hereafter, "full sample" refers to a set of all parent-child pairs and "co-resident" sample refers to a set of parent-child pairs that co-reside in the same household.

4.2.3 Methodology

4.2.3.1 Intergenerational Regression Coefficient (IRC)

Following Hertz et al., IRC is measured as

$$EDU_i^C = \propto + \beta^P EDU_i^P + \varepsilon_i \qquad (4.1)$$

where EDU_i^C is education of the i^{th}child and EDU_i^P is the education of her parent. ε_i is the estimated regression error term. β^P , our main parameter of interest is the IRC estimate which shows the average change in the education of the child corresponding to a one - year change in parental education. This measure is very easy to interpret. The value of β^P shows the degree of association between the education levels of parent and children. High value of β^P suggest that child's educational outcome highly dependent on parental education, and vice versa. In short, the IRC shows the degree of association between the parental education level and educational attainment of the child.

4.2.3.2 Intergenerational Correlation Coefficient (ICC)

The ICC shows the similarity between the ranking of parents and children in their respective education distributions. This measure accounts for changes in inequality across generations (Bjorklund and Jantti, 2009). To estimate ICC, we standardize the education variable by dividing it with the standard deviation of the education distribution. The model used to estimate ICC is as follows

$$(EDU_{adj})_i^C = \alpha^a + \rho\ (EDU_{adj})_i^P + \ \varepsilon_i \tag{4.2}$$

where $(EDU_{adj})_i^C$ and $(EDU_{adj})_i^P$ denote standardized education for child and parent of the i^{th} household, respectively. The education variable is adjusted by dividing it by the standard deviation of education distribution, β^a. ρ is the ICC estimate, which can be interpreted as the standard deviation increase in $(EDU_{adj})_i^C$ due to a one standard deviation increase in $(EDU_{adj})_i^P$. The two mobility measures are related as follows

$$\rho = \left(\frac{\sigma_i^P}{\sigma_i^C}\right)\beta^P \tag{4.3}$$

Here the superscript P refers to parents (fathers or mothers) and the superscript C refers to children. We generate separate mobility estimates for five decadal birth cohorts. For a specific birth cohort, σ_i^C is the standard deviation of child education. σ_i^P is the standard deviation of parental education.

Generally, ICC is the preferred measure of mobility if we want to compare mobility estimates across different groups. Furthermore, it can also be useful in analysing the trend in IEM estimates. A change in IRC could be a due to relative change in the distribution of education over time. But, the ICC measures similarity in the rank of parent and child within the educational distribution of their respective cohorts.

Further, we decompose the ICC in order to identify the sources of persistence. In order to facilitate this decomposition, we make five ordinal categories for the level of educational attainment: 0 years (No education), 1-5 years (Primary education), 6-8 years (Middle education), 9-12 years (Secondary education), 13-16 years (College education). Parental education is the main explanatory variable, and its regression coefficient provides a measure of IEM. In addition, we also incorporate several control variables that may affect the education attainment of the child. These include age, household size, ethnicity, social status, city of residence (metro/non-metro), location (rural/urban), and state of residence.

Additionally, we also provide Lewbel-IV estimates which accounts for biasness of mobility estimates due to potential endogeneity of parental education variable (refer Section 3.4 in Chapter 3 for a description of the Lewbel-IV identification strategy).

4.3 Results

4.3.1 Regression based estimates of IEM

Table 4.3 reports the univariate regression estimates of mobility for full sample of father-son pairs across different birth cohorts of sons. For each univariate regression, we regress son's education on a constant and father's education, where years of schooling is used to measure education.

For the full sample (Panel A), all IRC and ICC estimates are positive and statistically significant; however, their magnitudes have declined over time, which indicates that IEM has increased over time. Nonetheless, even for the youngest birth cohort, father education still has a positive and significant association with son's education. The temporal increase in mobility is largely concentrated in the rural sector, whereas, the mobility estimates for the urban sector have remained stable over time.

Table 4.4 reports the magnitude of IEM using six different mobility indicators (refer section 3.3 of Chapter 3 for the definition of these mobility indicators). At an all- India level, the IRC estimate changes substantially and significantly over the time period while the ICC estimate does not see any major improvement.

Table 4.5 reports multivariate regression estimates of mobility for father-son pairs. We generate estimates using OLS-regressions and the synthetic instrumental variable method (IV-Lewbel) that accounts for potential endogeneity of father's education. Correlation coefficients report the standardized regression coefficients. All models are estimated with the complete set of controls specified in Table 4.2.

In all cases, we find that the estimated coefficient for father's education is positive and significant. For example, the IV-Lewbel estimate of ICC suggests that a one standard deviation increase in father's education increases son's years of schooling by 0.32 standard deviations. The bottom panel of Table 4.5 reports some diagnostic statistics to examine the validity of the IV-Lewbel approach. The Durbin-Wu-Hausman test indicates that father's education is potentially endogenous, and therefore OLS based mobility estimates are likely to be biased. The Breusch-Pagan heteroskedasticity test confirms the presence of heteroskedasticity in the error process. The Kleibergen-Paap LM statistic indicates that the synthetic instruments are relevant, i.e., the instruments are associated with parental education – the endogenous variable of interest. The Hansen J test suggests that overidentifying restrictions are valid. The results reported in both tables are consistent, although the magnitude of Lewbel IV estimates of IRC and ICC are smaller than the corresponding OLS estimates. This suggests that regression-based measures may potentially overestimate intergenerational educational persistence due to heterogeneity in unobserved characteristics such as ability and preference. Notwithstanding, the Lewbel IV models, which address the omitted ability bias and potential measurement error in parental education, also confirm significant intergenerational persistence.

The coefficients of the control variables suggest that membership of marginalized castes (OBC, ST, or SC) reduces the educational attainment of the child. Children residing in rural areas are less educated than their urban counterparts. We examine whether the degree of IEM varies across social groups by using interactions of father's education with social group dummy variables. In all models, these interaction coefficients are negative and significant, suggesting that membership of marginalized social groups reduces the IEM.

4.3.2 Matching based estimates of IEM

To examine the robustness of regression-based estimates we generate alternative estimates of IEM using matching estimators. We employ six matching estimators, namely, Nearest Neighbour (NN) matching, Mahalanobis matching, and four kernel based matching estimators based on the Epanechnikov kernel, the Biweight kernel, the Normal kernel, and the Uniform kernel. High father's education is used as the treatment based on three different thresholds: 1) If the father's education is primary & above it is coded as 1, and, zero otherwise; 2) If the father's education is middle & above it is coded as 1, and zero otherwise; 3) If the father's education is secondary & above it is coded as 1, and zero otherwise. Primary & above refers to at least 1 year of schooling, Middle & above refers to at least 6 years of schooling, and Secondary & above refers to at least 9 years of schooling.

With three different definitions for the treatment variable and six different matching methods, we estimate a total of 18 matching models, and the ATT estimates of these models are reported in Table 4.6. The ATT estimates represent the effect of receiving treatment on son's educational attainment. All matching methods are estimated with a covariate set comprising the full set of control variables specified in Table 4.2. For brevity, we only report the ATT estimates in Table 4.6. Detailed estimation results of the matching models including the tests of covariate balance are available in Tables A.1 to A.18 of the appendix.

All ATT estimates are positive and significant at 1% level of significance, and they remain largely stable regardless of the choice of the matching method or the choice of the education threshold used to define the treatment. Ceteris paribus, having a highly educated father increases son's educational attainment by at least 2 years. Thus, consistent with the regression-based estimates of IRC and ICC, the ATT estimates also indicate significant intergenerational educational persistence. To test whether the ATT estimates are robust to potential endogeneity of parental education, we conduct Rosenbaum's sensitivity analysis. The sensitivity analysis re-estimates the ATT reported in Table 4.6 under different assumptions for unobserved heterogeneity. The result of sensitivity analysis is available in supplementary appendix, Tables A.1 to A.18. We find that ATT estimates are robust to differential assignment to treatment due to unobserved factors such as genetic ability or parental attitude towards education.

4.3.3 The bias induced by the co-residence restriction

The co-residence restriction induces selection bias in the mobility estimates. This is because the co-resident sample over-represents younger, unmarried and less-educated individuals, who are more likely to reside in their parental households. Unfortunately, most household survey datasets suffer from data truncation due to the co-residence restriction. With the exception of the mother-son pairs, we are able to relax the co-residence restriction and generate full samples for all other parent-child combinations.

Table 4.7 reports the mobility estimates for all parent-child pairs and quantifies the bias induced in these estimates due to co-residence restriction. Panel A presents mobility estimates for the coresident sample and the full sample. Columns 1 and 2 (3 and 4) report the IRC (ICC) estimates for the full sample and the co-resident sample, respectively. The mobility estimates are based on the IV-Lewbel model estimated with all controls described in Table 4.2. The estimates remain robust, regardless of the method used to estimate mobility. For brevity of presentation, we do not report the coefficients of controls. Finally, we report tests of equality which compare the mobility estimates of the full sample and the co-resident sample.

For all parent-child combinations, there is evidence of significant intergenerational persistence (low IEM). The co-residence restriction induces a downward bias on the estimated coefficients, and the IRC estimates are more sensitive to this bias as compared to the ICC estimates.

In all comparisons, the tests of equality reject the null hypothesis that the mobility estimate for the full sample is equal to that for the co-resident sample. The results suggest that co-residence restriction induces a large and significant downward bias in the IRC and the ICC estimates. Thus, studies that do not correct for the co-residency bias are likely to underestimate IEM. However, the statistical rejection of the equality tests must be interpreted with some caution. With a large set of parent-child pairs, the standard errors of IRC and ICC estimates are quite small, and therefore, even minor differences in these estimates can lead to a rejection of null hypothesis in the equality tests.

Therefore, we opt a different approach to quantify the magnitude of this bias by calculating a measure of normalized bias in IRC estimates (and similarly for the ICC estimates) as follows

$$Normalized\ Bias = \left|\frac{(IRC_F - IRC_{CR})X\ 100}{IRC_{CR}}\right| \tag{4.4}$$

where $IRC_F\ and\ IRC_{CR}$ refer to IRCs estimated using the full sample and the coresident sample, respectively. Normalized bias, which is induced in the mobility estimates when the co-residence restriction is imposed, is reported in Panel B. In all comparisons, we find that the coresident samples induce downward bias on IRC and ICC estimates. For the IRC estimates, the bias is highest for mother-daughter pairs (30.42%) followed by father-son pairs (28.78%) and father-daughter pairs (25.79%). Unlike the IRC estimates, the ICC estimates obtained from the full sample are quite close to the corresponding estimates obtained from the coresident sample. The average normalized bias in ICC estimates is 5.84% whereas it is 28.35% for the IRC estimates. The highest bias observed for the ICC estimates (6.81% for the father-daughter pairs), is lower than the lowest bias observed for the IRC estimates (25.79% for the father-daughter pairs). The results suggest that the co-residence restriction can induce large downward bias in the IRC estimates, whereas, the ICC estimates are less sensitive to this bias. Thus, if it is not possible to control for the co-residence restriction in the data, the ICC is preferable over the IRC as the measure of intergenerational educational persistence.

4.4 Decomposition Analysis: Sources of IEM

The intergenerational transmission coefficient captures may hide the difference in mobility pattern across the educational distribution. A higher value of intergenerational transmission coefficient suggests lower mobility, but it does not reveal the direction of such mobility. Therefore, we identify the sources of intergenerational persistence by decomposing the ICCs following the approach of Checci et al. (2008). The results of ICC decomposition for father–son pairs across different birth cohorts of sons are reported in Table 4.8. Using Equation (3.7), the ICC for each child cohort has been disaggregated into 25 different components, which correspond to associations between to the five education levels of fathers, each associated with five education levels of sons. For example, the first panel of the Table 4.8 decomposes the ICC of the group of sons having non-literate fathers in five categories corresponding to five education levels of sons. The vertical sum of these figures in each column gives the correlation coefficient for the specific child cohort. The ICC has declined from 0.621 (oldest cohort) to 0.431 (youngest cohort). This implies that the IEM has increased over time. However, it is interesting to identify the source of persistence as this increase in mobility could be attributed to two reasons. First, it could be upward mobility due to less educated fathers who have highly educated children. Second, it could be downward mobility due to highly educated fathers having less educated children. Generally, upward mobility should be desirable as compared to downward mobility. However, upward mobility can exacerbate inequality in education attainment if it is primarily displayed by children of highly educated fathers but not by the children of less educated fathers.

The results reported in Table 4.8 indicate that the main source of persistence in IEM is the positive contribution of highly educated fathers having children who are either equally well-educated or better educated than their fathers. There is an increasing proportion of low-educated fathers having children who are also low-educated and declining proportion of low-educated fathers having children who are highly educated. The positive contribution to ICC by the group of sons with illiterate fathers has increased from 48% (oldest cohort) to 62% (youngest cohort), whereas for the group of sons with college educated fathers, it has declined from 37% to 27%. A key challenge in terms of IEM is the presence of high degree of persistence in the left tail of the educational distribution.

We find some improvements in mobility measures, as the magnitude of upward (downward) mobility has increased (decreased) over time. Nonetheless, there remains substantial persistence in both right and left tails of the educational distribution. The contribution due to pairs where both generations have below-mean education has steadily increased over time. Thus, the children with the least educated fathers have become less mobile over time. Similarly, the contribution from pairs where both generations have high education is large and accounts for over one-third of the overall persistence.

4.5 Conclusion

In this chapter, we analyzed the trends in IEM in India. We draw three major conclusion from the analysis. First, although the intergenerational persistence has decreased overtime; parental education remains significant predictor of child education. However, estimated transmission coefficnet does not inform us about the direction of mobility. Therefore, we also apply the decomposition technique to examine the direction of mobility over time. The ICC decomposition reveals that although the intergenerational persistence has declined over time, implying an increasing mobility; it remains sticky in the tails of the education distribution. The positive persistence at the upper end of the education distribution has increased from 43% to 61% over time, while negative persistence at the lower end has increased from 5% to 22%.

Third, we also try to obviate the issue of endogeneity by using heteroscedasticity-based identification of Lewbel (2012). The traditional instrumental variable approach is of limited use with secondary databases, as it is difficult to identify variables which are independent of error term and which do not affect the dependent variable when independent variable is held constant (exclusion restriction). We use an alternative identification strategy proposed by Lewbel (2012) which replaces endogenous regressors, such as parental education, with synthetic instrumental variables constructed using linear combinations of exogenous regressors. The major advantage of this identification strategy is that it does not rely on the standard exclusion restriction. After controlling for potential endogeneity of parental education level, we find that the OLS estimates are larger than the corresponding Lewbel IV estimates. This suggests that regression-based measures may potentially overestimate intergenerational educational persistence due to heterogeneity in unobserved characteristics such as ability and preference. Notwithstanding, the Lewbel IV models, which address the omitted ability bias and potential measurement error in parental education, also show a significant intergenerational persistence.

The results from matching techniques confirm the robustness of our regression based estimates. Results from Rosenbaum sensitivity analysis show that the effect of parental education is robust to different assumptions regarding the unobserved heterogeneity. The estimated ATT bounds show that the effect of parental education remain significant even after assuming substantial hidden selection bias.

Table 4.1: Descriptive Statistics of school enrolment, Age (20-29)

	Sample Size		**Number of children enrolled**		**Share enrolled (%)**		**Number of children enrolled[a]**		**Share enrolled[a](%)**	
	Age (20-24)	Age (25-29)	Age (20-24)	Age (25-29)	Age (20-24)	Age (25-29)	Age (20-24)	Age (25-29)	Age (20-24)	Age (25-29)
Father-Son	6,576	4,520	1,883	236	28.63	5.22	1,432	98	21.77	2.16
Father-Daughter	3,147	778	2,195	112	92.63	14.4	865	42	27.49	5.41
Mother-Son	6,398	4,334	1,861	231	29.09	5.33	1,414	97	22.10	2.24
Mother-Daughter	2,881	742	1,294	109	44.91	14.69	863	41	29.95	5.52

Notes: The superscript “a” refers to enrolled children who have not attained highest achievable education level (16 years).

Table 4.2: Variable Definitions

Variables	Definition
Child education	Completed years of education of the son.
Father education	Completed years of education of the parent.
Control variables	
Child age	Age of the child in years at the time of IHDS-II survey.
Child age^2	Square of Child age.
Father age	Age of the parent (either father or mother) in years at the time of IHDS-II survey.
Hindu (Omitted)	Equals 1 if the individual belongs to Hindu religion, 0 otherwise.
Muslim	Equals 1 if the individual belongs to Muslim religion, 0 otherwise.
Other religion	Equals 1 if the individual belongs to other religion, 0 otherwise.
Forward Caste (Omitted)	Equals 1 if the individual belongs to the forward caste including brahmins, 0 otherwise.
ST	Equals 1 if the individual belongs to the Scheduled Tribe category, 0 otherwise.
SC	Equals 1 if the individual belongs to the Scheduled Caste category, 0 otherwise.
OBC	Equals 1 if the individual belongs to the Other Backward Class category, 0 otherwise.
Urban (Omitted)	Equals 1 for urban households, 0 otherwise.
Rural	Equals 1 for rural households, 0 otherwise.
Metro (Omitted)	Equals 1 for households in metro areas, 0 otherwise.
Non metro	Equals 1 for households in non-metro areas, 0 otherwise.

Notes: We have taken the time-invariant variables as our explanatory variables in the regression analysis.

Table 4.3: Univariate regression estimates of IEM

Birth cohort (son)	ICC (ρ)	IRC (β)	Average years of schooling (son)	Average years of schooling (father)	Standard deviation of education (son)	Standard deviation of education (father)
			Total			
1947-1956	0.58	0.65	9.81	5.74	4.74	3.91
1956-1966	0.48	0.52	9.88	5.82	4.84	4.01
1967-1976	0.39	0.45	9.91	5.91	4.88	4.04
1977-1986	0.40	0.46	10.07	6.18	4.93	4.16
Overall	0.46	0.52	9.94	5.91	4.84	4.03
			Rural			
1947-1956	0.53	0.57	7.81	4.01	4.36	3.62
1956-1966	0.38	0.36	8.79	4.21	4.37	3.81
1967-1976	0.37	0.37	9.21	4.63	4.41	4.02
1977-1986	0.35	0.38	9.38	4.65	4.62	4.11
Overall	0.41	0.42	8.86	4.37	4.45	3.89
			Urban			
1947-1956	0.42	0.53	10.96	7.59	4.27	3.23
1956-1966	0.49	0.61	11.26	8.11	4.66	3.68
1967-1976	0.39	0.49	11.82	8.17	4.72	3.74
1977-1986	0.43	0.51	11.98	8.52	4.82	3.81
Overall	0.43	0.53	11.51	8.09	4.61	3.61

Notes: All Observations have been weighted using IHDS-II weights to reflect 2011 Indian population. The sample include coresident father-son pairs.

Table 4.4: IEM Using Different Mobility Indicators

	All	1947-56	1957-66	1967-76	1977-86	All	1947-56	1957-66	1967-76	1977-86
			IRC					*ICC*		
All India	0.604	0.738	0.645	0.553	0.524	0.552	0.621	0.474	0.413	0.431
Hindu	0.608	0.738	0.65	0.544	0.505	0.412	0.623	0.475	0.412	0.442
Non-Hindu	0.642	0.738	0.624	0.593	0.6	0.434	0.635	0.475	0.424	0.511
Rural	0.601	0.708	0.611	0.539	0.486	0.361	0.527	0.362	0.354	0.379
Urban	0.534	0.604	0.531	0.497	0.514	0.436	0.625	0.472	0.448	0.518
Forward Caste	0.495	0.608	0.515	0.433	0.413	0.429	0.641	0.484	0.423	0.442
Non-Forward Caste	0.609	0.719	0.616	0.545	0.521	0.382	0.572	0.412	0.368	0.431
			BUM					*UCP*		
All India	0.275	0.213	0.245	0.297	0.33	0.823	0.833	0.82	0.819	0.826
Hindu	0.286	0.22	0.252	0.311	0.344	0.833	0.854	0.828	0.822	0.839
Non-Hindu	0.224	0.177	0.21	0.227	0.264	0.766	0.728	0.778	0.798	0.753
Rural	0.225	0.149	0.178	0.255	0.296	0.747	0.739	0.691	0.772	0.748
Urban	0.385	0.363	0.377	0.388	0.408	0.869	0.865	0.873	0.85	0.881
Forward Caste	0.451	0.384	0.433	0.48	0.508	0.878	0.874	0.866	0.862	0.899
Non-Forward Caste	0.234	0.166	0.197	0.256	0.295	0.781	0.785	0.77	0.783	0.781
			M1					*M2*		
All India	4.361	3.821	4.169	4.669	4.6	3.777	3.893	4.047	3.651	3.389
Hindu	4.444	3.91	4.257	4.747	4.687	3.882	3.999	4.137	3.759	3.513
Non-Hindu	3.943	3.358	3.714	4.275	4.191	3.251	3.396	3.593	3.088	2.74
Rural	4.196	3.357	3.763	4.592	4.765	3.654	4.1	4.014	3.268	2.981
Urban	4.661	4.744	4.853	4.808	4.298	4.001	3.513	4.107	4.296	4.201
Forward Caste	4.887	4.883	5.03	5.034	4.605	4.261	3.881	4.315	4.493	4.392
Non-Forward Caste	4.211	3.484	3.908	4.567	4.599	3.639	3.896	3.972	3.395	3.071

Notes: This table reports estimates of IEM using six different mobility indicators. The sample includes father-son pair which also includes information on fathers who don't coreside with their children from IHDS-II dataset.

Table 4.5: Regression-based estimates of IEM

Variables	Ordinary Least Square		IV-Lewbel	
	Correlation coefficient (1)	Regression coefficient (2)	Correlation coefficient (3)	Regression coefficient (4)
Father education	0.415***	0.459***	0.327***	0.392***
	(0.005)	(0.007)	(0.040)	(0.053)
Child age	0.068***	0.019***	0.035***	0.010***
	(0.004)	(0.001)	(0.006)	(0.002)
Child age^2	-0.005***	-0.001***	-0.004***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Father age	0.006*	0.001*	0.022***	0.006***
	(0.003)	(0.001)	(0.004)	(0.001)
Rural	-0.474***	-0.134***	-0.238***	-0.068**
	(0.054)	(0.015)	(0.108)	(0.031)
OBC	-0.588***	-0.166***	-0.182**	-0.052**
	(0.059)	(0.017)	(0.083)	(0.024)
SC	-0.887***	-0.249***	-0.049	-0.014
	(0.071)	(0.020)	(0.132)	(0.038)
ST	-0.653***	-0.184***	-0.332*	0.095**
	(0.102)	(0.029)	(0.169)	(0.048)
Muslim	-1.112***	-0.313***	-0.481***	-0.138***
	(0.076)	(0.022)	(0.115)	(0.033)
Other religion	-0.162	-0.046	-0.256**	-0.073**
	(0.109)	(0.031)	(0.123)	(0.035)
Non metro	-0.433***	-0.123***	-0.103	-0.029*
	(0.106)	(0.029)	(0.123)	(0.035)
OBC × Father education	-0.019*	-0.005*	-0.013*	-0.002**
	(0.0116)	(0.003)	(0.009)	(0.001)
SC × Father education	-0.060***	-0.017 ***	-0.042***	-0.014***
	(0.014)	(0.004)	(0.003)	(0.002)
ST × Father education	-0.038**	-0.010*	-0.025**	-0.008**
	(0.022)	(0.066)	(0.013)	(0.004)
Observations	39,297	39,297	39,297	39,297
Adjusted R-squared	0.3198	0.3198	0.3121	0.3121

Table 4.5 (continued).

Diagnostic statistics for IV-Lewbel models

Breush-Pagan LM statistic	Kleibergen-Paap LM statistic	Hansen J statistic	Durbin-Watson Statistic
46.34^{***}	360.87^{***}	1.21	72.95^{***}

Notes: All models are estimated using the full sample of father-son pairs. The dependant variable is son's years of schooling. PSU-clustered standard errors reported in parentheses. *, **, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively. Observations have been weighted using IHDS-II weights to reflect the 2011 Indian population.

Table 4.6: ATT estimates of IEM

	Threshold for high vs low father's education (treatment)		
Matching Method	Primary & above vs. Below primary	Middle & above vs. Below middle	Secondary & above vs. Below secondary
Nearest Neighbour (NN)	2.156***	2.510***	2.900***
	(0.124)	(0.123)	(0.185)
Mahalanobis matching	2.104***	2.542***	2.689***
	(0.123)	(0.117)	(0.176)
Kernel-based matching estimators			
Epanechnikov kernel	2.271***	2.607***	2.871***
	(0.096)	(0.091)	(0.111)
Biweight kernel	2.262***	2.601***	2.864***
	(0.096)	(0.091)	(0.112)
Normal kernel	2.864***	2.641***	2.995***
	(0.112)	(0.088)	(0.108)
Uniform kernel	2.289***	2.617***	2.881***
	(0.095)	(0.090)	(0.111)

Notes: This table reports the ATT estimates representing the effect of father's education on son's educational attainment. The matching models are estimated using the full sample of father-son pairs, with a covariate set comprising the full set of control variables described in Table 1. PSU-clustered standard errors reported in parentheses. *, **, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively. Primary & above refers to at least 1 year of schooling, Middle & above refers to at least 6 years of schooling, and Secondary & above refers to at least 9 years of schooling.

Table 4.7: Mobility estimates for different parent-child pairs and the bias induced by the co-residence restriction

Panel A: IEM estimates for different parent-child combinations

	IRC			ICC		
	Full	Co-resident	Test of Equality(χ2)	Full	Co-resident	Test of Equality(χ2)
	(1)	(2)	(1) = (2)	(3)	(4)	(3) = (4)
Father-Son	0.604***	0.469***	85.43	0.552***	0.529***	12.66
	(0.040)	(0.007)		(0.053)	(0.008)	
Observations	39,297	9,496		39,297	9,496	
Father-Daughter	0.551***	0.438***	61.57	0.502***	0.470***	41.58
	(0.002)	(0.004)		(0.021)	(0.002)	
Observations	32,675	2,288		32,675	2,288	
Mother-Daughter	0.733***	0.562***	134.61	0.618***	0.581***	34.92
	(0.007)	(0.028)		(0.004)	(0.023)	
Observations	31,713	2,155		31,713	2,155	
Mother-Son		0.362***			0.411***	
		(0.008)			(0.009)	
Observations		7,225			7,225	

Panel B: Normalized bias induced in mobility estimates by imposing the co-residence restriction

	Bias (IRC)	Bias (ICC)
Father-Son	28.78%	4.34%
Father-Daughter	25.79%	6.81%
Mother-Daughter	30.42%	6.37%

Notes: All coefficients are based on the IV-Lewbel model estimated with the full set of control variables. Full sample is not available for the Mother-Son pairs. PSU-clustered standard errors reported in parentheses. *, **, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively. Observations have been weighted using IHDS-II weights to reflect the 2011 Indian population.

Table 4.8: Decomposition of ICC for father-son pairs

Child's education/birth cohort	1947-1956	1957-1966	1967-1976	1977-1986	Overall
Father: No Education					
C:No education	0.000	0.000	0.013	0.003	0.009
C: Primary	0.180	0.124	0.113	0.134	0.148
C:Middle	0.118	0.084	0.068	0.077	0.068
C:Secondary	0.001	0.017	0.023	0.009	-0.018
C:College	0.000	-0.028	-0.048	-0.046	-0.036
Total contribution	0.299	0.197	0.169	0.176	0.171
Father: Primary					
C:No education	0.000	0.000	0.000	0.000	0.000
C: Primary	0.024	0.020	0.017	0.015	0.009
C:Middle	0.006	0.009	0.010	0.011	0.005
C:Secondary	0.000	0.003	0.005	0.002	-0.002
C:College	-0.015	-0.016	-0.015	-0.013	-0.006
Total contribution	0.015	0.017	0.017	0.016	0.007
Father: Middle					
C:No education	0.000	0.000	-0.001	-0.001	-0.001
C: Primary	0.000	-0.005	-0.006	-0.008	-0.013
C:Middle	-0.007	-0.007	-0.005	-0.009	-0.011
C:Secondary	0.000	-0.003	-0.004	-0.002	0.076
C:College	0.018	0.025	0.017	0.020	0.077
Total contribution	0.010	0.009	0.001	0.000	0.128
Father: Secondary					
C:No education	0.000	0.000	-0.001	-0.001	-0.001
C: Primary	0.000	-0.017	-0.009	-0.009	-0.012

Table 4.8 (continued)					
C:Middle	-0.008	-0.002	-0.011	-0.013	-0.013
C:Secondary	-0.003	-0.009	-0.014	-0.007	0.065
C:College	0.119	0.14	0.111	0.126	0.151
Total contribution	0.108	0.112	0.075	0.097	0.111
		*Father**:** College*			
C:No education	0.000	0.000	0.000	0.000	0.000
C: Primary	0.000	-0.008	-0.002	-0.001	-0.001
C:Middle	0.000	0.000	-0.004	-0.003	-0.002
C:Secondary	-0.001	-0.004	-0.007	-0.002	0.035
C:College	0.189	0.161	0.175	0.149	0.164
Total contribution	0.189	0.150	0.162	0.143	0.135
Correlation Coefficient	0.621	0.474	0.413	0.431	0.552

Notes:

The education variable is converted into discrete education levels.

No education: 0 years; Primary: 1-5 years; Middle: 6-8 years; Secondary: 9-12 years; and College: 13-16 years.

Chapter 5

Intergenerational Educational Mobility among Migrants

5.1 Introduction

Indian society is characterized by large socioeconomic disparities due to caste, religion and gender based social stratification (Munshi & Rosenzweig, 2016) and an uneven distribution of economic growth (Chaudhuri & Ravallion, 2006). These socioeconomic disparities persist over generations and create a barrier to IEM. IEM of education among Indian children is a well-studied topic (see, for example, Asadullah & Yalonetzky, 2012; Azam, 2016; Azam & Bhatt, 2015; Borkotoky et al., 2015; Choudhary & Singh, 2017; Emran et al., 2020; Emran & Shilpi, 2015). We describe the role of migration experience on IEM in India.

The effect of migration on intergeneration educational mobility is complex. Borjas (1989) shows that, among migrants, IEM is moderated by the human capital of their ethnic community. If the migrant population faced difficulty in integrating with the local society, the parental education level plays a dominant role in determining their offspring's education level. Further, migrants often lack access to public resources available to the natives that acts as a barrier and prevents their children from climbing up the social ladder. Schneebaum et al. (2015) observes that migrants have limited access to public resources, due to information asymmetry and non-membership of local social networks. As a result, educational attainment of migrant children depends more on private investments, such as household assets or parental education level, rather than public investments (Ammermueller, 2007; Schneeweis, 2011). Despite these disadvatages migrant children can display higher upward IEM than their native-born counterparts (Farley & Alba, 2002; Luthra & Soehl, 2015). Several explanations have been proposed for explaining this paradox such as parental self selection and differences in educational aspirations, expectations and cultural values such as long term orientation between migrant and native populations (Beine et al., 2020; Feliciano, 2005; Feliciano & Lanuza, 2017; Figlio et al., 2019). A different but related strand of literature examines whether migration can be a substitute for education in the pursuit of improving economic status. Ward (2020) found that internal migration led to substantial economic gains in the United States. On an average the economic gains of migration were more than three times larger than those obtained through one year of additional education. These contrasting findings suggest that further study of the role of migration experience on child education is warranted. We use a novel strategy to identify children who experienced migration when they were enrolled in school to examine the role of migration experience on IEM in India.

To our knowledge, the effect of migration on IEM in India remains unexplored. This is surprising given the vast population of migrant workers in India. The Census of 2011 found that there are 13.9 crore internal migrants in India.

Most of the migrant workers start working at a very young age, usually experience no (or even downward) economic mobility, and are engaged in low-paying informal sector jobs for their entire work-life (Sharma, 2017). With manual labour taking its toll, and poor access to public health services, migrant workers are often forced to go back to their hometown due to health problems. This lowers their household income, forcing their children to start migrating for work at a relatively young age. This vicious cycle has perverse intergenerational implications for health, education, and economic outcomes of the children.

In India, migrant children face learning barriers due to substantial language and cultural differences, and quite often a change of curriculum when migrating across states. In addition, migrant households have limited access to public resources as state government jobs and local social welfare schemes tend to have significant reservations for the native residents of the state. For example, migrant families that are unable to produce required proof of identify and residence would be unable to receive social welfare entitlement such as subsidized food under public distribution system. Therefore, a lack of access to public resources tends to make migrant children more dependent on private investments by their parents. Conversely, migrating to more developed areas may improve the quality of available school facilities for migrant children. Zhou & Xu (2017) document substantial upward mobility among children of rural-to-urban migrants, and found that longer stays in urban areas improves the educational outcome for migrant children. Internal migration also provides economic gains for migrants (Bhavnani & Lacina, 2017), and enables them to afford better quality education for their children (Bouoiyour et al., 2016).

Using data from the IHDS, we estimate interngenrational educational persistence among 39,297 father-son pairs. In addition, we provide new evidence on the impact of migration experience on IEM, which is the main contribution of this study. Most of the extant works on IEM in India relies on household survey data from the NSS, and more recently from the IHDS. Since these household surveys impose a coresidency restriction to identify members of the household, the mobility estimates suffer from coresident sample selection bias (Asadullah & Yalonetzky, 2012; Emran & Shilpi, 2015; Hnatkovska et al., 2013). Using datasets from Bangladesh and India, Emran et al. (2015) found that mobility estimates are generally underestimated due to coresidency restriction characterizing most household survey datasets. We follow Azam and Bhatt (2015) to remove coresidency bias from IHDS data. We use additional information available in the IHDS data that allows us to identify the education of father, even when he does not reside in the same household as his son[1]. Thus, our IEM estimates are robust to any selection-bias induced by the coresidency restriction.

Establishing a causal effect of parental education is problematic as both parental and child education are influenced by selection on unobserved variables like genetic ability,

[1] Similar procedure to relax coresidency restriction in IHDS data has been used by Azam and Bhatt (2015). They show that the coresidency restriction leads to a substantial loss of observations and an overrepresentation of younger individuals who tend to reside with their parents.

parental attitude towards education or social networks. For example, the traditional ordinary least squares (OLS) estimator can overestimate intergenerational educational persistence due to selection on unobserved ability. We use two strategies to address potential endogeneity problems. First, we implement heteroscedasticity-based identification strategy of Lewbel (2012) which uses heteroskedastic covariance restriction to generate internal instruments when standard exclusion restrictions are not available. These internal instruments are then used in the two-stage instrumental variable regression to provide a causal interpretation to IEM estimates. Second, we use matching estimators to generate alternative estimates of mobility. Then, we implement Rosenbaum's sensitivity analysis (Rosenbaum, 2002) to check the robustness of these matching estimates to selection on unobserved heterogeneity. In the present chapter, we examine how exposure to migration during schooling has affected the educational attainment of an individual by family background, where the latter is proxied in terms of parental educational attainment.

The remainder of this chapter is organized as follows. Section 5.2 describes the dataset and empirical methodology. In Section 5.3, we study the degree of IEM among migrants. Section 5.4 presents the decomposition analysis to identify the source of persistence. Finally, last section presents concluding remarks.

5.2 Dataset and Methodology

5.2.1 Dataset

We use data from both rounds of IHDS surveys, namely, IHDS-I (2004-05) and IHDS-II (2011-12). We construct a panel dataset by combining both waves of IHDS surveys by using the households that were interviewed under both rounds.

As discussed in the previous chapter, we use a sample of individuals aged between 25 and 65 years, who are expected to have completed their education. The survey contains question about the migration history of the households. The panel information on household and individual characteristics allows us to identify children who have had any migration experience during schooling. IEM is being examined for all the parent-child pairs. However, our focus is on estimating the degree of heterogeneity in IEM estimates for father-son pair, for both non-migrants and migrants. For father-son pairs, we are able to identify migrant children without imposing coresidency restriction. However, for other parent-child pairs, the estimates are obtained after imposing coresidency restriction. Therefore, the results for other parent-child pairs (mother-son, mother-daughter and father-daughter) are to be viewed as purely suggestive as their mobility estimates may be biased owing to small sample size and co-residency restriction.

Variable definitions are provided in Table 5.1. The dependent variable used in our analysis is the completed years of education of sons aged between 25-65 years at the time of IHDS-2, and for whom father's education was available. Substantial evidence suggests that socioeconomic features, such as caste, ethnicity, and household resources have a significant relation with the educational attainment of the child. For example, children from rural areas are less-educated as compared to their urban counterparts (Hnatkovska et al., 2013). Therefore, we use caste, religion, rural residence, and household assets[2] as socioeconomic control variables.

[2] Household assets is a score ranging from 0 (lowest) to 33 (highest) representing the goods owned by the household (such as electric fan, refrigerator, pressure cooker, TV etc.) and the quality of housing (such as cemented roof, separate kitchen, piped indoor water). IHDS contains three principal measures for the economic status of household –income, consumption expenditure, and assets. Of these three, household assets is the least volatile and it also measures the household wealth accumulated over several years. The income and expenditure measures show large year-on-year fluctuations especially in rural areas with farm based income, and they are also subject to measurement errors due to underreporting (Hurst et al., 2014)

Age of the child at migration serves as a proxy for many migration related factors, such as differences in language and culture, that can affect the educational outcomes of migrant children (Basu, 2018). Gonzalez (2003) documents a non-linear relationship between age at migration and educational attainment. It is useful to examine whether these findings can be extended in the case of internal migration. Among migrant children, we examine the effect of 'age at migration', and its square, on child's education. We define four categories for the type of migration, namely, urban-to-rural, rural-to-urban, rural-to-rural, and urban-to-urban[3]. The type of migration may influence child education. For instance, children who migrate from rural to urban regions may get better access to schooling as compared to those who migrate from rural to rural regions.

5.2.2 Identification of Migrants

We measure the degree of IEM among children of migrant households. This analysis identifies migrants as those children who migrated between the two survey rounds of IHDS: IHDS-I and IHDS-II, and who were enrolled at the time when IHDS-I survey was conducted but had completed their education before the time when IHDS-II survey was conducted. This sample of children allows us to measure the impact of migration on IEM of education. Although, the IHDS surveys provide enrollment status at the time of the survey, this variable is sparsely populated. To overcome this limitation, we use the following step-by-step approach. First, using the dataset of IHDS-I survey, we identify the starting year of education as the 2005 – number of years of education reported in IHDS-I survey. Second, we identify the ending year of education as 2005 + (number years of education in IHDS-II - number years of education in

[3] IHDS surveys contain information about the current place of residence and the last place of residence from where the family have permanently migrated.

IHDS-I). Second, we select only those children for which the ending year of education is less than 2012, which implies that they finished their education before 2012. Third, we identify children whose families migrated when they were enrolled. IHDS-II includes a question "How many years ago did your family first come to this village/town/city?" This allows us to calculate the year of migration for the household. Finally, we identify children whose families migrated when they were enrolled using the following rule: the starting year of education is less than the year of migration, and the year of migration is less than the ending year of education. As mentioned earlier, we also classify the type of migration into four categories based on whether the last and current place of residence belong to rural or urban areas.

We identify a set of 39,297 father-son pairs (including those who do not co-reside in the same household) for which data is available for all variables specified in Table 1). Henceforth, referred to as the full sample. In 1,229 pairs, the child had some migration experience during schooling, and in the remaining 38,068 pairs, the child had no migration experience during schooling. Henceforth, we refer to the former set as the migrant sample and latter set as the non-migrant sample. Table 5.2 reports the descriptive statistics for these samples and provides a comparison between non-migrants and migrants.

A comparison of mean years of schooling of parent and their children shows that education level has substantially increased across generations for both non-migrants and migrants. On an average, the migrant children are younger and they have more years of schooling then the non-migrant children. Since younger cohorts tend to be more educated due to a general increase in educational attainment over time, this finding is expected. This is because the set of migrant children is limited to those enrolled at the time of IHDS-1 (year 2004-05), whereas this restriction is not imposed on the set of non-migrant children. Likewise, the parents of migrant children are younger and more educated than those of the non-migrant children. The socioeconomic characteristics of both groups are similar. There is little difference between the household assets of non-migrants and migrants, and about 61 percent of parent-child pairs belong to rural areas, for both groups. The distribution across different castes and religions is approximately similar for both groups, with one exception[4]. The average age at migration among migrant children is 12.6 years. Majority of migrants relocate to urban areas – around 45 percent migrated from urban to urban areas, 34 percent migrated from rural to urban areas, and only 7 percent migrated from urban to rural areas. Thus, while migrant children face significant hurdles due to abrupt changes in language and curriculum, they could also potentially reap benefit from better quality education in the urban regions (Zhang, 2017).

5.2.3 Methodology

To estimate IEM among non-migrants and migrants, we use Ordinary Least Square (OLS) method. To control for endogeneity of parental education, we apply Lewbel-IV method. These methods have been described in chapters 3 and 4. Similarly, we also apply the decomposition technique to identify the source of persistence among migrants. However, the decomposition analysis has been restricted to father-son pairs. The small sample size for other parent-child pairs prevents us from looking at the source of persistence among them.

Rosenbaum's sensitivity analysis

We also use matching estimators to generate alternative estimates of mobility, by representing parent education as a binary treatment that equals 1 if parental education exceeds a specific threshold, and 0 otherwise. We use three thresholds for defining high parental education: primary and above (father education > 0 years), middle & above (father education

[4] 40.5 percent of the non-migrants belong to other backward castes (OBC) whereas the corresponding proportion for migrants is 35.6 percent (significant at 1% level).

> 5 years), and secondary & above (father education > 8 years). The matching estimators relax the linear functional form assumption of usual regression specification and allow us to examine the common support assumption and covariate balance for all observed characteristics. Although matching estimators provide alternate estimates of mobility to examine the robustness of regression-based IEM estimates, they do not confirm causality of the relation. This is because matching estimators assume treatment allocation is random conditional on a set of observed characteristics. However, in the presence of selection on unobserved characteristics, such as ability, the treatment effect estimates obtained from matching estimators may suffer from hidden selection bias. We implement Rosenbaum's sensitivity analysis (Rosenbaum, 2002) to test whether the treatment effect estimates obtained from these matching estimators are robust to unobserved heterogeneity. In this procedure, a sensitivity parameter Γ denotes the effect of unobserved characteristics, such as genetically transmitted ability, on the odds ratio of receiving treatment (high parental education). $\Gamma = 1$ means that there is no selection on unobserved characteristics, whereas $\Gamma = 1.5$ implies that selection on unobserved characteristics increases the odds of receiving treatment (high parental education) by 50 percent. Thus, higher values of Γ represent more severe hidden selection bias. The sensitivity analysis procedure estimates the treatment effects under different assumptions of the magnitude of the hidden selection bias. The treatment effect estimates are highly sensitive to selection on unobserved characteristics if the qualitative conclusions change with Γ being marginally greater than 1, whereas they are insensitive if conclusions change only for very large values of Γ.

5.3 Results

5.3.1 IEM among Migrants

To examine the role of migration background in IEM, we identify children who have had some migration experience during their schooling years. The step-by-step procedure outlined in sub-section 5.2.3 is used to identify migrant children, i.e., children whose family migrated during their schooling. We identify two samples of father-son pairs, namely, the migrant sample and the non-migrant sample. The migrant sample includes those father-son pairs where the family migrated during son's education, and other pairs comprise the non-migrant sample.

OLS-regression estimates are reported in Table 5.3. We find that membership of socially disadvantaged castes and minority religion has a negative association with child education, whereas household assets have a positive association with child's education. These results corroborate prior evidence on educational inequalities in India based on caste, religion and economic status (see, for example, Borooah & Iyer, 2005; Desai & Kulkarni, 2008; Deshpande, 2000).

For both non-migrants and migrants, the measures of intergenerational persistence in education are statistically significant. The IRC estimates imply that a one-year increase in schooling of father improves son's attainment by 0.489 years for non-migrants, and by 0.386 years for migrants. While IRC measures the relation between educational attainment of both generations, ICC measures the similarity between parents and children in terms of their rank in the education distribution for their generation. For example, an ICC of one implies that there is no positional mobility, i.e., both son and father have identical ranks in the education distribution of their generation. A similar pattern is observed with the ICC estimates where we find that a one-standard-deviation increase in father education increases son's education by

0.405 (0.326) standard deviations among non-migrants (migrants). Thus, both measures suggest that non-migrant children are relatively less mobile than migrant children. The tests of equality show that this difference in mobility is statistically significant.

Table 5.4 re-estimates OLS-regressions reported in Table 5.3 using the instrumental variable method of Lewbel (2012), hereafter referred to as Lewbel IV. The results reported in both tables are consistent, although the magnitude of Lewbel IV estimates of IRC and ICC are smaller than the corresponding OLS estimates. This suggests that regression-based measures may potentially overestimate intergenerational educational persistence due to heterogeneity in unobserved characteristics such as ability and preference. Notwithstanding, the Lewbel IV models, which address the omitted ability bias and potential measurement error in parental education, also confirm intergenerational persistence.

The diagnostic tests suggest that the Lewbel IV models are well specified. The Breusch-Pagan/Cook-Weisenberg test statistic confirms that the exogeneous variables are heteroscedastic, which validates the heteroscedastic errors assumption of the Lewbel IV estimation[5]. In all Lewbel IV models, the instruments pass tests for relevance (using the Kleibergen-Paap rank Wald test) and exogeneity (using the Hansen J-test for overidentifying restrictions). The Durbin–Wu–Hausman test indicates that father's education is potentially endogenous, and therefore the Lewbel IV estimates are likely to be more reliable than the OLS estimates.

Next, we estimate IEM using a set of matching estimators. We emphasize that matching estimators are only used to examine the robustness of regression-based estimates, and they do not address the potential endogeneity issues such as selection on unobserved ability. We use

[5] The Breusch-Pagan/Cook-Weisenberg tests reported in the paper are based on the set of all exogenous variables. To investigate the source of heteroscedasticity in our data, we also conducted the test with each individual variable (not reported). We find that most of the heteroscedasticity comes from variables for age, household assets and the rural dummy.

high father's education as the treatment which is 1 if father's education is higher than a specified threshold, and 0 otherwise. The ATT (average treatment effect on treated) estimates for 18 matching models with six matching estimators and three thresholds for high father's education are reported in Table 5.5. The ATT estimates represent the effect of receiving treatment on son's educational attainment. All matching methods are estimated with a covariate set comprising all controls specified in Table 5.1.

We find that all ATT estimates are positive & statistically significant, and they remain largely stable regardless of the choice of the matching method or the choice of the high education threshold used for fathers (treatment). Ceteris paribus, having a highly educated father increases the son's educational attainment by at least 2 years. The size of the effect generally increases when the threshold for high father's education is increased. For example, with the Mahalanobis estimator, we find that for non-migrant (migrant) children, attainment increased by 2.104 (2.077) years when the threshold is primary and above and by 2.689 (2.474) years when it is secondary and above.

Parental education is defined using discrete indicator variables in matching models, whereas it is defined as a completed years of schooling in the regression-based models. Nonetheless, the findings from both approaches are largely similar. Similar to the estimates reported in Table 5.3 (OLS) and Table 5.4 (Lewbel IV), the ATT estimates also indicate significant intergenerational educational persistence. In addition, the ATT estimates for migrants are generally smaller than the corresponding estimates of non-migrants, suggesting that migrants are more mobile.

To test whether the ATT estimates are robust to potential endogeneity of parental education, we conduct Rosenbaum's sensitivity analysis. The sensitivity analysis re-estimates the matching models reported in Table 5.5 under different scenarios with varying levels of unobserved heterogeneity. The parameter is Γ (≥ 1) is a measure of the magnitude of hidden selection bias, with $\Gamma = 1$ representing the baseline model assuming no hidden selection bias. For each value of Γ, the analysis reports the significance levels of the lower and upper bounds of the ATT estimate. Since the estimated effect is positive and the selection on unobserved characteristics is likely to result in an upward bias, we focus only on the significance of upper bound of the ATT estimates. The results for the Rosenbaum sensitivity analysis are presented in Table 5.6. Γ is the log odds of selection into treatment (high parental education) due to unobserved factors and Sig+ is the *p*-value representing significance of the upper bound estimate of ATT. For all models, the estimates of ATT are significant at 5 percent level with $\Gamma = 1$ to 1.8, and even at $\Gamma = 2.0$ all ATT estimates are still significant at 10 percent level. The results show that the positive effect of high father's education is robust to selection on unobserved ability. Thus, results from both Lewbel IV regressions and the Rosenbaum sensitivity analysis suggest that the estimated intergenerational links are robust to endogeneity concerns.

Finally, we examine the impact of migration characteristics (refer Table 5.7) on attainment of migrant children. The models reported in Table 5.3 are re-estimated by including migration characteristics as additional independent variables. The coefficient for age at migration is positive, whereas it is negative for its square term. Similar negative and convex relationship between cage at the time of migration and attainment has been observed in studies on immigrant children (see, for example, Basu, 2018). Regardless of age at migration, migrant children in India can potentially face abrupt changes in language, culture and curriculum that can affect their educational outcomes. Our findings suggests that children that migrate at a young age are the worst affected in terms of their educational attainment. Hu and Szente (2010) study the challenges faced by young Chinese migrant children and make similar observations. They observe that older migrant children tend to overcome some of these challenges by demanding equal opportunities and social equity, whereas the youngest children tend to have little say in the early childhood education they receive. The younger migrant children also face these disadvantages for a greater proportion of their schooling experience as compared to older migrant children. While our findings are similar, we also document a beneficial impact of migration experience. Regardless of their place of origin, children who migrated to urban areas have higher educational attainment that those who migrated to rural areas. This finding indicates that migrant children could potentially benefit from better quality education in the urban regions as noted by Zhang (2017).

5.3.2 Gender, Migration and IEM

In this section, we study the impact of migration and gender on IEM. To make the results comparable across different parent-child pairs, we limit our analysis to co-resident parent-child pairs. Our migrant sample comprise of around 3% of the total sample size of father-son pairs. We have information on educational attainments for 9,496 father-son pairs- that is, with data available on at least one male individual in each generation for two generations, and who are residing in the same household. For migrant sample, there are 302 households where educational attainment information is available for at least one male member of the household in two consecutive generations.

We study how the intersection of gender and migration background affects mobility. This design is motivated by some earlier results that found relation between migration background and IEM differs for men and women. For example, Schneebaum et al. (2015) found that migrant men tend to be the more mobile then migrant women. Examining IEM among Canadian migrant , Abada and Tenkorang (2009) found that boys have a lower likelihood of obtaining higher education than girls, and Aydemir et al. (2009) found that migrant women are the most mobile group, especially relative to their mother’s education.

Intersection of gender and migration background provides a more nuanced view of their combined effect on mobility. Additionally, we examine the role of gender of both parents and their children by examining all four parent-child pairs. Overall, this results in eight mutually exclusive identities for children and their mobility estimates are reported in Table 5.8. Due to a limited number of observations for migrant parent-child pairs, the instrumental variable approach could be reliably estimated only for migrant father-son pairs. To facilitate comparison across all eight sets of parent-child pairs, we restrict the results presented in Table 5.8 to OLS estimates that provide a conservative estimate of the level of intergenerational education persistence.

The results suggest significant intergenerational persistence, and there is evidence that IEM follows gender lines. We find that father's (mother's) education correlates more to education level of sons (daughters) than that of daughters (sons).

In all comparisons, the coefficients estimated for the migrant sample are lower than those estimated for the non-migrant sample, which suggests that migrant children are more mobile than non-migrant children. Migrant daughters are particularly mobile relative to their fathers. It should be noted as a caveat that since these estimates are based on co-resident samples, which are available for all four types of parent-child pairs, it is possible that these estimates are biased due to over-representation of young unmarried daughters that co-reside with their parents.

Irrespective of whether the father-son sample is full sample (without imposing coresidency restriction) or coresident sample, all the above results indicate natives are less mobile than migrants.

5.4 Decomposition Analysis: Sources of IEM

Table 5.9 shows the univariate estimates of ICC for migrant and non-migrant sample. Findings show that natives are less mobile than migrants. However, the intergenerational transmission coefficient provides an incomplete representation of IEM. For instance, migrant children may be upward mobile simply because their parents have considerably lower attainment than those of native children. Moreover, lower value of intergenerational transmission coefficient implies greater mobility but doesn't reveal the direction of mobility. We decompose the ICC at different levels of father education to present a more complete representation of IEM. In doing so, we compare the mobility pattern of non-migrants and migrants whose fathers have attained same education level.

We identify the sources of intergenerational persistence by decomposing the ICCs following the approach of Checci et al. (2008). Table 5.10 shows decomposition of ICC for father–son pairs across different birth cohorts of sons. Using Equation (3.7), the correlation coefficient for each child cohort has been divided into 25 different components, which correspond to associations between to the five education levels of fathers, each associated with five education levels of sons. For example, the first panel of the Table 5.10 decomposes the ICC of the group of sons having non-literate fathers in five categories corresponding to five education levels of sons. The vertical sum of these figures in each column gives the correlation coefficient for the specific child cohort. For migrant sample, small sample size for older birth cohorts restricts us to perform the decompostion analysis for only three birth cohorts: 1967-1976, 1977-1986 and 1987-1996.

The ICC has declined from 0.507 (oldest cohort) to 0.373 (youngest cohort). This implies an increase in IEM over time. However, it is interesting to identify the source of persistence as this increase in mobility could be attributed to two reasons. First, it could be upward mobility due to fathers who are illiterate or who have attained primary education and whose children have a high level of education. Second, it could be downward mobility due to fathers who have a high level of education and whose children have achieved education level which is below the mean education level for their own cohort. Generally, upward mobility should be desirable as compared to downward mobility. However, upward mobility can exacerbate inequality in education attainment if it is primarily displayed by children of highly educated fathers but not by those of less educated fathers.

The main source of persistence in IEM is the positive contribution of highly educated fathers having children who are either equally well-educated or better educated than their fathers. There is an increasing proportion of low-educated fathers having children who are also low-educated and declining proportion of low-educated fathers having children who are highly educated. The total positive contribution to the ICC from "sons with non-literate fathers" has increased from 48% (oldest cohort) to 62% (youngest cohort), whereas for sons with college educated fathers, it has declined from 37% (oldest cohort) to 27% (youngest cohort). The decline in the proportion of positive contribution to the ICC at right tail of the educational distribution is somewhat offset by an increase in contribution for sons with secondary educated fathers – 19% (oldest cohort) to 40% (youngest cohort). Therefore, the primary challenge in terms of IEM is the presence of high degree of persistence for sons with less educated fathers.

Table 5.11 sheds light on how migration affects the intergenerational persistence by looking at different components of persistence. Among migrant father-son pairs, we find that that a substantial source of persistence emanates from IEM from highly-educated fathers to their highly-educated sons. Pairs where both generations have a college education contribute 0.216 which is 58% of the estimated ICC. There is some evidence of upward mobility – pairs where the father has no education and sons have a college education reduce ICC by -0.047, approximately -12% of the estimated ρ. Overall, fathers with low education have a total contribution of 0.137 (36.8% of ρ) and those with high education have a contribution of 0.146 (39.2% of ρ). Thus, the two extreme levels of father's education contribute around three quarters of the overall persistence, whereas persistence is weak when father has some intermediate level of education (1 to 12 years of schooling). For example, fathers with middle education (6 to 8 years of schooling) have the lowest total contribution of 0.008 (2.1% of ρ).

Among non-migrant father-son pairs, the decomposition pattern is similar, however, there are two notable differences. Overall, the estimated ρ for non-migrants is 0.427 whereas it is 0.372 for migrants suggesting that migrants tend to be more mobile than non-migrants. Non-migrant pairs where both generations have a college education contribute 0.134 (31.4% of ρ), which is significant but much smaller than in case of migrants (58.1% of ρ). For non-migrants, the largest contribution to persistence (0.148, 34.7% of ρ) comes from fathers with no education and sons with primary education. The corresponding figure for migrants is 0.056 (15.1% of ρ). Thus, there is significant persistence at both tails of educational distribution among non-migrants.

5.4 Conclusion

Despite a large population of internal migrants in India, the effect of migration on IEM in India remains largely unexplored. This study presents new evidence that migration experienced during schooling affects IEM. We find that migrant children display lower intergenerational persistence than non-migrants and this difference is statistically significant. A decomposition of ICC reveals that migrants have a greater likelihood of being downward mobile than non-migrants. For both groups, most of the persistence emanates from tails of the educational distribution, either where both generations have low education or where both generations have high education. This represents two separate failures of the Indian education system: the former indicates that the education policies fail to offset for the deficiencies in parental educational inputs; the latter highlights difficulty in obtaining college education for children with less-educated parents.

Migrant children in India face substantial challenges due to abrupt changes in language, curriculum, and culture. Using "age at the time of migration" as a proxy of these migration related factors, we find that those children who migrate at a young age fare worse in terms of their eventual attainment. Younger migrant children face these disadvantages for a greater proportion of their schooling experience, and unlike older children, they are unable to overcome some of these challenges by demanding equal opportunities and social equity. We recommend some policy measures to address the educational disadvantages faced by migrant children and ensuring better assimilation of migrant children at host areas.

Due to a lack of local proofs of identify and residence, migrant families are often unable to receive social welfare entitlement such as subsidized food under public distribution system. These requirements also hinder access to banking services, and connections for utility services such as cooking gas. A government initiative that integrates access to all public services in a single platform that uses the Aadhaar information for authentication will obviate the need for local proofs of residence and identity. Second, policy response needs to address the vicious cycle where poor occupational health of migrant parents eventually forces their children to migrate at a young age, usually working in similar unskilled, low-paying jobs as their parents with poor prospects of upward mobility. Government investment in affordable public health services, especially focused in urban areas with the highest concentration of migrant population would be well placed. Third, there is a need to improve the effectiveness of the existing legal framework in resolving informal sector disputes. Migrant workers routinely face workplace disputes related to non-payment of wages, compensation for workplace accidents and even deaths. A significant institutional reform could be establishing a National Commission for

Migrant workers that represents the right of migrant workers and provides advisory to state and central governments on all policy matters that affect migrant workers. Fourth, the government educational initiatives should aim to sensitize school authorities about the various disadvantages faced by migrant children. Migrant children face added challenges of adapting to a new learning environment with different linguistic and academic practices, which leads to an increase in dropout rates among them. Dropout rates among migrant children can be reduced if schools make concerted efforts to improve awareness of migrant parents regarding the economic benefits of educating their children and the different support schemes available to them. For example, schools may have policies that provide financial support to children belonging to low-income families, but often the migrant families are not able to avail them due to language barriers, lack of familiarity with the administration process (Ainscow & Hargreaves, 2016) or due to the stigma attached to claiming financial support (Baumberg Geiger, 2016). Simple measures such as translating standard school textbooks in all major regional languages and providing an open access to them through knowledge portals can help migrant children overcome the linguistic barriers. School administration should also focus on preventing discrimination against migrant children and make efforts towards promoting community cohesion to enable better integration of students from different backgrounds.

Table 5.1: Variables Definition

Variables	Definition
Child education	Completed education for son (years)
Father education	Completed education for father (years)
Child age	Child age (years)
Child age^2	Child's age square
Father age	Father age (years).
Hindu (Omitted)	Equals 1 for Hindus, 0 otherwise.
Muslim	Equals 1 for Muslims, 0 otherwise.
Other religion	Equals 1 for other religions, 0 otherwise.
Forward Caste (Omitted)	Equals 1 for forward caste including brahmins, 0 otherwise.
ST	Equals 1 for Scheduled Tribes, 0 otherwise.
SC	Equals 1 for Scheduled Castes, 0 otherwise.
OBC	Equals 1 for Other Backward Classes, 0 otherwise.
Urban (Omitted)	Equals 1 for urban households, 0 otherwise.
Rural	Equals 1 for rural households, 0 otherwise.
Household Assets	Score between 0 (lowest) to 33 (highest) representing number of household assets and the quality of housing.
Migration Characteristics	
Age at migration	Child age at migration (years).
Age migration^2	Square of age at migration.
<u>*Type of Migration*</u>	
Rural to Rural (Omitted)	Equals 1 if the household migrated from rural to rural region, 0 otherwise.
Rural to Urban	Equals 1 if the household migrated from rural to urban region, 0 otherwise.
Urban to Urban	Equals 1 if the household migrated from urban to urban region, 0 otherwise.
Urban to Rural	Equals 1 if the household migrated from urban to rural region, 0 otherwise.

Table 5.2: Descriptive Statistics

	Full Sample $\hat{\theta}_0$	Non Migrants $\hat{\theta}_1$	Migrants $\hat{\theta}_2$	Test of equality $\hat{\theta}_1 - \hat{\theta}_2 = 0$
Number of father-son pairs	39297	38068	1229	
Son education (years)	9.820	9.776	11.176	-3.192***
Son age (years)	29.558	29.635	27.189	-5.095***
Proportion of sons with:				
No schooling: 0 years	0.010	0.010	0.005	1.570
Up to primary level: 1-5 years	0.130	0.133	0.038	55.691***
Middle school completion: 6-8 years	0.214	0.218	0.082	76.136***
Up to Higher secondary: 9-12 years	0.449	0.449	0.476	2.095
Graduate and above: 13-16 years	0.197	0.190	0.399	195.6***
Father education (years)	5.791	5.786	5.953	-2.044**
Father age (years)	57.816	57.882	55.771	-4.948***
Household Assets	15.413	15.411	15.472	-0.944
Rural	0.614	0.614	0.610	1.013
Proportion of households belonging to different caste groups				
Scheduled Castes (SC)	0.213	0.213	0.218	0.082
Scheduled Tribes (ST)	0.086	0.086	0.087	0.065
Other backward castes (OBC)	0.404	0.405	0.356	7.086***
Forward caste	0.283	0.281	0.324	6.245**
Others	0.014	0.014	0.013	0.050
Proportion of households belonging to different religions				
Hindus	0.782	0.783	0.752	3.693*
Muslims	0.147	0.146	0.159	0.843
Other religion	0.072	0.071	0.089	3.057*

Migration Characteristics (Based on Migrant subsample)

	Age at Migration	Rural to Urban	Rural to Rural	Urban to Rural	Urban to Urban
Mean	12.61	0.34	0.13	0.07	0.45
Standard Deviation	3.58	0.47	0.35	0.25	0.49

Notes: We use chi-squared tests for comparing characteristics represented as proportions, and Welch's t-test for comparing characteristics represented by mean values. The last column "Tests of Equality" reports the test statistic. *,**, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively. All statistics are calculated using IHDS-2 sampling weights.

Table 5.3: OLS estimates of IRC and ICC for Non-Migrants and Migrants

	IRC			*ICC*		
	Non-Migrants $\hat{\theta}_0$	Migrants $\hat{\theta}_1$	Test of Equality $\hat{\theta}_0 - \hat{\theta}_1 = 0$	Non-Migrants $\hat{\theta}_2$	Migrants $\hat{\theta}_3$	Test of Equality $\hat{\theta}_2 - \hat{\theta}_3 = 0$
Father education	0.489***	0.386***	14.534***	0.405***	0.326***	9.011***
	(0.005)	(0.027)		(0.007)	(0.042)	
Child age	0.068***	0.143***	-11.735***	0.019***	0.138***	-5.550***
	(0.004)	(0.037)		(0.001)	(0.012)	
Father Age	-0.001	-0.002	0.205	-0.007	-0.002	1.569
	(0.011)	(0.017)		(0.019)	(0.006)	
Rural	-0.474***	-0.273***	-16.725***	-0.134***	-0.155***	4.831***
	(0.054)	(0.041)		(0.015)	(0.015)	
OBC	-0.588***	-0.051	-30.781***	-0.166***	-0.221**	1.367
	(0.059)	(0.072)		(0.017)	(0.104)	
SC	-0.887***	-0.225**	-7.049***	-0.249***	-0.077*	-5.381***
	(0.071)	(0.129)		(0.020)	(0.042)	
ST	-0.653***	-0.655*	0.010	-0.184***	-0.223*	0.592
	(0.102)	(0.372)		(0.029)	(0.131)	
Muslim	-1.112***	-0.807*	-2.213**	-0.313***	-0.276*	-0.786
	(0.076)	(0.483)		(0.022)	(0.165)	
Other Religion	-0.162	0.391	-3.683***	-0.046	0.134	-3.485***
	(0.109)	(0.526)		(0.031)	(0.181)	
Household assets	0.375***	0.433***	-1.400	0.123***	0.289**	-11.744***
	(0.106)	(0.144)		(0.029)	(0.125)	

Notes: PSU-clustered standard errors reported in parentheses. *,**, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively. The regression also includes child's age square and state dummies as additional regressors.

Table 5.4: Lewbel IV estimates of IRC and ICC for Non-Migrant and Migrant sample

	IRC			ICC		
	Non Migrants $\hat{\theta}_0$	Migrants $\hat{\theta}_1$	Test of Equality $\hat{\theta}_0 - \hat{\theta}_1 = 0$	Non Migrants $\hat{\theta}_2$	Migrants $\hat{\theta}_3$	Test of Equality $\hat{\theta}_2 - \hat{\theta}_3 = 0$
Father education	0.402***	0.343***	18.393***	0.327***	0.247***	7.299***
	(0.005)	(0.030)		(0.006)	(0.038)	
Child age	0.066***	0.323***	-24.300***	0.022***	0.158***	-37.053***
	(0.004)	(0.037)		(0.001)	(0.013)	
Father Age	-0.001	-0.002	0.159	-0.008	-0.002	-3.195***
	(0.011)	(0.020)		(0.022)	(0.005)	
Rural	-0.457***	-0.289***	-12.899***	-0.151***	-0.143***	-1.808*
	(0.063)	(0.044)		(0.015)	(0.016)	
OBC	-0.577***	-0.041	-22.254***	-0.159***	-0.216**	1.930*
	(0.056)	(0.084)		(0.016)	(0.104)	
SC	-1.053***	-0.237*	-20.085***	-0.234***	-0.089*	-13.453***
	(0.072)	(0.142)		(0.021)	(0.038)	
ST	-0.624***	-0.665*	0.407	-0.208***	-0.218*	0.269
	(0.118)	(0.358)		(0.029)	(0.125)	
Muslim	-0.927***	-0.696*	-1.502	-0.337***	-0.266*	-1.862*
	(0.067)	(0.439)		(0.019)	(0.134)	
Other Religion	-0.131	0.448	-3.355***	-0.041	0.140	-3.702***
	(0.094)	(0.606)		(0.028)	(0.171)	
Household assets	0.446***	0.474***	-0.820	0.135***	0.307**	-4.515***
	(0.100)	(0.119)		(0.029)	(0.133)	

Diagnostic statistics for Lewbel IV regressions

Sample	Breusch-Pagan Cook-Weisberg test statistic	Kleibergen-Paap test statistic	Hansen J test (*p*-value)	Durbin-Wu-Hausman test statistic
Non Migrants	27.57***	10.804***	0.612	7.742***
Migrants	77.14***	24.709***	0.373	10.301***

Notes: PSU-clustered standard errors reported in parentheses. *,**, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively

Table 5.5: High parental education's effect on child educational attainment (Matching-based estimates)

	Non Migrants			Migrants		
	Threshold for high father's education			Threshold for high father's education		
Matching Estimator	Primary & above	Middle & above	Secondary & above	Primary & above	Middle & above	Secondary & above
Nearest Neighbour (NN)	2.156***	2.510***	2.900***	2.061***	2.418***	2.767***
	(0.124)	(0.123)	(0.185)	(0.570)	(0.500)	(0.805)
Mahalanobis Matching (MM)	2.104***	2.542***	2.689***	2.077***	2.478***	2.574***
	(0.123)	(0.117)	(0.176)	(0.610)	(0.579)	(0.738)
Kernel-based matching estimators						
Epanechnikov kernel	2.271***	2.607***	2.871***	2.130***	2.497***	2.787***
	(0.096)	(0.091)	(0.111)	(0.478)	(0.445)	(0.507)
Biweight kernel	2.262***	2.601***	2.864***	2.402***	2.484***	2.707***
	(0.096)	(0.091)	(0.112)	(0.435)	(0.405)	(0.532)
Normal kernel	2.664***	2.841***	2.995***	2.583***	2.640***	2.852***
	(0.112)	(0.088)	(0.108)	(0.504)	(0.412)	(0.446)
Uniform kernel	2.289***	2.617***	2.881***	2.089***	2.494***	2.677***
	(0.095)	(0.090)	(0.111)	(0.423)	(0.442)	(0.483)

Notes: This table reports the average treatment effect on treated (ATT) estimates representing the effect of high father's education on son's educational attainment. All matching models were estimated with a covariate set comprising the full set of control variables with the full sample of father-son pairs. PSU-clustered standard errors reported in parentheses. *, **, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively. Primary & above refers to at least 1 year of schooling, Middle & above refers to at least 6 years of schooling, and Secondary & above refers to at least 9 years of schooling.

Table 5.6: Rosenbaum bounds sensitivity analysis

	Non-Migrant Sample						Migrant Sample					
	NN Estimator	MM Estimator	Kernel Estimators				NN Estimator	MM Estimator	Kernel Estimators			
			Epanechnikov	Biweight	Normal	Uniform			Epanechnikov	Biweight	Normal	Uniform
Γ	**Sig+**	**Sig+**	**Sig+**	**Sig+**	**Sig+**	**Sig+**	**Sig+**	**Sig+**	**Sig+**	**Sig+**	**Sig+**	**Sig+**
Panel A: Primary schooling as the threshold for high father education												
1.0	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.001	0.001	0.000	0.002	0.002
1.2	0.001	0.001	0.002	0.002	0.002	0.001	0.002	0.004	0.008	0.018	0.001	0.015
1.4	0.002	0.003	0.004	0.002	0.003	0.002	0.017	0.008	0.024	0.019	0.016	0.018
1.6	0.005	0.004	0.008	0.004	0.006	0.004	0.011	0.017	0.033	0.024	0.029	0.012
1.8	0.009	0.008	0.016	0.009	0.012	0.009	0.040	0.026	0.040	0.039	0.030	0.036
2.0	0.019	0.017	0.033	0.020	0.026	0.018	0.042	0.030	0.052	0.047	0.034	0.042
Panel B: Middle school as the threshold for high father education												
1.0	0.001	0.001	0.001	0.000	0.000	0.001	0.001	0.000	0.002	0.001	0.000	0.000
1.2	0.002	0.001	0.002	0.000	0.001	0.001	0.007	0.003	0.022	0.014	0.023	0.004
1.4	0.003	0.003	0.004	0.002	0.002	0.003	0.026	0.016	0.007	0.023	0.004	0.009
1.6	0.006	0.005	0.006	0.004	0.004	0.006	0.022	0.028	0.016	0.008	0.032	0.023
1.8	0.011	0.010	0.013	0.007	0.009	0.012	0.035	0.033	0.041	0.032	0.038	0.041
2.0	0.025	0.020	0.027	0.017	0.018	0.024	0.038	0.044	0.058	0.034	0.045	0.050
Panel C: Secondary schooling as the threshold for high father education												
1.0	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.001	0.002	0.002	0.002
1.2	0.001	0.002	0.001	0.002	0.002	0.001	0.014	0.006	0.014	0.007	0.002	0.017
1.4	0.002	0.003	0.002	0.003	0.004	0.002	0.014	0.027	0.013	0.008	0.004	0.014
1.6	0.004	0.006	0.004	0.005	0.007	0.005	0.029	0.017	0.008	0.017	0.013	0.027
1.8	0.007	0.011	0.009	0.010	0.013	0.011	0.037	0.040	0.028	0.034	0.030	0.037
2.0	0.014	0.023	0.020	0.021	0.028	0.020	0.044	0.045	0.043	0.039	0.044	0.047

Notes: Sig+ is the p-value representing the significance of upper bound of ATT.

Table 5.7: Effects of migration characteristics on child's education attainment

	IRC		ICC	
	OLS	Lewbel IV	OLS	Lewbel IV
Father education	0.418***	0.397***	0.372***	0.388**
	(0.129)	(0.138)	(0.142)	(0.219)
Child age	0.397***	0.364***	0.116***	0.109***
	(0.037)	(0.045)	(0.011)	(0.014)
Age at migration	0.233***	0.270***	0.179***	0.193***
	(0.078)	(0.089)	(0.013)	(0.031)
Age migration^2	-0.012***	-0.012***	-0.004***	-0.004
	(0.002)	(0.002)	(0.000)	(0.000)
Muslim	-0.785*	-0.698*	-0.245	-0.288
	(0.441)	(0.402)	(0.166)	(0.208)
Other religion	0.323	0.297	0.180	0.161
	(0.382)	(0.365)	(0.176)	(0.220)
SC	-0.225*	0.170	-0.082	-0.085
	(0.129)	(0.119)	(0.102)	(0.118)
ST	-0.655	-0.108	-0.256	-0.255
	(0.472)	(0.081)	(0.218)	(0.220)
OBC	-0.251**	-0.134*	-0.244*	-0.288*
	(0.132)	(0.073)	(0.140)	(0.169)
Household assets	0.325***	0.411***	0.244**	0.251**
	(0.189)	(0.156)	(0.117)	(0.122)
Rural to Urban	0.857**	0.796**	0.293**	0.233**
	(0.423)	(0.403)	(0.145)	(0.117)
Urban to Urban	0.861**	0.808**	0.295*	0.237**
	(0.435)	(0.396)	(0.157)	(0.120)
Urban to Rural	0.137	0.224	0.047	0.076
	(0.597)	(0.638)	(0.204)	(0.218)

Notes: PSU-clustered standard errors reported in parentheses.. *, **, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively.

Table 5.8: Migration background, gender and IEM

	Father-Son		Father-Daughter		Mother-Son		Mother-Daughter	
	ICC	IRC	ICC	IRC	ICC	IRC	ICC	IRC
Non-migrants	0.452^{***}	0.489^{***}	0.336^{***}	0.451^{***}	0.345^{***}	0.374^{***}	0.399^{***}	0.411^{***}
	(0.005)	(0.007)	(0.011)	(0.012)	(0.006)	(0.007)	(0.012)	(0.012)
Observations	9,194	9,194	2,288	2,288	7,225	7,225	2,155	2,155
Migrants	0.354^{***}	0.365^{***}	0.133^{***}	0.151^{***}	0.201^{***}	0.296^{***}	0.253^{***}	0.336^{***}
	(0.025)	(0.038)	(0.085)	(0.056)	(0.044)	(0.065)	(0.084)	(0.112)
Observations	302	302	86	86	168	168	82	82

Notes: All coefficients are based on the OLS regression estimated with the full set of control variables. To facilitate comparison between migrants and non-migrants, this analysis has been restricted to parent-child pairs who co-reside in the same household. PSU-clustered standard errors reported in parentheses. *, **, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively.

Table 5.9: ICCs for father-son pairs

	Migrants	*Non-migrants*
Correlation Coefficient (ICC)	0.372***	0.427***
	(0.042)	(0.007)
No. of observations (N)	1,229	38,068

Notes: *** p<0.01, ** p<0.05, * p<0.1; PSU-clustered standard errors in parentheses.

Table 5.10: Decomposition of persistence measured by correlation coefficient for the migrant sample (father-son pairs)

Child's education/birth cohort	1947-1956	1957-1966	1967-1976	1977-1986	1987-96	Overall
Father: No Education						
C:No education	-	-	0.000	0.000	0.000	0.000
C: Primary	-	-	0.000	0.000	0.080	0.056
C:Middle	-	-	0.000	0.074	0.079	0.074
C:Secondary	-	-	0.221	0.000	0.033	0.054
C:College	-	-	-0.049	-0.058	-0.018	-0.047
Total contribution	-	-	0.172	0.016	0.174	0.137
Father: Primary						
C:No education	-	-	0.000	0.000	0.000	0.000
C: Primary	-	-	0.000	0.000	0.011	0.008
C:Middle	-	-	0.070	0.000	0.023	0.024
C:Secondary	-	-	0.089	0.121	0.039	0.070
C:College	-	-	-0.039	-0.050	-0.008	-0.035
Total contribution	-	-	0.120	0.071	0.065	0.068
Father: Middle						
C:No education	-	-	0.000	0.000	0.000	0.000
C: Primary	-	-	0.000	0.000	0.004	0.005
C:Middle	-	-	0.000	0.000	0.000	0.000
C:Secondary	-	-	0.000	0.067	0.005	0.014
C:College	-	-	-0.004	-0.011	-0.009	-0.011
Table 5.10 (continued).						
Total contribution	-	-	-0.004	0.056	0.000	0.008
Father: Secondary						
C:No education	-	-	0.000	0.000	0.000	0.000
C: Primary	-	-	0.000	0.000	-0.004	-0.002
C:Middle	-	-	0.000	0.000	-0.008	-0.004

C:Secondary	-	-	-0.021	-0.001	-0.026	-0.026
C:College	-	-	0.047	0.001	0.064	0.045
Total contribution	-	-	0.026	0.000	0.027	0.012
Father: College						
C:No education	-	-	0.000	0.000	0.000	0.000
C: Primary	-	-	0.000	0.000	0.000	0.000
C:Middle	-	-	0.000	0.000	-0.019	-0.014
C:Secondary	-	-	0.000	-0.025	-0.045	-0.056
C:College	-	-	0.193	0.181	0.172	0.216
Total contribution	-	-	0.193	0.157	0.108	0.146
Correlation Coefficient	-	-	0.507	0.304	0.373	0.372

Table 5.11: Decomposition of ICC for father-son pairs,

Panel A: Decomposition of ICC for migrant father-son pairs

Son's Education	Father's Education					
	No Education	Primary Education	Middle Education	Secondary Education	College Education	Total contribution
No education	0.000	0.000	0.000	0.000	0.000	0.000
Primary	0.056	0.008	0.005	-0.002	0.000	0.067
Middle	0.074	0.024	0.000	-0.004	-0.014	0.080
Secondary	0.054	0.070	0.014	-0.026	-0.056	0.056
College	-0.047	-0.035	-0.011	0.045	0.216	0.168
Total contribution	0.137	0.068	0.008	0.013	0.146	0.372

Panel B: Decomposition of ICC for non-migrant father-son pairs

Son's Education	Father's Education					
	No Education	Primary Education	Middle Education	Secondary Education	College Education	Total contribution
No education	0.009	0.000	-0.001	-0.001	0.000	0.007
Primary	0.148	0.009	-0.013	-0.012	-0.001	0.131
Middle	0.068	0.005	-0.011	-0.013	-0.002	0.047
Secondary	-0.018	-0.002	0.006	0.015	0.005	0.006
College	-0.036	-0.006	0.022	0.121	0.134	0.235
Total contribution	0.171	0.007	0.003	0.111	0.135	0.427

Chapter 6

Schooling Progression in India: Does Return Migration Affect Mobility?

6.1 Introduction

In the wake of low enrolment rate and high drop-out rates, Government of India had initiated Sarva Shiksha Abhiyan (SSA) in 2001. Although the primary enrolment rate has drastically been improved in the last few years, but high level of drop-out rates in higher education still poses a serious challenge (Sikdar & Mukherjee, 2012). In India, the Gross Enrolment Ratio (GER) at secondary level hovers around 40% , which is quite low in comparison to other countries in East Asia and Latin America whose average gross enrolment ratio ranges between 70 to 80% (World Bank, 2009). The difference in gross enrolment ratio at primary and secondary level indicates that child's transition between these two successive educational levels may be affected by the same factors, but differently. There have been prior empirical studies that have explored the determinants of schooling progression in developing countries(Hanushek et al., 2008; Levy, 1971; Ray, 2002). Research in the field of educational stratification has emphasized on the importance of family structure in shaping the outcome of future generations(Duncan, 1967; Hauser & Featherman, 1976; Mare, 1980). Their findings highlight the importance of household characteristics as well as socio-economic status of the family. Being one of the important household characteristics, migration history of households also plays a significant part in shaping the future of upcoming generations.

Increasing inequality in the recent decades has substantially increased the short-term migration in India. There is substantial empirical research focusing on temporary migration. Most of these studies have relied on data provided by the NSS. The NSS defines short-term migrants as individuals who were living away from their "usual place of residence[6]" for a time period ranging anywhere between one to six months during one year prior to survey, for employment or in search of employment. However, this definition precludes accurate estimation of short-term migrants. Therefore, the estimates of short-term migrants tend to be contentious due to reliance on six-month criteria to define usual place of residence (Deshingkar, 2015). Lack of adequate secondary datasets which can provide comprehensive information on migration history of individuals has made it difficult to analyze various dimensions of migration. Some recent studies have used primary surveys to assess the impact of short-term migration(Coffey, 2013; Coffey et al., 2015; Dodd et al., 2016). Yet, this topic holds many open questions which may constitute new research topics. One such question may be the association between short-term migrants and IEM.

The IHDS-II (2011-12) questionnaire offers some crucial information regarding the migration history of the individuals. The IHDS-II questionnaire asks respondents: " Have you or any member of your household left to find seasonal/short-term work during the last 5 years and returned to live here?" This study focuses on individuals who migrated for a while but returned back home to live but their stay away from home before returning back could range anywhere between one month and 5 years, therefore, for the rest of the paper we will refer these short-term migrants as 'return migrants'[7].

There appears to be no previous cited empirical research on the impact of return migration on IEM in India. We fill this void in literature by investigating the determinants and consequences of return migration in India. Using sequential logistic regressions, we examine the impact of return

---

[6] Usual place of residence refers to the place where an individual has spent six months or more with the exception of newborn infants

[7] First,the IHDS-II question regarding the migration history of individuals who spent some time away from home but returned back home to live in last 5 years doesn't fit into the existing criteria used by NSSO to define short-term migrants. To avoid ambiguity, we, therefore, use the term 'return migrants'. Second, the question in itself filters out individuals who have any intention to migrate in immediate future.

migration on schooling progression of the child. This study is novel in its approach of investigating the determinants of different educational transitions in India, while offering robust estimates by conducting sensitive analysis under different scenarios of unobserved heterogeneity.

The rest of the chapter is framed as follows. Section 6.2 presents past literature related to return migration. Section 6.3 describes the data source and methodology. Section 6.4 presents the results. Section 6.5 discuss the results from sensitivity analysis. Last section summarizes the findings.

6.2 Return Migration: Determinants and Consequences

Many people think that migration is a one-way phenomenon. However, many migrants move to other places for a short span of time with an intent to get back to their place of origin. Depending on the place of origin and their place of destination, return migrants can be roughly classified into two categories: international return migrants and internal return migrants. International return migrants are individuals returning to their home country after staying away in the foreign land for some time, while internal return migrants represents individuals who return to their home after staying away from home for a brief period, within the same geopolitical entity.

International return migrants, also known as Immigrants, contribute significantly to the economic development of their host economies. When migrants return to their home country, the accumulated savings and acquired skills and knowledge play as a potential driver of economic development in the source country. Theoretically, under a general equilibrium framework, (Djajić, 2014) suggests that welfare of non-migrant individuals in the origin country of the migrants is maximized if the migrants are employed in other countries for a duration of between 8-12 years.

Many previous empirical studies have explored the determinants of return migration and its impact on the home economy(Bijwaard et al., 2014; Borjas, 1985; Gibson et al., 2010; Ilahi, 1999; Kırdar, 2009; Yang, 2006).

All the above-mentioned research studies have focused on international return migrants. Being one of the key fields of migration, internal return migration remains a neglected topic. Internal return migration is still viewed as a no profit phenomenon for the local regional community. Some of the recent research in the field of internal return migrants make an attempt to show that even internal return migrants can positively affect the economic development in their regional community(Chen & Wang, 2019; Démurger & Xu, 2011; Wang & Fan, 2006). Recent literature in the field of internal migration try to plug the void in literature by scrutinizing the impact of local and regional return migrants. Newbold's (2001) research on return migrants is one of the pioneer studies that tries to categorize return migrants in Canada into migrants who return to the same residence and others who return to the same region. Results show that return migrants who return to the same residence tend to be younger and more educated compared to their counterparts who return to the same region but to a different residence. Another interesting finding is that employed people are less prone to coming back to their usual place of residence and more likely to relocate to a new residence but within the same province. In the last few years, some more research studies have analyzed the determinants and impact of internal return migrants (Démurger & Xu, 2011; Junge et al., 2015; Piotrowski & Tong, 2010; Thanh et al., 2019; Wang & Fan, 2006; Zhao, 2002).

In developing countries like India internal return migration takes the form of seasonal or circular movement between rural and urban areas. Internal migration has seen an upward trend in the last two decades (Keshri & Bhagat, 2013). Overtime, an extensive literature has evolved on the determinants of internal migration in India. Haberfeld et al., 1999) analyzes the attributes of rural household who have at least one seasonal migrant in their family, and find that probability of the household having a seasonal migrant is negatively associated with the number of educated members, number of workers and the total income earned from primary activities. (Parida & Madheswaran, 2011) find that age is positively related to the decision to migrate. Age is being used as substitute for work experience suggesting the earning potential of the individual, therefore positive correlation shows that older individuals tend to migrate in search of better income prospects. While, age-squared comes out to be negatively related with the decision to migrate showing that even though the earning potential of an individual increases with age, it is subject to diminishing returns. They also show marital status as a determining factor of migration where married individuals are more likely to migrate than their unmarried counterparts. Bhattacharya (1983) finds that household characteristics are more important than individual characteristics in determining the flow of migration.(Nayyar & Kim, 2018) use both rounds of IHDS (IHDS-I and IHDS-II) to identify migrants. They use information from both round of IHDS: focusing on households with no non-resident members in the family in 2004-05 survey and studying the same households for presence of any non-resident members in 2011-12. The panel regression findings show that total income of the household has a very strong correlation with the likelihood of having a migrant family member. In addition, results also suggest that household size and decision to migrate are non-linearly related to each other. Precisely speaking, smaller households have higher chance of having a migrant member, but after a certain threshold as the household size increases the chance of the household having a migrant family member increase. The latter reason can be supported by the fact that after a certain household size, the decrease in the per capita land holdings motivate the family members to search for additional income sources outside their home region

(Hay, 1980; Nabi, 1984; Singh 1988). Regarding social groups, their findings suggest that the household association with a particular social group does not have any statistically significant influence one's decision regarding migration. The results are in conflict with previous empirical studies which argues that individuals belonging to lower caste households are less likely to migrate as compared to individuals belonging to the upper caste households. (Dubey et al., 2004) categorize internal migrants into two categories: members belonging to upper caste households who migrate to enhance their standard of living and those who belong to lower caste households who migrate for the sake of survival. Other studies find that distance to destination area is negatively correlated to internal migration (Kone et al., 2018; Parida & Madheswaran, 2011).However, there is no study which has empirically analysed internal return migration, in an Indian context.

Return migrants do not solely have an immediate effect on earnings or wealth accumulation. They tend to pass on their accumulated wealth, skills and other socio-economic characteristics to the next generations. Therefore, it becomes important to study the link between return migration and IEM. The term 'IEM' refers to the transmission of socioeconomic characteristics across generations within the same family. The topic of IEM has been extensively researched (Björklund et al., 2006; Black & Devereux, 2010; Chetty et al., 2014; Holmlund et al., 2011a; Lindahl et al., 2012; Solon, 1992b, 2002), but less is known about the association between the family's migration history and its impact on IEM.

The topic of IEM has also been well researched in India(Azam & Bhatt, 2015; Emran & Shilpi, 2015; Hnatkovska et al., 2013; Motiram & Singh, 2012; Reddy, 2015). Most of the Indian studies have focused on IEM due to access to reliable information on individual's completed years of education. All studies find that parental education level significantly affects the education level of their child. Nearly all studies report that socioeconomic status of the family, like caste, religion, and household income acts as an important determent of child's educational level. Some recent studies have analyzed the impact of migration on child's educational level (Battistella & Conaco, 1998; Coffey, 2013). Still, there barely exist any study which analyzes the association between family's migration status and IEM in India. This paper fills a void in IEM literature by addressing three research questions – (i) What are the factors which affect return migration in India; (ii) Does parental education affect the schooling progression of their children; and (iii) How does the migration status (whether return migrant or not) of the parents interact with their own education level in affecting the schooling progression of their child.

The current study is contributing to the existing literature on IEM in number of ways. First, existing studies on IEM focus on years of education completed by the child making it hard to capture the impact of specific household and schooling characteristics. Therefore, we focus on school progression as an indicator of child's schooling. Using sequential logit model, we explore the determinants of child schooling in India. Second, we utilize the information available on individuals who had migrated in the last five years at the time of survey but returned to their home to live to identify return migrants. This gives us an opportunity to examine the different factors which affect return migration and to assess its impact on schooling progression of the child. Third, we also explore the effect of return migration on IEM by including an interaction effect between parental education and return migration. Last, considering the fact that our estimates may be biased due to endogenous nature of our main variables of interest (parental education and return migration), we run sensitivity analyses to find out the sensitivity of estimates of our main variables of interest under different scenarios of unobserved heterogeneity.

6.3 Data and Methodology

6.3.1 Data Source

This analysis in this chapter relies on data collected by IHDS-II . The IHDS offer nationally representative data that covers multiple topics related to human development. Unlike single topic surveys, this data collects information on multiple topics covering different dimensions of human development in India. IHDS-I was conducted in 2011-12 where around 42,000 households were surveyed. IHDS-II was conducted in 2011-12, where around 83% of the households were surveyed which were also the part of IHDS-I survey covering around 41,554 urban and rural households in all states surveyed.

In the household interview, information is available at different levels, such as, household and individual level. Individual information offers data on personal characteristics of the individuals, for example, gender, age, education, caste, religion, marital status, occupation, migration status, etc. On the other hand, information on household includes place of residence, social status, total household income, number of family members, activity status and other demographic characteristics of the households. Additionally, a part of questionnaire focuses on the migration history of the household. It offers information regarding the household members who left the household to find short-term work in the last five years but returned to stay at their usual place of residence. Table 6.1 defines all the variables used in the analysis.

6.3.2 Defining Return Migrants

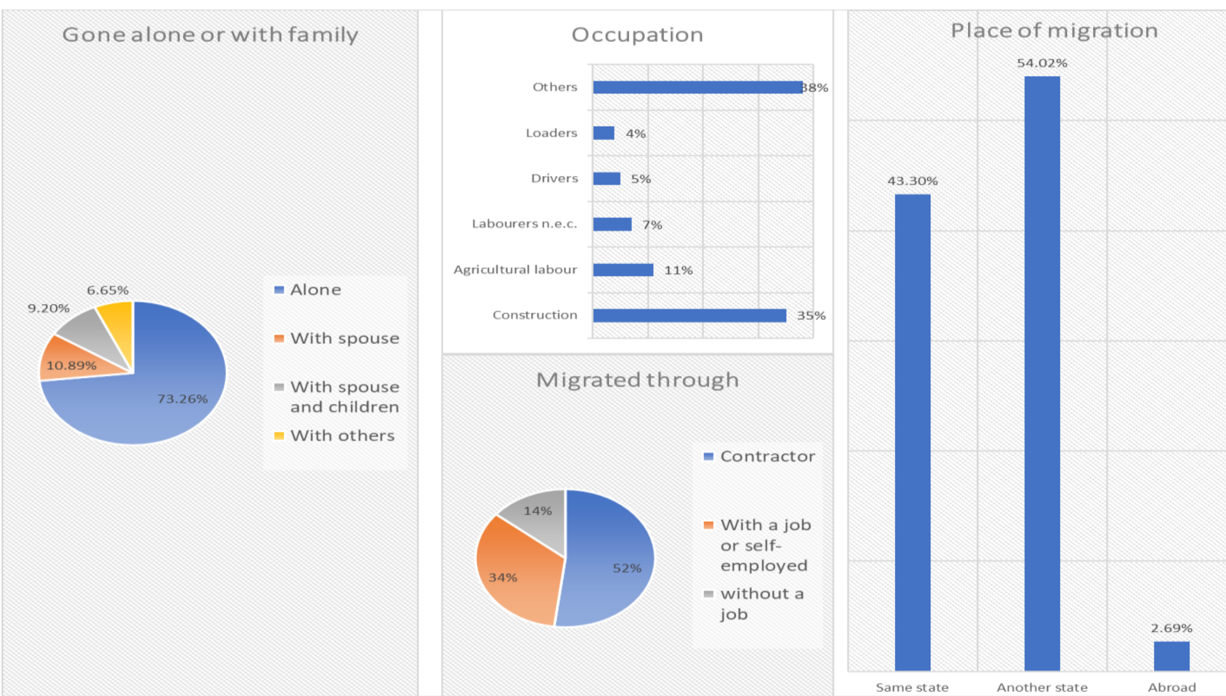
We use data provided by IHDS-II (2011-12) to identify return migrants. The IHDS-II questionnaire asks the respondents: "Have you or any member of your household left to find seasonal/short-term work during the last 5 years and returned to live here?" We use the term 'return migrants' to refer all the people who left the household in search of short-term work but returned to their place in the past five years prior to IHDS-II survey. Out of the total households surveyed in 2011-12, approximately 9% of them reported of having at least one return migrant. However, the share of total return migrants account is less than 2 per cent of the total population.

6.3.3 Characteristics of return migrants

The IHDS data on short-term migrants show that around two-thirds of the short-term migrants moved to urban region. Some of the characteristics of return migrants are shown in Figure 6.1. Data on return migrants reveals that almost two-thirds of total return migrants moved from rural sector to urban sector. Most of these workers travelled alone (73%) to destination area with majority of them migrating between the states and, on an average, stayed away from their home for a duration of 7.5 months. Only 9% of the workers migrated with their spouses and children. About one-third of the total short-term migrants (35%) migrated to work in the construction sector. Of the total households that reported at least one member as a return migrant, 85% belongs to Scheduled category (SC/STs). Regarding the education level, around 85% are illiterate or have completed primary education. Relevant data on return migrants shows that contractors or middlemen play a crucial role in the movement of migrants, where around fifty-two% of the return migrants had migrated with their help. Contractors or middlemen help in sourcing and recruiting the workers for internal migration to destination areas.

Descriptive statistics show that only 5% of the households have a father-son pair where father has been identified as a short-term migrant. On an average, these fathers were away from their home for 51 months in the last 5 years preceding the IHDS-II survey. Whereas, the proportion of households where mother has been identified as a short-term migrant is less than 1% (refer Table 6.4).

Figure 6.1: Characteristics of Short-term migrants



Source: Indian Human Development Survey, 2011-12 (IHDS-II).

6.3.4 Determinants of Return Migrants

First, we define return migrants as individuals who have migrated in the last five years preceding the 2011-12 survey but returned to their home before the survey. The IHDS questionnaire also provides information regarding household members who are non-resident. This provide information on migration stock (total number of migrants at a particular point of time), but it has its own shortcomings. First, the data doesn't provide information on the destination of migrants (whether they moved within the Indian territory or moved abroad). Second, it does not offer information regarding the year in which the household members migrated. Third, the non-resident questionnaire does not account for the type of migration (whether migrated alone or with some other household members). Therefore, comparing return migrants with migrants may lead to biased results.

To facilitate the analysis, we categorize return migrants into two categories: relatively permanent return migrants and temporary return migrants. We distinguish between relatively permanent and temporary return migrants. We define permanent return migrants as individuals who had migrated for any number of months but came back home one year prior to the 2011-12 survey. This definition allows information on returnees who have spent last one year at home and therefore, this increases their probability of participating in the local economic activities. Therefore, they can be seen as key actors of change in rural areas. Temporary return migrants represent individuals who had recent migration experience. In other words, workers who had migrated in the last one year but returned back home at the time of IHDS-II survey can be defined as temporary return migrants. In this paper, we refer temporary return migrants as continuing migrants as their frequent visits can be associated with probability of seeking migration work again. Non-migrant households are defined as those where no family member has any kind of migration experience.

We gauge the relation between child's schooling and parental socio-economic status. Therefore, we put the following restrictions on the observation to obtain a sub-sample that can address our research problem. First, we restrict our sample to children aged between 6-20 years and who were enrolled at some level of education and were not involved into any kind of labor market activities at the time of survey. Second, we consider only those return migrants who are male parent of the child. Therefore, non-migrants are defined as fathers who did not experience migration between the two survey rounds of IHDS (IHDS-I and IHDS-II).

Table 6.2 compares the individual and household level characteristics of return migrant with continuing migrants and non-migrants. The final sample comprises of 40,922 children aged between 6-20 years old. Out of 40,9222 children, 2709 children have fathers who are return migrants. Returnees are predominantly married. Regarding the educational achievement, continuing migrants are more competitive than permanent return migrants but less competitive than non-migrants. Results also show that the proportion of households belonging to minority social groups substantially vary between return migrants and non-migrants. Return migrants have higher probability of belonging to marginalized section of the society. The same difference exists between permanent return migrant and continuing migrant, where the proportion is high for the former group.

Regarding migration characteristics of different groups show that proportion of workers who moved with their family to destination area with a job is high among continuing migrants than permanent return migrants. Results also show that there are also differences in household characteristics. Composition of family also differs across different groups. On the other hand, the average number of elderly aging above 65 years is high in the non-migrant sample than the non-migrant counterparts. The average number of children is high among migrant household as compared to non-migrant counterparts. Similarly, the proportion of young children (Children under 6) is quite high among return migrant households. The average land holding in migrant households is lower than the non-migrant counterparts. The average number of adult laborers in the families of return migrants is low than in the families of non-migrants as well as continuing migrants.

Regarding the place of origin characteristics, the villages where permanent return migrants exist have higher proportion of non-farm labor than continuing migrants, but both have less proportion than villages which have no migrants. The same pattern is true for average per-capita income at the place of origin, however the difference between permanent return migrants and continuing migrants is not significant.

We use binomial multivariate logit regression to identify factors influencing the odds of being a return migrant. The dependent variable, return migrant, is a dichotomous variable. It takes the value 1 if an individual is a permanent return migrant and 0 if an individual is a continuing migrant. The model controls for individual and household level characteristics and state fixed effect. The inclusion of state fixed controls for all state-level unobservable factors.

Our final sample comprises of 2,709 observations. We estimate the following equation using binomial logit model:

$$P_i = \beta_1 + \beta_2\, Individual_i + \beta_3 Household_i + \beta_4 POr_i + \beta_5 State_i \qquad (6.1)$$

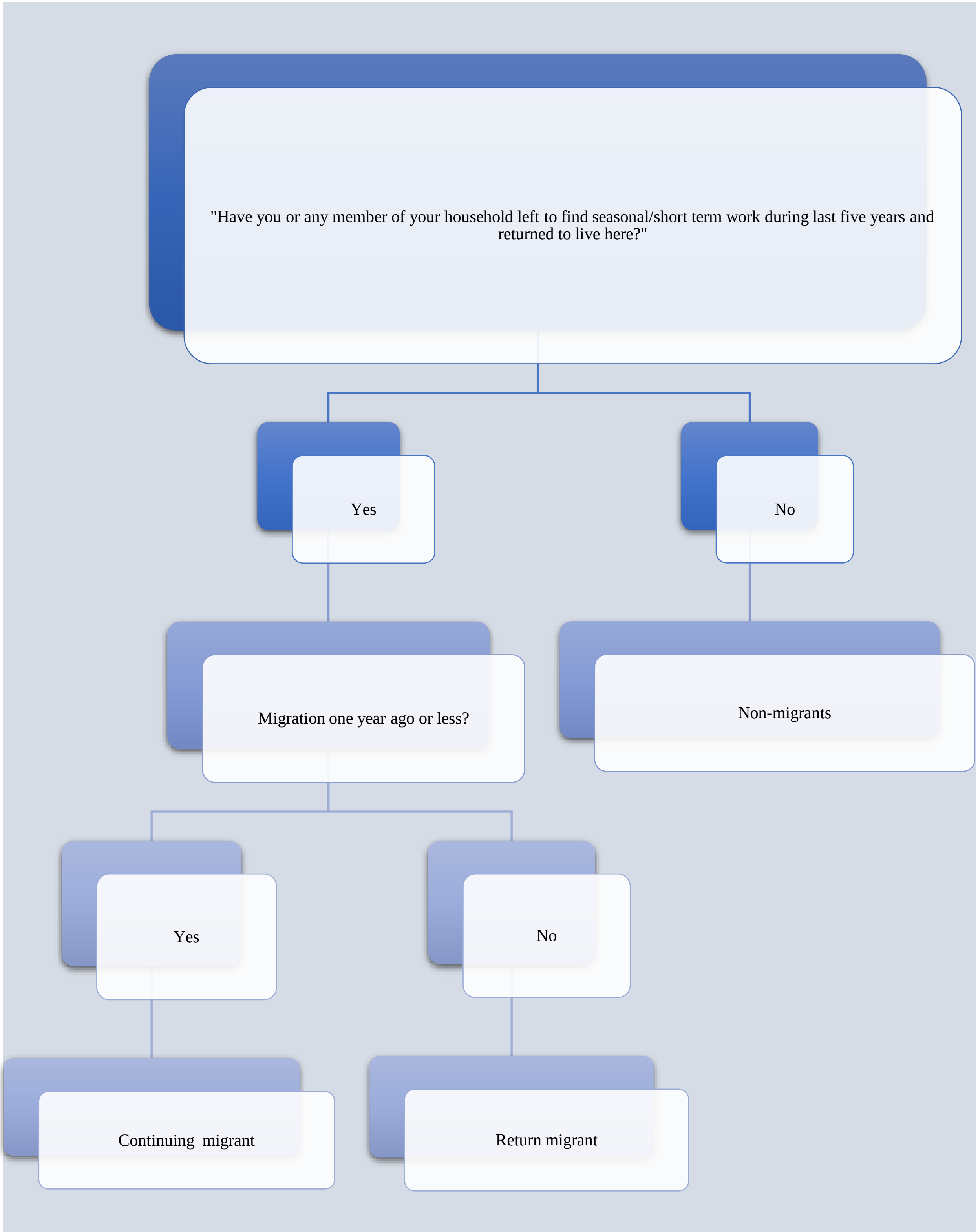


Figure 6.2: Identification of return migrants, continuing migrants, and non-migrants

Where P_i is the probability of individual being a permanent return migrant. Individual variable comprises of individual level characteristics like age, square of age, education level, and marital status. Household variable comprises of household level characteristics like composition of family, household assets (proxy for household welfare) and other socioeconomic status of the family. POr variable comprises of characteristics of migrant's place of origin like per capita income and proportion of non-farm laborers at the source area of the migrants. 'State' variable controls for the fixed effect of different states.

6.3.5 Empirical Framework: A Sequential Logit Model of School Progression

This chapter also explore the effect of return migration on transmission of education across generations. Following Waelbroeck (2003), we apply sequential logit modelling framework for our analysis. Many other studies have used this flexible methodology that does not restricts the distributional effect a-priori (Alpu and Fidan, 2004; Pal, 2004). This study considers three events, y_1 , y_2 y_3 (i.e., three schooling transitions) that occur sequentially. We categorize the educational attainment of children in four categories, namely, non-literate (children with zero years of schooling), primary education (children with 1 to 5 years of schooling), secondary education (children with 6 to 12 years of schooling) and post-secondary education (children with more than 12 years of schooling). Let y_{1i} =1 if the ith child has some level of primary education, 0 otherwise. Similarly, $y_{2i} = 1$ if the ith child has some level of secondary education, 0 otherwise. Finally, $y_{3i} = 1$ if the ith child has some level of post-secondary education. Then, the educational attainment is modeled through a sequence of three transitions: from non-literation to primary education, from primary education to secondary education, and from secondary education to post-secondary education. Our main variables of interest are parental (father or mother) education, return migrant (whether the male parent is a return migrant or not) and their interaction term.

This will help in controlling for any kind of bias which may arise due to inclusion of female return migrants. The adverse circumstances faced by females may force them to migrate temporarily in search of work. These unobserved factors may are likely to affect non-random selection of female migrant and child's schooling progression, simultaneously. Third, we include only those return migrants who are also the father of the child. This allows us to investigate the interplay between the father's education and migration status in affecting child's schooling progression.

Table 6.4 shows the summary statistics. There are around 40,922 male and female children aged between 6-20 years in our sample. These children constitute the potential population that should be enrolled in various level of education. We consider three sequentially related decision related to education transitions: non-literate to primary, primary to secondary and secondary to post-secondary. Summary statistics show that of the total sample of children in the age-group of 6-20 years, 9% have never joined school or have dropped out too early and only 13% of the potential population have some level of post-secondary education. Those who did not make to post-secondary school includes children who could not finish secondary education as well as those who completed secondary education but did not wish to take up post-secondary education. The sample shows that the average years of education attained by fathers and mothers is 5.8 years and 3.7 years, respectively. Summary statistics also show that only 5% of the total households have a parent-child pair where the male parent has been identified as a return migrant. On an average, these fathers were away from their home for 51 months in the last 5 years preceding the IHDS-II survey. Whereas the proportion of households where mother has been identified as a return migrant is less than 1%. Therefore, our final sample consists of 21,861 sons and 19,061 daughters who were enrolled at some level of education during the survey round and had father who is a return migrant.

We estimate sequential logit models to describe the relation between father's education and the likelihood that the child would transit from a lower education level to a higher education level. The sequential logit model is defined as follows

$$\Pr\left(pass_{1,i} = 1 \,\middle|\, X_i\right) = \hat{p}_{1i} = \Lambda(\hat{\alpha}_1 + \hat{\beta}_1 X_i) \tag{6.2}$$

$$\Pr\left(pass_{2,i} = 1 \,\middle|\, X_i\right) = \hat{p}_{2i} = \Lambda\left(\hat{\alpha}_2 + \hat{\beta}_2 X_i\right), if\ pass_{1,i} = 1 \tag{6.3}$$

$$\Pr\left(pass_{3,i} = 1 \,\middle|\, X_i\right) = \hat{p}_{3i} = \Lambda\left(\hat{\alpha}_3 + \hat{\beta}_3 X_i\right), if\ pass_{2,i} = 1 \tag{6.4}$$

Where $\Lambda(.) = \frac{\exp(.)}{1+\exp(.)}$ is the standard logistic function. $pass_{k,i}$ are indicator variables which indicate whether individual *i* passes the k^{th} transition (k=1,2,3), and $\hat{p}_{ki}$ is the conditional probability that individual *i* passes the k^{th} transition. X_i is the vector of explanatory variables, and $\hat{\beta}_k$ is a vector of estimated coefficients. This function makes sure that predicted probabilities ranges between 0 and 1. The coefficients of independent variables represent the log odds ratio, while the constant represents the baseline odds of successfully clearing the different schooling transitions.

6.3.6 Sample-Selectivity and Unobserved Heterogeneity: Challenges and Strategies

The sequential logit estimates could be biased due to unobserved heterogeneity which can be explained with two main points. First, the case of selective attrition where some children have higher probability to drop-out at successive educational transitions than others. This makes it more likely to introduce correlation between unobserved variables which affect the probability of the child making the first transition with observed and unobserved variables that affect the likelihood of the child making the higher educational transitions (Buis, 2011; Cameron & Heckman, 1998; Holm & Jaeger, 2011; Mare, 1980). If one could control for all the observed and unobserved variables, there will still be variation between children, which can be referred as 'idiosyncratic error'. This could produce biased estimates as the odds are likely to differ at different transitions ?due to variance in this idiosyncratic error. Therefore, Angrist & Pischke (2008) & Mood (2010) "suggest to resolve this issue by focusing more on the odds rather than the unobserved variables.

Second, when the model does not control for omitted variables then the estimated effect of the main variable of interest is averaged over the variables that have been omitted in the model rather than estimating its causal impact, resulting in biased coefficients. Several methodological strategies have been proposed to address the bias due to unobserved heterogeneity in educational transition models (see, for example, Buis, 2011; Holm & Jaeger, 2011; Lucas, 2001; Lucas, Fucella, & Berends, 2011; Mare, 1993). However, it is problematic to control for omitted variables.

The biggest challenge in analysing the impact of return migration is posed by selectivity issue. A common research strategy is to compare the outcome of households where at least one member is a return migrant to those where no one has migrated. Such comparisons are prone to serious methodological issues: first, households self-select into migration; second, some migrant household move with the family (which never gets recorded into the household survey dataset) while other migrant household leave some of the members behind; third, some migrants choose to return back home (some may be forced to return due to adverse conditions in the destination area, while others may self-select to return back home), so all the return migrant households may be considered equally affected by migration experience. In all the above discussed scenarios, there is presence of unobserved attributes, like personality type, ability, motivation and other unobserved institutional factors which are correlated with both the decision to migrate for a short period of time and return back home, and the educational transition of children of that households. Returnees could also have an indirect effect, spreading changing attitude towards education to the non-migrant households. If return migrants have a diffuse impact on the non-migrant households, even then the comparison of migrant households with non-migrant household in empirical analysis will provide us the lower bound of the direct effect of return migration. Parental education variable also poses the risk of endogeneity where unobserved factors like genetic ability are likely to be correlated to both parental education as well as the schooling progression of their children. In essence, our main variables of interest, parental education and return migration are potentially endogenous. However, it is problematic to control for all the variables. Some variables remain omitted from the analysis but might affect the dependant variable. This makes it difficult to estimate the causal impact of our main variable.

Therefore, instead of identifying the causal impact of our main variables of interest, we assess how different assumptions regarding the unobserved variable, u affects our conclusion. Using a novel approach proposed by Buis (2011), we perform the sensitivity analysis which examine the overall robustness of our estimates assuming various levels of unobserved heterogeneity. The sensitivity analysis comprises re-estimating the sequential logit model under a wide range of scenarios assuming different magnitudes for unobserved heterogeneity term, and by varying the correlation of the unobserved heterogeneity term with our main variables of interest.

One of the key issues while assuming a range of scenarios concerning unobserved heterogeneity is to create reasonable scenarios. Before creating different scenarios, we assume the unobserved heterogeneity variable as a standardized normal variable, u, that represents a weighted total of all the variables that remain unobserved, but we would like to include them in our sequential logit specification. This will facilitate in comparing the effect of u across different educational transitions with the effect of standardized observed variables. Then, β_u will be the estimated effect size (log odds ratio) for the unobserved heterogeneity term, u.

6.4 Results

6.4.1 Determinants of Return Migration

The findings are shown in Table 6.3. Age is found be positively associated with one's probability to return home after temporary migration. This shows that older individuals have higher chance to migrate on a short-term basis than their younger counterparts. The findings are consistent with many previous studies on return migration (Massey, 1987; Sharda, 1984). We run two different regressions to make our results comparable to other empirical studies. The results are shown under two different specifications: Model 1 includes the education variables as levels of schooling while in model 2, education has been included as a continuous variable. Results from model 1 show that high-educated individuals are more likely to return back home than less-educated individuals. Model 2 provides the same result for education variable. This could be attributed to educated individuals being more risk averse by preferring employment at their source area rather than migrating to a new place. Migrants in host areas are usually employed in unskilled jobs that does not reward education. Therefore, educated migrants are more likely to take non-farm employment at their place of origin rather than wasting time at host area in finding suitable jobs that fit their educational qualification (Zhao, 1999). Another interesting finding from the results show that individuals belonging to households who have only sons tend to return back after migrating for a short span of time.

Individuals who migrate along with their family are 34% less prone to coming back home permanently, while those who migrate to new location without a prior job are 32% more likely to return back home than those who migrate with a job. Results also show that individuals belonging to places with high average per-capita income and proportion of non-labour farm work have higher probability to permanently return back home.

6.4.2 Consequence of Return Migration on Child's Schooling Progression

Results from a sequential analysis is shown in Table 6.5. Tables 6.5 reports the exponentiated coefficients for the sequential logit models that represent the effect of parental education on child's schooling transitions. These exponentiated coefficients can be interpreted as odds ratio which are estimated using maximum likelihood approach. Whenever we study impact on educational transitions, it is better to present the effects as odds ratio rather than marginal effects. It helps to net out the effect of educational expansion from other changes. In estimation of sequential logit model, an odds ratio helps to control for educational expansion by controlling for changes in baseline odds for each educational transition. In other words, it measures the relative effect. The original education variable is a continuous variable ranging from 0 to 16 but is standardized to have 0 mean and a standard deviation of 1. Further, we use survey weights in estimating the model.

We find that irrespective of child's gender, an increase in the level of education attained by parents is associated with an increase in the odds of their child transitioning from a lower education level to a higher education level. For example, a one-year increase in father's education increases the odds of transitioning from primary to secondary education by 4.4% (5.3%) for sons (daughters). Similarly, a one-year increase in mother's education increases the odds of transitioning from primary to secondary education by 4.6% (5.2%) for sons (daughters). The odds of moving to a upper education level increase with an increase in household assets, for both sons and daughters. The presence of younger siblings lowers the odds of transition to higher levels of education for both sons and daughters. The children of urban households and Hindu households have higher odds for all three transitions relative to the children of rural households and non-Hindu households, respectively. There are very small differences across the various castes which are statistically insignificant with one exception—-the daughters from OBC households have 30.4% lower odds of the transitioning from non-literate to primary education as compared to those from forward caste households (reference category).

We also investigate the effect of father's migration experience on IEM of education and find mixed results. We find that father's migration experience increases the effect of father's education for the primary to secondary transition for both sons and daughters, and it reduces the effect of father's education for secondary to post-secondary transition of sons. However, among uneducated fathers (Father Education = 0 years), father's migration experience reduces the odds of primary to secondary transition for both sons and daughters. Nonetheless, one must be cautious in interpreting these effects as causal as both parental education attainment and migration decision are potentially endogenous and therefore these estimates could be biased due to unobserved heterogeneity. In the next section, we analyze the robustness of these sequential logit estimates to the presence of unobserved heterogeneity by implementing the sensitivity analysis of Buis (2011).

6.4.3 Additional Determinants of Child's Schooling Progression among Return Migrants (Gender Disaggregated Analysis)

Section 6.4.2 shows that the degree of impact of parental education on child's schooling transitions varies by the gender of the child. This makes it important to scrutinize the issue by analyzing how other factors of child's schooling progression are linked with the child's gender. Table 6.8 shows the results (marginal effects) for sequential model used to study child's educational transitions. The analysis is being done by taking the sub-sample of return migrants. Results show that parental education, child's age and its square, place of residence (urban/rural), and household assets alter the transition probabilities of sons and daughters in the same manner. The age of the child is significantly positive for all the three transitions, indicating that children at the upper end of the education ladder are more likely to pass the transitions than the younger ones. The negative sign for age-squared term suggest the non-linearity in the effect of age across the educational transition. Similarly, positive marginal effect for urban shows that urban children have higher probability of passing all the three transitions than their rural counterparts. However, being a daughter belonging to minority group (Non-Hindu,SC/STs,OBCs) significantly reduces the likelihood of passing all the three transitions. In short, girl child faces double disadvantage (once by gender, once by caste/ethnicity) as compared to a male child. In India, widespread caste and gender based inequality leads to higher discrimination faced by women from minority castes. Increasing crime and sexual violence against Dalit women coupled with absence of parent due to migration forces them to drop out of school. The findings also suggest that for sons, presence of young children (children under 6) in the household considerably lessens the likelihood of passing the first two transitions. This could be attributed to several reasons. First, in migrant families, parental absence may force the male child to take the responsibility of the household members to ease the financial burden on the household. Second, although remittance may improve the financial condition of the household and motivate the parents to invest in child's schooling but successful migration may also compel them (or their children) to use their financial resources to create an alternative to schooling. They could invest their earnings on non-farm activities which can be seen as better alternative which is less risky than educational investment (Curran et al. 2004; Kandel and Kao,2000). However, in case of daughters, it comes out to be significant only for the first transitions. In households where there are young children, older daughters are assigned the responsibility to take care of the younger siblings. Additionally, migrant families where the likelihood of both parents working is high, the responsibility of daughters to perform household duties increases manifold. Presence of school age (6-14 year olds) siblings significantly increases the likelihood of the first two transitions for sons. Number of siblings seems to be associated with lowering the likelihood of the first two transitions. Higher number of siblings will lead to competition for resources and dilution of resources will adversely impact the educational progress of the children. However, at the higher transition the number of siblings increases the likelihood of passing the third transition. These results may indicate that once a child proves his academic competence, his education at the higher level is seen as a better investment option yielding higher return in future. However, for daughters, the impact of number of siblings is insignificant for all the three transitions.

6.5 Sensitivity Analysis

Following Buis (2011), we assume that the unobserved heterogeneity and its effect size remains constant across the transitions. Some studies attempt to use past academic performance to represent unobserved heterogeneity such as genetic endowments and motivation (Erikson et al., 2005; Kloosterman et al., 2009). These studies observed that the effect size of standardized variable of past academic performance is approximately 2.5 in terms of log odds ratio. This is a sizeable effect which suggests that a one standard deviation increase in the unobserved heterogeneity increases the odds of moving from a lower education level to a higher education level by a factor of 12 ($e^{2.5}$). In the first set of scenarios for the sensitivity analysis (Appendix B, Tables B.1 to B.6), we re-estimate the sequential logit model assuming β_u varies from 0 to 5. $\beta_u = 0$ implies that unobserved heterogeneity does not change the odds of the transition ($e^{\beta_u} = 1.0$) , and $\beta_u = 5$ is a extreme assumption which implies that a one standard deviation increase in unobserved heterogeneity increases the odds of transition by a factor of 148 ($e^{\beta_u} = 148.4$). In the second set of scenarios for the sensitivity analysis (Appendix B, Tables B.7 to B.24), we re-estimate the sequential logit model assuming an effect size of $\beta_u = 0$, and by varying the correlation between the unobserved heterogeneity term and our main variables of interest—parental education and return migrantion—from -0.6 to 0.6. Overall, we re-estimate the sequential logit model under 192 unique scenarios by varying the effect size of unobserved heterogeneity term and its correlation with the different variables of interest. The sensitivity analysis provides a range of estimates depending on the degree of unobserved heterogeneity assumed.

Tables 6.6 and 6.7 summarize the result of the sensitivity analysis for sons and daughters, respectively. The Baseline 1 model reported in column 2 (ρ=0; βu=0) of Tables 6.6 and 6.7 is same as the initial sequential logit model estimates reported in Table 6.5. The Baseline 1 model introduces unobserved heterogeneity term that is uncorrelated with our variables of interest, and which has an effect size of $\beta u=2.5$. In both Baseline models, we find that irrespective of the gender of the child, an increase in parental education is correlated with an increase in the odds of their child transitioning from a lower education level to a higher education level. However, sensitivity analysis suggests that after accounting for potential bias due to unobserved heterogeneity the effects of parental education on child's educational transition are weaker. For instance, an increase in father's education is associated with an increase in odds of primary to secondary and secondary to post-secondary transitions for sons, but it is not associated with an increase in odds of non-literate to primary transition (refer last column of Table 6.6, maximum p-value). Similarly, an increase in mother's education is associated with better odds of non-literate to primary and primary to secondary transitions for sons, and not with the secondary to post-secondary transition (refer last column of Table 6.6, maximum p-value). We find that the relation between father's education and daughter's education is not robust to unobserved heterogeneity for all three educational transitions (refer last column of Table 6.7, maximum p-value). However, an increase in mother's education improves the odds of all three educational transitions for daughters, and these estimates are robust to unobserved heterogeneity (refer last column of Table 6.7, maximum p-value).

We find that the estimated effect for the Migrant dummy is highly sensitive to the unobserved heterogeneity when the unobserved heterogeneity term is correlated with the migration decision (refer Appendix B, Tables B.1 to B.24). Therefore, the baseline model results that among the uneducated fathers (Father Education = 0 years), father's migration experience reduces the odds of primary to secondary transition for both sons and daughters must be interpreted with caution. We find that the "Father Education × Migrant" interaction term is statistically significant and robust to unobserved heterogeneity for primary to secondary and secondary to post-secondary transitions of sons (refer Table 6.6). The minimum and maximum effect size (odds ratio) for the interaction term are 1.057 and 1.177 for the primary to secondary transition, and they are 0.800 and 0.903 for the secondary to post-secondary transition. This implies that father's migration experience reduces the IEM (increases the effect of father's education) for primary to secondary transition of sons, and it increases IEM (reduces the effect of father's education) for the secondary to post-secondary transition. We find that the interaction between father's education and migration status is insignificant for daughters for all three transitions at a 5% significance level (refer last column of Table 6.7, maximum p-value).

6.5 Conclusion

In this paper, we investigate the determinants of child progression in schooling using sequential logit modelling technique. We also explore the effect of parental temporary migration status on child's schooling progression. Using data from the IHDS-II, the results suggest that parental education is one of the significant determinants of child's schooling progression. The other important factors affecting the odds of child's transition to different educational levels include socioeconomic status of the household, sibling composition and place of residence. However, different factors affect the schooling progression of the child differently. Parental education is found to be positively linked with child's schooling transition to primary, secondary and post-secondary educational level. Child age is positively associated with all the educational transitions. Findings show that in case of sons, household income is significantly associated with child's transition to secondary and post-secondary level of education, while for the lowest educational transition the impact of household income turns out to be insignificant. The latter could be attributed to the low or almost free entry to primary education level. On the contrary, for daughters the impact of household income is insignificant for all the three education transitions.

Furthermore, we find that household assets indicating the overall welfare status of the family positively affects the child's schooling progression at all the three education levels. It is also found that, number of siblings in the household negatively affects the likelihood of child's schooling progression, but only for sons. It may be that large household size leads to competition for resources, which reduces the total welfare of the household leading to slow schooling progression.

Additionally, we perform a sensitivity analysis to scrutinize the consequences of unobserved variables on our main variables of interest i.e., parental education and migration status. However, the effect of both father's education and mother's education are likely to get underestimated in the simple sequential logit model, as these effects are consistently larger in settings with more unobserved heterogeneity. Furthermore, we find that these estimates are quite sensitive to variables which remain unobserved in the analysis.

Table 6.1: Variable definitions

Variables	Description of the variables
Individual characteristics	
Male	Dummy variable=1 for male, 0 otherwise.
Married	Dummy variable=1 for married individual, 0 otherwise.
Schooling	Non-Literate = 0, Primary = 1, Secondary =2 and Post-secondary=2
Child Education	Continuous variable: Completed years of education of the child.
Child Age	Continuous variable: Child's age in years at the time of survey
Child Age^2	Square of Child's age
Non literate	Dummy variable = 1 for Non-literate child, 0 otherwise.
Primary	Dummy variable = 1 for child with primary education, 0 otherwise.
Secondary	Dummy variable = 1 for child with secondary education, 0 otherwise.
Post-secondary	Dummy variable = 1 for child with post-secondary education, 0 otherwise.
Father education	Continuous variable: Completed years of education of father.
Mother education	Continuous variable: Completed years of education of mother.
Father age	Age of the father (in years).
Mother age	Age of the mother (in years).
Household characteristics	
Assets	Number of household assets with a maximum score of 33.
Urban	Dummy variable = 1for urban households, 0 otherwise.
Non-Hindu	Dummy variable = 1 for Non-Hindu Households, 0 otherwise.
OBC	Dummy variable = 1 for Other backward castes, 0 otherwise.
SC & ST	Dummy variable = 1 for Scheduled Castes and Tribes, 0 otherwise.
Children under 6	Dummy variable = 1 if household has children under 6, 0 otherwise
6-14 year olds	Dummy variable = 1 if household has 6-14 year olds, 0 otherwise.
Debt	Logarithm of total outstanding household debt at the time of survey.
Remittances	Logarithm of total remittances received by the household.
Only sons	Dummy variable = 1 if the father has only sons, 0 otherwise.
Only daughters	Dummy variable = 1 if the father has only daughters, 0 otherwise.
Siblings	Number of siblings in the household.
Migration characteristics	
Father migrant	Dummy variable = 1 if father is a return migrant, 0 otherwise.
Mother migrant	Dummy variable = 1 if mother is a return migrant, 0 otherwise.
Months of migration	Dummy variable = 1 if father is a return migrant, 0 otherwise.
Migrated with family	Dummy variable=1 if the return migrant had migrated with family, 0 otherwise.
Migrated without job	Dummy variable: If the return migrant had migrated without job, 0 otherwise

Table 6.1 (continued).

Place of origin characteristics	
Per-capita income	Per-capita income at the place of source of the migrants (in Rs. Thousands).
Non-farm	Proportion of non-farm workers to total workers at the place of source of the migrants

Table 6.2: Descriptive statistics for Non-migrant, Migrant, Continuing Migrant and Return Migrant samples

	All Workers			Migrant Workers		Tests of Equality	
	Total	non migrants	all migrants	continuing migrants	return migrants	$\widehat{\theta_1} - \widehat{\theta_2} = 0$	$\widehat{\theta_3} - \widehat{\theta_4} = 0$
	$\widehat{\theta_0}$	$\widehat{\theta_1}$	$\widehat{\theta_2}$	$\widehat{\theta_3}$	$\widehat{\theta_4}$	test statistic	test statistic
Number of workers	40,922	38,213	2709	1576	1133		
Personal characteristics							
Married (%)	76.22	76.14	77.45	76.27	79.08	2.39	2.99
Age	39.56	39.94	33.50	32.56	34.81	27.96**	-5.09**
Schooling	5.41	5.45	4.77	4.53	5.11	7.95**	-3.48**
Non-Literate (%)	32.95	32.84	34.70	36.38	32.36	3.98*	4.70*
Primary (%)	19.18	19.01	21.95	21.84	22.10	14.22**	0.03
Secondary (%)	36.61	36.59	36.95	36.95	36.96	0.14	0.00
Higher Secondary (%)	6.58	6.69	4.69	3.75	6.01	16.60**	7.56**
Graduate & above (%)	4.68	4.86	1.70	1.08	2.56	57.07**	8.68**
Migration characteristics							
Months of migration (months)	-	-	8.35	7.94	8.92	-	-2.40*
Migrated with family (%)	-	-	22.09	24.56	18.66	-	13.32**
Migrated without job (%)	-	-	11.23	9.69	13.41	-	8.72**
Household characteristics							
Assets	1.81	1.77	1.19	1.20	1.13	45.52**	3.36**
Urban							
Debt	114.26	113.31	112.88	123.95	97.11	0.03	1.08
Children under 6	0.55	0.55	0.76	0.77	0.75	0.04	-1.19
6 -14 year olds	0.94	0.93	1.10	1.11	1.08	-3.44**	0.38

Table 6.2 (continued).

Hindu (%)	84.28	84.14	86.49	87.06	85.70	10.61**	5716.94**
Muslim (%)	9.11	9.05	10.11	9.20	11.39	3.48	856.82**
General (%)	21.95	22.54	12.51	10.85	14.83	149.61**	1192.85**
SC/ST (%)	36.49	35.96	45.07	47.08	42.28	91.30**	2555.89**
OBC (%)	40.21	40.11	41.71	41.69	41.75	2.71	2746.63**
Only Sons (%)	14.94	14.85	16.52	17.15	15.82	4.54*	0.73
Only Daughters (%)	20.51	20.40	21.85	20.73	23.62	2.67	2.77
Place of origin characteristics							
Per-capita income	22.68	22.79	15.92	14.34	15.88	9.13**	-1.50
Non-farm (%)	17.37	17.68	13.23	11.17	15.90	7724.39**	1610.94**

Notes: We use chi-squared tests for comparing characteristics represented as proportions, and Welch’s t-test for comparing characteristics represented by mean values

Table 6.3: Marginal effects for determinants of Return Migration

Explanatory variables	Model 1		Model 2	
	Coefficients	Marginal Effects (%)	Coefficients	Marginal Effects (%)
Constant	-1.483**	–	-1.527**	–
Male	0.032	0.734	0.015	0.349
Married	0.079	1.824	0.081	1.864
Age	0.020**	0.471	0.021**	0.491
Age^2	-0.001**	-0.021	-0.001**	-0.021
Primary	0.151	3.492	–	–
Secondary	0.145	3.347	–	–
Higher secondary	0.450*	10.400	–	–
Graduate & above	0.705*	16.310	–	–
Years of education	–	–	0.029**	0.679
Months of Migration	0.008	0.180	0.008*	0.181
Migrated with family	-0.349**	-8.064	-0.342**	-7.908
Migrated without job	0.320*	7.400	0.321*	7.429
Debt	-0.007	-0.155	-0.007	-0.163
Children under 6	0.006	0.130	0.007	0.169
6–14 year olds	-0.081	-1.868	-0.081	-1.885
Muslim	0.258	5.974	0.274	6.349
SC & ST	-0.191	-4.426	-0.181	-4.178
OBC	-0.253	-5.840	-0.246	-5.703
Only daughters	0.023	0.532	0.024	0.553
Only sons	0.234*	5.411	0.230*	5.330
Per capita income	0.019**	0.445	0.019**	0.451
Non-farm	0.558**	12.900	0.542**	12.540
Number of Observations	2,709		2,709	

Notes: This table reports the marginal effects for the logit models with the dummy for return migration as the dependent (permanent return migrant=1, continuing migrant=0). *,**,*** indicates statistical significance at 10%,5%,1%, respectively.

Table 6.4: Descriptive statistics

Variable	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
Individual characteristics					
Child Gender	40922	0.53	0.50	0.00	1.00
Child Education	40922	6.00	3.79	0.00	16.00
Child Age	40922	13.20	4.18	6.00	20.00
Child Age^2	40922	191.55	110.66	36.00	400.00
Non-Literate	40922	0.09	0.29	0.00	1.00
Primary	40922	0.37	0.48	0.00	1.00
Secondary	40922	0.41	0.49	0.00	1.00
Post-secondary	40922	0.13	0.34	0.00	1.00
Father Education	40922	5.85	3.56	0.00	16.00
Mother Education	40922	3.71	4.49	0.00	16.00
Household characteristics					
Assets	40922	14.61	6.30	0.00	33.00
Urban	40922	0.32	0.47	0.00	1.00
Non-Hindu	40922	0.22	0.41	0.00	1.00
OBC	40922	0.43	0.49	0.00	1.00
SCST	40922	0.32	0.46	0.00	1.00
Children under 6	40922	0.26	0.43	0.00	1.00
6-14 year olds	40922	0.80	0.39	0.00	1.00
Debt	40922	5.15	5.28	0.00	16.11
Remittances	40922	0.58	2.33	0.00	13.22
Only Sons	40922	0.19	0.40	0.00	1.00
Only Daughters	40922	0.09	0.28	0.00	1.00
Siblings	40922	2.55	1.91	0.00	15.00
Migration characteristics					
Father Migrant	40922	0.06	0.23	0.00	1.00
Father's Months of Migration	40922	0.52	3.27	0.00	60.00
Mother Migrant	40922	0.01	0.10	0.00	1.00
Mother's Months of Migration	40922	0.09	1.38	0.00	60.00

Notes: We identify all parent child pairs IHDS-II dataset where, at the time of the survey, the child follows the following three conditions: (i) The child is aged between 6 to 20 years; (ii) The child is enrolled in the school; (iii) The child is not engaged in any labor market activity. We also omit those pairs where any of the variables described in Table 1 are not available. IHDS-II sampling weights are used to calculate all statistics.

Table 6.5: Estimated Odds ratio for Schooling Progression of the Child

Explanatory variables	Sons: Non-Literate to Primary	Sons: Primary to Secondary	Sons: Secondary to Post-secondary	Daughters: Non-Literate to Primary	Daughters: Primary to Secondary	Daughters: Secondary to Post-secondary
Father Education	1.066***	1.044***	1.084***	1.073***	1.053***	1.059***
	(0.008)	(0.007)	(0.008)	(0.009)	(0.007)	(0.009)
Father Migrant	0.829	0.718*	1.299	0.834	0.706*	0.769
	(0.119)	(0.111)	(0.354)	(0.125)	(0.109)	(0.291)
Father Education × Father Migrant	1.006	1.057*	0.903*	1.038	1.055*	1.006
	(0.025)	(0.026)	(0.036)	(0.030)	(0.028)	(0.051)
Mother Education	1.027**	1.046***	1.060***	1.047***	1.052***	1.088***
	(0.010)	(0.008)	(0.009)	(0.010)	(0.009)	(0.010)
Remittances	0.974*	0.971**	1.004	0.970**	0.967***	0.982
	(0.011)	(0.009)	(0.012)	(0.011)	(0.010)	(0.013)
Child Age	1.141***	1.384***	2.135***	1.130***	1.374***	2.147***
	(0.004)	(0.010)	(0.082)	(0.004)	(0.010)	(0.086)
Child Age^2	1.000***	0.999***	0.998***	1.000***	0.999***	0.998***
	0.000	0.000	0.000	0.000	0.000	0.000
Months of Migration	0.988	0.993	1.005	0.997	0.992	0.990
	(0.009)	(0.009)	(0.014)	(0.009)	(0.010)	(0.017)
Assets	1.063***	1.088***	1.093***	1.075***	1.095***	1.105***
	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.008)
Non-Hindu	0.615***	0.586***	0.829*	0.640***	0.648***	0.704***
	(0.041)	(0.035)	(0.062)	(0.045)	(0.041)	(0.058)
OBC	0.847*	1.151*	1.064	0.696***	0.947	1.000
	(0.064)	(0.071)	(0.076)	(0.058)	(0.064)	(0.080)
SC & ST	0.950	1.021	0.874	0.838	1.024	0.858
	(0.079)	(0.068)	(0.070)	(0.076)	(0.075)	(0.076)
Urban	0.680***	0.604***	0.799***	0.673***	0.573***	0.857*
	(0.102)	(0.097)	(0.084)	(0.111)	(0.112)	(0.088)
Children under 6	0.754***	0.712***	0.648***	0.684***	0.584***	0.751**
	(0.045)	(0.044)	(0.060)	(0.043)	(0.035)	(0.075)
6-14 year old	0.651***	0.804*	0.886	0.551***	0.717**	0.806**
	(0.077)	(0.069)	(0.058)	(0.071)	(0.074)	(0.059)
Debt	1.004	1.005	1.021***	0.994	1.000	1.016**
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
Only Sons / Only Daughters	0.847*	0.778***	0.875*	1.026	1.137	1.177
	(0.057)	(0.043)	(0.059)	(0.080)	(0.078)	(0.102)
Siblings	0.824***	0.819***	0.907**	0.984	0.948***	1.009
	(0.018)	(0.018)	(0.028)	(0.013)	(0.012)	(0.018)

Notes: This table reports exponentiated logit coefficients (odds ratios) for the sequential logit models of schooling progression. IHDS-II sampling weights are used. Robust standard errors clustered by primary sampling unit (PSU) are reported in parentheses. *, **, and *** denote significance at 5 percent, 1 percent, and 0.1 percent levels, respectively.

Table 6.6: Sensitivity analysis for Sons

	Baseline 1	Baseline 2	Odds ratio		p-value	
Explanatory variables	$\rho = 0;\ \beta_u = 0$	$\rho = 0;\ \beta_u = 2.5$	Minimum	Maximum	Minimum	Maximum
Non-literate to primary transition						
Father Education	1.066***	1.113***	0.804***	1.502***	0.000	0.939
Father Migrant	0.829	0.700	0.001***	482.7***	0.000	0.951
Father Education × Father Migrant	1.006	1.021	1.006	1.043	0.557	0.807
Mother Education	1.027**	1.042**	1.023*	1.453***	0.000	0.033
Primary to secondary transition						
Father Education	1.044***	1.090***	1.033**	1.471***	0.000	0.004
Father Migrant	0.718*	0.552*	0.001***	388.2***	0.000	0.946
Father Education × Father Migrant	1.057*	1.100*	1.057*	1.177*	0.034	0.046
Mother Education	1.046***	1.077***	1.022*	1.495***	0.000	0.020
Secondary to post-secondary transition						
Father Education	1.084***	1.165***	1.064***	1.561***	0.000	0.000
Father Migrant	1.299	1.318	0.002***	846.5***	0.000	0.984
Father Education × Father Migrant	0.903*	0.864*	0.800*	0.903*	0.005	0.012
Mother Education	1.060***	1.111***	0.783***	1.534***	0.000	0.566

Notes: This table presents a summary of results obtained from sensitivity analysis conducted for parent-son pairs. For each educational transition, we consider a total of 32 scenarios. The first eleven scenarios assume that the unobserved heterogeneity (u) is uncorrelated ($\rho = 0$) with our main explanatory variables, and the effect of u in term of log odds, β_u, is varied from 0 to 5 (with β_u increasing in increments of 0.5 starting from β_u = 0, β_u = 0.5, and so on, till β_u = 5). Next, we estimate seven scenarios where β_u is fixed at 2.5, and the correlation (ρ) of u with father's education is varied from -0.6 to +0.6 (with ρ increasing in increments of 0.2 starting from ρ = -0.6, ρ = -0.4, and so on, till ρ = +0.6). Similarly, we compute seven scenarios where u is correlated with mother's education and another seven scenarios where u is correlated with father's return migration. This results in a total of 32 logit coefficients for each variable, out of which we report the minimum & maximum odds ratio (exponentiated logit coefficient), and minimum & maximum p-value indicating the significance of the logit coefficients. We also report odds ratios for two baseline models: Baseline 1 model, with $\rho = 0$ and $\beta_u = 0$; and Baseline 2 model, with $\rho = 0$ and $\beta_u = 2.5$.

*, **, and *** denote significance at 5 percent, 1 percent, and 0.1 percent levels, respectively.

Table 6.7: Sensitivity analysis for Daughters

Explanatory variables	Baseline 1 $\rho = 0;\ \beta_u = 0$	Baseline 2 $\rho = 0;\ \beta_u = 2.5$	Odds ratio Minimum	Odds ratio Maximum	p-value Minimum	p-value Maximum
Non-literate to primary transition						
Father Education	1.073***	1.126***	0.811***	1.520***	0.000	0.356
Father Migrant	0.834	0.771	0.001***	458.3***	0.000	0.951
Father Education × Father Migrant	1.038	1.078	1.038	1.144	0.138	0.209
Mother Education	1.047***	1.068***	1.031**	1.482***	0.000	0.007
Primary to secondary transition						
Father Education	1.053***	1.109***	0.799***	1.496***	0.000	0.868
Father Migrant	0.706*	0.548*	0.001***	342.6***	0.000	0.941
Father Education × Father Migrant	1.055*	1.102*	1.055*	1.184*	0.047	0.058
Mother Education	1.052***	1.094***	1.035***	1.508***	0.000	0.001
Secondary to post-secondary transition						
Father Education	1.059***	1.125***	0.809***	1.515***	0.000	0.426
Father Migrant	0.769	0.543	0.001***	344.6***	0.000	0.975
Father Education × Father Migrant	1.006	1.034	1.006	1.063	0.682	0.906
Mother Education	1.088***	1.161***	1.029***	1.590***	0.000	0.013

Notes: This table presents a summary of results obtained from sensitivity analysis conducted for parent-daughter pairs. For each educational transition, we consider a total of 32 scenarios. The first eleven scenarios assume that the unobserved heterogeneity (u) is uncorrelated ($\rho = 0$) with our main explanatory variables, and the effect of u in term of log odds, β_u, is varied from 0 to 5 (with β_u increasing in increments of 0.5 starting from β_u = 0, β_u = 0.5, and so on, till β_u = 5). Next, we estimate seven scenarios where β_u is fixed at 2.5, and the correlation (ρ) of u with father's education is varied from -0.6 to +0.6 (with ρ increasing in increments of 0.2 starting from ρ = -0.6, ρ = -0.4, and so on, till ρ = +0.6). Similarly, we compute seven scenarios where u is correlated with mother's education and another seven scenarios where u is correlated with father's return migration. This results in a total of 32 logit coefficients for each variable, out of which we report the minimum & maximum odds ratio (exponentiated logit coefficient), and minimum & maximum p-value indicating the significance of the logit coefficients. We also report odds ratios for two baseline models: Baseline 1 model, with $\rho = 0$ and $\beta_u = 0$; and Baseline 2 model, with $\rho = 0$ and $\beta_u = 2.5$.

*, **, and *** denote significance at 5 percent, 1 percent, and 0.1 percent levels, respectively.

Table 6.8: Determinants of Schooling Progression among Return Migrants, Marginal Effects

	Sons			Daughters		
	Non-Literate to Primary	Primary to Secondary	Non-Literate to Primary	Non-Literate to Primary	Primary to Secondary	Secondary to Post-secondary
Father Education	0.007***	0.006***	0.010***	0.010***	0.006***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)
Mother Education	0.003***	0.005***	0.006	0.006***	0.009***	0.010***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Child Age	0.086***	0.122***	0.116***	0.116***	0.060***	0.036***
	(0.004)	(0.006)	(0.005)	(0.005)	(0.001)	(0.001)
Child Age^2	-0.002***	-0.003***	-0.004***	-0.004***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban	0.006**	0.003	-0.004	0.004	-0.002**	-0.005
	(0.003)	(0.003)	(0.005)	(0.003)	(0.001)	(0.004)
Months of Migration	-0.001	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Assets	0.007***	0.006***	0.005**	0.005**	0.006***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)
Non-Hindu	-0.023	-0.049**	-0.044**	-0.044**	-0.050***	-0019***
	(0.021)	(0.022)	(0.022)	(0.022)	(0.004)	(0.004)
OBC	-0.016	0.005	-0.054*	-0.054*	.008*	-0.008**
	(0.024)	(0.025)	(0.029)	(0.029)	(0.004)	(0.003)
SC Sr ST	0.006	-0.017	-0.039	-0.039	-0.013***	-0.006
	(0.026)	(0.026)	(0.030)	(0.030)	(0.005)	(0.004)
Children under 6	-0.035**	-0.055***	-0.037**	-0.037**	-0.095***	-0.046***
	(0.016)	(0.019)	(0.018)	(0.013)	(0.004)	(0.003)
6-14 year old	0.147***	0.068***	0.029	0.029	-0.021***	-0.048***
	(0.026)	(0.026)	(0.030)	(0.030)	(0.005)	(0.003)
Debt	0.000	0.000	0.000	0.000	0.001*	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)
Siblings	-0.019***	-0.023***	0.001	0.001	0.000	0.001
	(0.006)	(0.006)	(0.003)	(0.003)	(0.001)	(0.001)

Notes: *,**,*** indicates statistical significance at 10%,5%,1%, respectively.

Chapter 7

Schooling Gap in India: Financial/Non-Financial Aid and Intergenerational Educational Mobility

7.1 Introduction

Income disparity among different socioeconomic groups is one of the most common phenomena of the Indian labour market. Research indicates that wage inequality is still persistent among workers with similar level of education, skills and other characteristics (Das, 2012; Kijima, 2006). Wage discrimination at the workplace has unfavourable effects on the economic outcomes of the society. A society with more inequality is said to be associated with less IEM (Piketty, 2000). This immobility increases the income consequences as result of inherent differences between the individuals.

Schooling is associated with many economic outcomes, including IEM. If an individual's schooling level exert a considerable impact on his/her income prospects and if level of schooling is strongly associated with the household characteristics, intergenerational correlation in income will be high. However, if household characteristics do not affect the schooling level of an individual, intergenerational correlation in income will be low.

Evidence suggests that children from poor family or those who belong to disadvantaged section of the society reap huge benefit if they attain higher education. Even though the expected returns to higher education is quite large for children belonging to poor families, still data shows that they are less likely to continue into post-compulsory education. Although the difference in the highest level of education attained by children belonging to high and low income households, in part, could be attributed to the inherited ability and effort of an individual, but still after controlling for these factors results show that high-income children fares well than low-income children in terms of grade completion. However, past research work indicates that these differences could be attributed to financial constraints faced by low-income families in bearing the underlying cost of higher education as well as lack of information and student aid.

Therefore, adhering to the concept of ‘equality of educational opportunity’ in promoting economic mobility among the disadvantaged sections of the society becomes inevitable. In this chapter, we investigate whether financial or non-financial aid affects the child’s educational attainment. We also investigate whether this effect differ based on parental education level and the given education level of the student. In short, this chapter addresss the following question:

i. If a student receives fellowship or other aid (free uniform/free books), does this affect the educational outcome of that children?
ii. Does the impact of encouragement (financial or non-financial aid) on child’s educational outcomes vary by parental education level?

This chapter is organized as follows. Section 7.2 discusses the theoretical background and previous research work. In Section 7.3, we explain the data source and the empirical model used for the analysis. Section 7.4 discusses the results. Section 7.5 concludes the chapter.

7.2 Theoretical literature

Higher education has become one of the major concerns for the economic development of an individual as well as the overall development of the society in which these individuals reside because higher education level tends to increase the earning potential of the individuals by making them more productive (P. W. Bauer et al., 2012; Faggian & Mccann, 2009). There are many studies that have analysed the positive association between educational level and wage in the labour market. It has been empirically shown that higher education increases the odds of securing high-wage occupation (Dickson & Harmon, 2011; Psacharopoulos & Patrinos, 2004). However, there are also many positive externalities linked to education. For example., educated individuals are capable of paying higher taxes and are likely to use less public services (Trostel, 2010). Many studies also show that increased regional proportion of highly-educated workers increases the probability of employment opportunities for other individuals, even if the latter are not well-educated (Glaeser & Resseger, 2010; Iranzo & Peri, 2009; Rauch, 1993)

Given the numerous benefits associated with higher education, it becomes important for policymakers to determine factors which alter distribution of education. One important factor could be merit-based financial aid to students who have high academic achievement score (Groen, 2011). However, the impact of these type of financial aid on higher education outcomes is still not empirically tested. There are limited research studies which have examined the role of financial support on higher education attainment. Dynarski (2008) and Sjoquist & Winters (2012) use 2000 Census data to analyse the effect of merit-aid programmes in Georgia and Arkansas. Dynarski (2008) find a positive and statistically significant association between merit-aid and college graduation rate while the latter results show an insignificant relationship.

It is also been argued that provision of school meals is closely associated with child's educational level (Afridi, 2011). This could be attributed to two major reasons. First, providing meals at school can fulfil the nutritional needs of the children, thereby increasing their learning capacity (Grantham-McGregor et al., 1997; Jacoby & And Others, 1996). However, some studies give a counterargument to the economic impact of provision of school meals. They argue that families whose children are beneficiary of school meals, make an adjustment to school meals by reducing the resources allocated to the children and transferring them to other members of the households. Second, this also provides an incentive for families to send their kids to school regularly.

Similarly, some of the recent public interventions allowed many researchers to analyse the short- term impact of free uniform provision on child's educational outcomes. In Kenya, a program to distribute two school uniforms during the last three years of primary level resulted in lower dropout rates, irrespective of child's gender. Another program in India provided free uniforms in government school, which resulted in improving the school enrolment rate of children belonging to the disadvantaged section of the society.

7.3 Data and Methodology

7.3.1 Data Source

We use data from both IHDS-I (2004-05) and IHDS-II (2011-12). The first part of analysis focuses on children aged between 10-21 years. Further, we also do the analysis independently for different age-groups: 10-12, 13-15, 15-18 and 18-21. The segregation of age group will allow us to control for the non-linear impact of family background on the schooling of the child, where family background is likely to have a considerable impact on schooling of older children. Another benefit of considering age-groups separately is that it mitigates the risk of biasness associated with unobserved heterogeneity and omitted variables.

7.3.2 Measuring Schooling Gap

We measure schooling gap of child as the difference between the anticipated years of schooling and the realised years of schooling attained by the child at the time of survey. The anticipated years of schooling is estimated as years of schooling that the child would have completed at the time of survey, had s/he started formal schooling at the age of 6 and progressed one class each year. In short, schooling gap is measured as follows

Schooling Gap = (Anticipated years of Schooling – Realised years of Schooling)

Where anticipated years of schooling is estimated as years of schooling that the child would have attained at the time of survey, had s/he started formal schooling when they turned 6 and progressed one grade each year and realised years of schooling is the actual educational level of the child at the time of survey.

7.3.3 Determinants of Schooling Gap

Studies show that the level of child's schooling is associated with parental education and other household variables. To capture the extent of IEM in terms of education, we explore the variation in schooling gaps at different age-groups by different level of parental education. It is important to note that this will allow us to capture the association between parental education and child schooling, and not causal effect. The estimation technique to study the impact of household structure on the schooling gap of the child is as follows:

$$SGAP_i = a_0 + a_1 Edu_{Fi} + a_2 Edu_{Mi} + a_3 Control_{hi} + e_i \tag{7.1}$$

We regress schooling gap of the i^th child ($SGAP_i$) on father's schooling (Edu_{Fi}), mother's schooling (Edu_{Mi}) , other household characteristics ($Control_{hi}$) and a stochastic disturbance term (e_i). Here, years of schooling gap (SGAP) is used as a proxy for child's educational achievement. The value of coefficients 'a_1' and 'a_2' indicate the impact of father's and mother's educational level on child's educational achievement, respectively.

We also employ the counterfactual framework of Rubin (1974) to estimate the causal impact of parental education on their child's schooling gap as an alternative strategy to check the robustness of our estimates. The parental education level is used as a dichotomous treatment variable which is 1 if the parental education exceeds certain threshold and 0 otherwise. The former set of children comprise the treatment group, whereas the latter set of children comprise the control group. The outcome variable is the education attainment of their child measures as years of schooling gap.

7.3.4 Estimating the Impact of Aid on Intergenerational Educational Mobility

7.3.4.1 Data Source

We use IHDS dataset to measure the impact of encouragement in the form of monetary or non-monetary aid on child's schooling gap. The source of information on aid is the IHDS-I dataset, and the data on schooling gap comes from the IHDS-II dataset. IHDS dataset provides information on whether the child received any kind of monetary (scholarship or free fees) or non-monetary (free books, free uniforms, etc.,) in the last one year from the date of survey. We exploit this information to produce dataset containing personal and socioeconomic traits of children and their respective parent who received any kind of aid. In addition, we also utilize information from IHDS-II to track down the progress of the child's educational progress for children whose information is available for both the rounds (IHDS-I and IHDS-II). We use schooling gap as an indicator of child's educational progress. To measure the impact of encouragement on schooling gap requires to calculate the years of gap in child's schooling which occurs between the two survey rounds. In short, we calculate the outcome variable as follows:

$$SGAP_i = ((EDU_{Ei}^{2012} - EDU_{Ai}^{2012}) - EDU_i^{2005}) \qquad (7.2)$$

Where $SGAP_i$ is the years of gap in schooling of the i^{th} child. EDU_{Ei}^{2012} is the estimated completed years of education of the i^{th} child if the latter continues his/her education progressing one class every year. EDU_{Ai}^{2012} is the realised years of education completed by the i^{th} child during IHDS-II survey.

Our sample is limited to children who were still enrolled during survey or just completed their education. Therefore, the upper age limit has been fixed at 15 years at the time of IHDS-I so that the child who is 15 years old in IHDS-I will turn 21-year-old in IHDS-II. That means, if this child has continued education advancing one grade per year then he would have been pursuing post-graduation at time of IHDS-II survey.

Further, we also analyse the impact of encouragement for different age-groups, separately. First, this allows us to control for factors that affects marginal decision to move to the next educational level. Second, it will help us to identify differential impact of encouragement on schooling gap of child belonging to different age-groups, if any.

7.3.4.2 Empirical Model

Our estimation is based on the assumptions that it is unlikely that children randomly avail aid which encourages them to continue to higher education. For example, government is more likely to support children from families belonging to disadvantaged section of the society or those with better academic history. Since, the education attainment of the child is positively correlated to the level of schooling gap, a simple correlation between those who received government benefit with those who didn't receive would not suffice. It may introduce biasness by overestimating the impact of aid on schooling progression. Further, there is suspicion of few other factors such as social class, ethnicity, annual household income, that can have an impact on the schooling progression of the child. However, our main hypothesis is that even after controlling for different sources of biasness, children who receive monetary or non-monetary benefits (fellowship, free books, uniforms, etc.,) are more likely to progress, without any discontinuation.

7.3.4.3 Model Variables

Outcome of Interest

Our outcome for the analysis is the schooling gap for the child between the two survey rounds of IHDS. The panel data allows us to calculate the gap in the schooling of the child between 2004-05 and 2011-12. We use this schooling gap as a proxy for educational progress of the child and our outcome of interest.

Key Explanatory variables

This analysis focuses on two main explanatory variables, each of which is a dummy variable. The first is whether the child received fellowship in 2005-06. The second is whether the child received any kind of non-monetary aid (free uniform/free books/fees).This binary variable is 1 if the child received any form of aid, 0 otherwise.

Model Controls

Our main aim is to study the causal impact of any kind of aid on child's schooling gap, which is being used as a proxy for child's educational progress. Therefore, it is necessary to control for other factors that are linked both to our treatment (child receiving any kind of aid) and outcome (schooling gap). We include caste, religion, gender, age, household size, whether the child ever failed/ repeated class, household income quintile, type of school as our controls for the model. All the information on these variables has been taken from IHDS-1(2004-05) survey, ensuring an overlap between the time when treatment is received by the child and the socio-economic condition of the household at that time.

7.3.4.4 Methodology

We employ the counterfactual framework of Rubin (1974) to estimate the effect of government aid on child's schooling gap. The 'government aid' variable is used as a binary treatment variable which is 1 if the child received any kind of government aid (fellowship/ free books/free uniform/ school fees) in 2004-05 and 0 otherwise. The former set of children comprise the treatment group, whereas the latter set of children comprise the control group. The outcome variable is the schooling gap in education between 2004-05 and 2011-12. Here, it is been assumed that any policy intervention comes with some level of lagged effect.

While the outcome variable for each child is observed only under one of the two possible treatment states, the counterfactual framework suggests that each child has a potential outcome under both treatment states (Morgan & Winship, 2015). More formally, suppose the outcome variable for child i is $Y_i(T)$, where T is the treatment variable which is 1 if the child government aid in 2004-05 and 0 otherwise. Then, the treatment effect for the child, τ_i, is defined as follows

$$\tau_i = Y_i(1) - Y_i(0) \tag{7.3}$$

For each child, one of the two terms is always missing in our data. For example, in the treatment group we observe $Y_i(1)$ but we do not observe $Y_i(0)$, whereas for the children in the control group we observe $Y_i(0)$ but we do not observe $Y_i(1)$. A naïve approach for solving this problem is to compare the average schooling gap of the treatment group with that of the control group. This approach can be applied if the treatment assignment is random, and the children in treatment and control group are similar in all other characteristics that may potentially affect their educational outcome (schooling gap), i.e., both groups have balanced covariate distributions. Since our study is not based on randomized trials but on observational data, the assumption of balanced covariate distributions is problematic due to the non-random assignment of treatment.

For example, suppose there is no impact of aid on schooling gap, however, individuals belonging to affluent households are more likely to have a higher educational attainment. In this case, we would see that, on an average, the treatment group comprises of parent-child pairs from more affluent households as compared to those in the control group. If follows that the parent-child pairs in the treatment group are more likely to have higher educational outcomes that those in the control group. In this scenario, a naïve comparison of the average educational outcome of the treatment and control groups would suggest a spurious causal relationship between parental education and the child education, while the actual relation is driven by the difference in household wealth between the treatment and control groups.

We use PSM to reduce the differences between the treatment group and the control group for a large set of potential confounders variables such as household income, prior academic achievement, etc. The matching approach allows us to estimate quasi-experimental contrasts between the outcomes of the treatment and the control groups by matching the individuals belonging to the treatment units with comparable individuals from the control group, and discarding the unmatched individuals in the control group (Morgan & Winship, 2015). The ATT is then estimated as follows

$$ATT = \frac{1}{N_T} \sum_{i,T_i=1} \left(Y_i(1) - \hat{Y}_i(0) \right) \tag{7.4}$$

Where N_T is the number of children in the treatment group. T_i is a binary treatment indicator that takes a value of 1 if the child received any kind of government aid in 2004-05 and 0 otherwise. ATT measures the average effect of treatment on only those children that received the treatment, and therefore, we restrict the summation to children belonging to the treatment group ($T_i = 1$). $Y_i(1)$ is the educational outcome (gap in schooling between IHDS-I(2004-05) and IHDS-II (2011-12)) of child i in the treatment group, and $\hat{Y}_i(0)$ is the counterfactual control outcome for the matched counterpart of the child i.

We estimate this counterfactual outcome using PSM. Using a big set of personal and household characteristics, we run a logistic regression model to determine each child's propensity score (P_i), i.e., the probability that child i receives the treatment

$$\log \frac{P_i}{1-P_i} = \alpha + \beta_k X_i + \varepsilon_i \tag{7.5}$$

Where β_k is a vector of k estimated coefficients corresponding to the individual and household characteristics, X_i is a vector of k individual and household characteristics for child i, and ε_i is a random error term that is logistically distributed.

7.4 Results

Table 7.1 presents summary statistics. Our main dependent variable is the years of schooling gap. Statistics show that the average years of schooling gap has declined from 2.2 years in 2004-05 to 1.8 years in 2011-12. The average gap in schooling is larger for sons than daughters, for both survey rounds. However, the gap has declined more for the sons, while for daughters the average schooling gap does not see any major improvement over the two survey rounds. Around 66% of the total households belong to the rural sector. Regarding household composition, around 20% of the total number of household members comprises of those over 65 years of age. In addition, there are on an average 2 children who are below the age of 15 in each household. The proportion of children who go to public schools is around 45% but this proportion is higher for males than females, for both the survey rounds. The average years of education has increased for both the parents over time.

Table 7.2 shows the years of schooling gap for children belonging to age-groups: 10 to 12 years, 13 to 15 years, 16 to 18 years and 19 to 21 years. This is done to dodge any kind of sample selection bias as children with higher ability have higher odds to continue higher education than children with low ability. In addition, it has also been found that family background effect is much larger at the lower end of educational distribution which gradually dissipates as children move up the education ladder (Cameron and Heckman, 1998). Statistics show that the size of schooling gap has reduced between the two survey rounds of IHDS, across all age-groups. The average schooling gap increases with the age of the child but remains on an average higher for sons than daughters. For example, the average schooling gap for sons of age-group 19-21 is 5.47 years in 2004-05 which reduced to 4.38 years in 2011-12 This means that on an average a male child belonging to this age-group was lagging more than 5 years behind than the expected schooling in 2004-05 while this lag on educational front improved by more than one year in 2011-12. For females this gap was 4.85 years in 2004-05 and 3.61 years in 2011-12. Results on test of equality show that the difference between average schooling gap of males and females become insignificant over time, with an exception in case of children belonging to the age-group 19-21. This suggest that gender inequality in higher education is still prevalent in the society. In addition, average schooling gap is more for children residing in rural areas than their counterparts. The marginal decision to continue schooling is likely to depend on family background characteristics. Previous studies find that parental education is the strongest factor affecting child's schooling. It may also represent genetic ability of the children. Table 7.3 and 7.4 shows the descriptive statistics regarding average schooling gap for children by father and mother's educational level, respectively. The average schooling gap reduces with parental education level. Figures show that differences between the mean schooling gap between the children of low-educated parents and children of high-educated parents has not reduced between the two rounds of survey. This signals that high degree of persistence is still prevalent in the economy where parental education is a strong determinant of child's schooling level.

Table 7.5 gives the average schooling gap by household income quintiles, for both sons and daughters. The average schooling gap declines with increasing household income quintile, regardless of child's gender. This suggests that household income is a crucial factor which affects the schooling gap in children. Statistics show that the difference between the mean schooling gap between children of bottom and top income quintile households has fell between two rounds of IHDS, both for sons and daughters. However, the reduction in difference is larger for sons than daughters.

7.4.1 Determinants of Schooling Gap

Table 7.6 reports the regression results for the schooling gap for both rounds of IHDS, separately for both sons and daughters. All regressions control for type of school dummy, age and age-squared of the child. In addition, we also control for other household characteristics like paternal education and their age and age-squared, sector (rural/urban), ethnicity, caste, household composition (number of elders and children). Our dependent variable is the years of schooling gap for the child and main explanatory variables are the educational level of both the parents. Results show that the impact of parental education is negative and significant at 1% level of significance. An extra year of father's education reduces the schooling gap by 0.21 and 0.17 years for sons and daughters in 2004-05, respectively. However, results for 2011-2012 data shows that the difference in the impact of father's education on sons and daughters has narrowed down. Residing in a rural sector raises the schooling gap for the child. Regarding social groups, significant and negative estimates for forward caste show that belonging to the latter decreases the average schooling gap for the child. Ethnicity also plays a vital role in shaping the educational progress of the child. The significant and positive coefficients for Muslims show that religious minority groups experience relative disadvantage in terms of schooling gap as compared to their dominant counterparts. Finally, the household composition variable comes out to be insignificant, implying no systematic difference in household compositions related to schooling gap. The coefficients of household income are significant and negative showing that higher household income reduces the average schooling gap of the child. However, the impact is much higher for daughters than sons. This means that household income is more important determinant of educational progress of the daughters than sons.

The age of the child is found to be positive and significant showing that older children have higher average schooling gap than their young counterparts. This result is in line with the human capital model which argues that children who complete their education at an appropriate age have greater advantage as they get more post-schooling time to maximize their returns to education. The coefficients of public school are significant and negative, i.e., studying in a public school reduces the schooling gap.

Table 7.7 estimates interaction-inclusive models to measure the effect of parental education on the child's schooling gap. We include the interactions of the parental education variable with household income, caste, religion and rural/urban indicator variable. We find that an increase in father's education reduces the schooling gap of their children, and this reduction is larger for Muslims and non-forward class households. The reduction in schooling gap is also larger for rural households as compared to urban households, however, the difference is not always significant.

For sons, the reduction in schooling gap with an increase in father's education is larger for higher income households. Similarly, increase in mother's education reduces the schooling gap of their children, and this reduction is also larger for Muslims, non-forward class and rural households. The interaction of mother's education with household income is generally insignificant.

7.4.2 Effect of Parental Education on Schooling Gap using Matching Estimator

Table 7.8 reports the ATT estimates based on PSM. The Abadie-Imbens standard errors are reported below the estimates in parentheses. Gamma (Γ) is the odds ratio of receiving treatment due to unobserved factors such as genetic endowments. Γ=1 implies that unobservable factors have no influence on treatment assignment. Γ=1.2 implies unobserved confounders can increase the odds for selection into treatment by 20%. "HL estimate bounds" are the Hodges-Lehmann bounds for the ATT estimate. "WSR p-value bounds" are the lower and upper bounds for the p-values from Wilcoxon's signed rank test. The treatment is defined as a binary variable that takes a value of 1 if parental education is high (defined as primary or above), and 0 otherwise. We estimate that the ATT of father's education on the child's schooling gap is -1.68 (-1.28) years for sons and -1.18 (-1.21) years for daughters for IHDS-I (IHDS-II), and the effect is statistically significant at the 1% level. The ATT of high maternal education on the child's schooling gap is -1.43 (-1.33) years for sons and -1.30 (-1.33) years for daughters for IHDS-I (IHDS-II), and the effect is statistically significant at the 1% level. The ATT estimates are robust to introduction of unobserved confounders. For example, even at Γ=1.6 which implies that unobserved confounders can increase the odds for selection into treatment by 60%, both HL estimate bounds and WSR p-value bounds indicate that the ATT is negative and statistically significant for all four parent-child combinations.

To assess the robustness of our results, we estimate ATT under two alternative definitions of high education, viz., middle and above and secondary and above. Additionally, we replicate the entire analysis with Mahalanobis matching to evaluate whether our results are vulnerable to the choice of matching methods. These additional robustness tests are reported in Tables 7.9 to 7.13. In all robustness tests, find that the results are qualitatively similar, the estimated ATT is negative and statistically significant for all four parent-child combination even at large values of the sensitivity parameter Γ.

7.4.3 Effect of Encouragement on Schooling gap

Tables 7.15 to 7.18 report effect of various modes of encouragement on the schooling gap of children. More specifically, we define five modes of encouragement, viz., scholarship (fellowship), mid-day meal, free books, free uniform and school fees provided by the government. Each of these modes is coded as a binary indicator variable that represents a treatment for our matching models. The covariates defined in Table 7.14 are used for matching the treatment units and control units. We use two matching estimators: Mahalanobis matching estimator and Epanechnikov kernel estimator with PSM. There are two major findings. First, among all the different modes of providing encouragement to the students, scholarship is the most effective method of reducing the schooling gap. Second, the benefits of providing encouragement in reducing the schooling gap are considerably larger for girls that those for boys. In Table 7.15, Panel A, we observe that scholarship reduces the schooling gap of sons by 0.32 years for the youngest cohort, aged 6 to 9 at the time of IHDS-I survey. Panel B and Panel C estimate the ATT for subsamples based on different levels of parental education. We find that for sons that have mothers with low education (1 to 5 years), the most effective modes of reducing the schooling gap are scholarship (ATT = -0.432 years) and school fees (ATT = -0.613 years). For sons that have mothers with high education (11 years and above), the most effective modes of reducing the schooling gap is free books (ATT = -0.733 years). In Table 7.16, we observe that regardless of the age of the daughter, scholarship reducing the schooling gap by around 0.4 year. The estimated ATT for daughters are more significant than those estimated for sons, for whom only the youngest cohort benefits from scholarship. In addition, we find that daughters of non-literate mothers that receive the scholarship have significantly lower schooling gap than those who do not receive scholarship. For all other forms of encouragement, we find weak evidence that they reduce schooling gap of children. For example, free books reduce schooling gap for daughters across all age-cohorts, but this effect is statistically significant only for daughters aged 10 to 12 years. The estimated ATT in tables 7.15 and 7.16 are based on the Mahalanobis matching estimator. We replicate these analyses with Epanechnikov kernel estimator with PSM. The results are qualitatively similar, and they are reported in Tables 7.17 and 7.18.

Tables 7.19 and 7.20 report the match quality for sons and daughters. Rubin (2001) specifies that a B value below 25 and value of R between 0.5 and 2, indicates a balanced control group. We find that Rubin's B and Rubin's R are within the acceptable range for all matching models which suggests that matching is efficient.

7.6 Conclusion

In this Chapter, we measured years of schooling gap of the child and studied the trend and pattern in schooling gap. The analysis in this chapter was based on panel information provided by two consecutive rounds of IHDS: IHDS-I and IHDS-II. Results show that the size of schooling gap has reduced between the two survey rounds of IHDS, across all age-groups. In addition, we find that the average schooling gap increases with the age of the child but remains on an average higher for sons than daughters. suggest that gender inequality in higher education is still prevalent in the society. In addition, average schooling gap is more for children residing in rural areas than their counterparts. The marginal decision to continue schooling is likely to depend on family background characteristics. Previous studies find that parental education is the strongest factor affecting child's schooling. It may also represent genetic ability of the children.

Results regarding the average schooling gap by household income quintiles show that the gaps are larger for the household who belong to the lower income quintile. The average schooling gap declines with increasing household income quintile, regardless of child's gender. This suggests that household income is a crucial factor which affects the schooling gap in children. Statistics show that the difference between the mean schooling gap between children of bottom and top income quintile households has fell between two rounds of IHDS, both for sons and daughters. However, the reduction in difference is larger for sons than daughters.

In this chapter, we also examined the impact of encouragement on child's educational outcomes. The educational outcome has been defined as the years of schooling gap of the child. We define five modes of encouragement, viz., scholarship (fellowship), mid-day meal, free books, free uniform and school fees which a child received. Our results showed that among all the different modes of providing encouragement to the students, scholarship is the most effective method of reducing the schooling gap. Second, the benefits of providing encouragement in reducing the schooling gap are considerably larger for girls that those for boys.

We also find that for sons that have mothers with low education (1 to 5 years), the most effective modes of reducing the schooling gap are scholarship and school fees. For sons that have mothers with high education (11 years and above), the most effective modes of reducing the schooling gap is free books. We also observe that regardless of the age of the daughter, scholarship reduces the schooling gap by around 0.4 year. The estimated ATT for daughters is more significant than those estimated for sons, for whom only the youngest cohort benefits from scholarship. In addition, we find that daughters of non-literate mothers that receive the scholarship have significantly lower schooling gap than those who do not receive scholarship. For all other forms of encouragement, we find weak evidence that they reduce schooling gap of children

Table 7.1: Descriptive statistics

	IHDS-I			IHDS-II			Tests of Equality		
Variables	Total	Sons	Daughters	Total	Sons	Daughters	$\theta_4 - \theta_1$	$\theta_2 - \theta_3$	$\theta_5 - \theta_6$
	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	(statistic)	(statistic)	(statistic)
Schooling gap (yrs)	2.252	2.568	1.615	1.807	1.985	1.591	-15.27***	20.76***	10.94***
Household Income(quintile)	3.001	3.000	3.002	3.000	3.000	3.000	-0.06	-0.11	0.00
Father Age (yrs)	44.889	45.301	44.058	45.913	46.114	45.670	14.85***	11.22***	5.01***
Mother Age (yrs)	39.468	39.876	38.645	40.715	40.923	40.462	19.88***	12.30***	5.71***
Child Age (yrs)	14.498	15.021	13.443	15.244	15.436	15.011	24.64***	34.75***	10.91***
Rural (%)	0.665	0.680	0.636	0.683	0.685	0.679	-16.30***	38.26***	1.18
Forward Caste (%)	0.183	0.181	0.186	0.247	0.244	0.251	-282.72***	0.67	1.70
Muslim (%)	0.154	0.147	0.168	0.155	0.151	0.160	-0.07	14.45***	4.36**
Public School (%)	0.452	0.412	0.533	0.465	0.439	0.496	-7.96***	262.35***	93.60***
Father Education (yrs)	5.386	5.087	5.990	5.650	5.505	5.825	6.12***	-12.78***	-5.79***
Mother Education (yrs)	2.890	2.639	3.396	3.396	3.284	3.531	13.25***	-12.04***	-4.84***
Elderly (%)	0.208	0.214	0.198	0.248	0.246	0.249	-101.61***	6.19**	0.33
Children below 15	2.141	1.988	2.451	1.745	1.588	1.935	-27.04***	-19.07***	-19.18***

Notes: This reports the mean values for the variables in the dataset. The tests of equality use chi-squared tests for comparing variable represented as proportions (indicator variables, reported as %), and the two sample Welch's t-test is used for other variables.

Table 7.2: Schooling gap for different child age-groups

Age Group (years)	Total		Rural		Urban		Tests of Equality		
	Sons	Daughters	Sons	Daughters	Sons	Daughters	$\theta_1 - \theta_2$	$\theta_3 - \theta_4$	$\theta_5 - \theta_6$
	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	(t-stat)	(t-stat)	(t-stat)
Panel A: IHDS-I Survey, 2004-05									
10 to 12	0.93	0.92	1.05	1.03	0.64	0.66	0.16	0.34	-0.29
13 to 15	1.67	1.43	1.89	1.82	1.20	0.79	3.34***	0.78	3.89***
16 to 18	3.26	2.60	3.65	3.44	2.44	1.58	5.40***	1.25	5.23***
19 to 21	5.47	4.85	6.06	6.44	4.33	3.17	2.68***	-1.18	4.07***
Panel B: IHDS-II Survey, 2011-12									
10 to 12	0.52	0.48	0.54	0.54	0.45	0.34	0.80	-0.03	1.62
13 to 15	0.90	0.90	1.00	1.06	0.65	0.52	0.01	-0.92	1.56
16 to 18	2.22	2.12	2.41	2.43	1.82	1.48	1.40	-0.28	2.88***
19 to 21	4.38	3.61	4.87	4.29	3.44	2.52	7.52***	4.38***	6.23***

Table 7.3: Father's education and child's schooling gap

Father's Education	Total		Rural		Urban		Tests of Equality		
	Sons	Daughters	Sons	Daughters	Sons	Daughters	$\theta_1 - \theta_2$	$\theta_3 - \theta_4$	$\theta_5 - \theta_6$
	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	(*t*-stat)	(*t*-stat)	(*t*-stat)
Panel A: IHDS-I Survey, 2004-05									
Non-literate	3.99	2.96	4.02	3.00	3.85	2.82	10.37***	9.21***	4.64***
1 to 5	2.67	1.87	2.73	1.97	2.52	1.61	7.68***	5.97***	4.89***
6 to 10	1.58	0.96	1.66	1.08	1.47	0.79	10.76***	7.47***	7.94***
11 to 15	0.83	0.43	1.07	0.75	0.65	0.21	5.36***	2.50**	4.92***
Panel B: IHDS-II Survey, 2011-12									
Non-literate	3.24	2.97	3.24	2.97	3.23	2.98	3.32***	3.04***	1.35
1 to 5	2.28	1.86	2.25	1.82	2.34	1.96	5.31***	4.71***	2.52**
6 to 10	1.23	0.87	1.21	0.96	1.27	0.71	8.04***	4.32***	7.69***
11 to 15	0.51	0.29	0.63	0.55	0.40	0.06	4.04***	0.91	5.35***

Table 7.4: Mother's education and child's schooling gap

Mother's Education	Total		Rural		Urban		Tests of Equality		
	Sons	Daughters	Sons	Daughters	Sons	Daughters	$\theta_1 - \theta_2$	$\theta_3 - \theta_4$	$\theta_5 - \theta_6$
	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	(t-stat)	(t-stat)	(t-stat)
Panel A: IHDS-I Survey, 2004-05									
Non-literate	3.34	2.96	3.43	3.00	3.04	2.82	10.37***	9.21***	4.64***
1 to 5	1.87	1.87	1.76	1.97	2.07	1.61	7.68***	5.97***	4.89***
6 to 10	0.98	0.96	0.98	1.08	0.99	0.79	10.76***	7.47***	7.94***
11 to 15	0.33	0.43	0.36	0.75	0.32	0.21	5.36***	2.50**	4.92***
Panel B: IHDS-II Survey, 2011-12									
Non-literate	2.83	2.97	2.80	2.97	2.94	2.98	3.32***	3.04***	1.35
1 to 5	1.41	1.86	1.30	1.82	1.65	1.96	5.31***	4.71***	2.52**
6 to 10	0.77	0.87	0.71	0.96	0.83	0.71	8.04***	4.32***	7.69***
11 to 15	0.13	0.29	0.02	0.55	0.17	0.06	4.04***	0.91	5.35***

Table 7.5: Schooling gap for different levels of household income

Household income (quintile)	Total		Rural		Urban		Tests of Equality		
	Sons	Daughters	Sons	Daughters	Sons	Daughters	$\theta_1 - \theta_2$	$\theta_3 - \theta_4$	$\theta_5 - \theta_6$
	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	(*t*-stat)	(*t*-stat)	(*t*-stat)
Panel A: IHDS-I Survey, 2004-05									
1	2.62	2.07	2.70	2.16	1.97	1.54	5.11***	4.64***	1.67*
2	2.86	1.98	2.97	2.02	2.42	1.87	8.45***	8.12***	2.48**
3	2.95	1.79	3.15	1.98	2.50	1.44	11.02***	8.78***	6.23***
4	2.60	1.39	2.97	1.74	2.08	1.04	11.86***	8.17***	7.93***
5	1.79	0.84	2.19	1.23	1.47	0.56	11.02***	6.75***	8.72***
Panel B: IHDS-II Survey, 2011-12									
1	1.93	1.85	1.99	1.92	1.46	1.27	0.99	0.82	0.96
2	1.98	1.68	2.06	1.78	1.73	1.32	3.69***	2.88***	2.67***
3	2.13	1.64	2.23	1.73	1.91	1.45	5.96***	4.88***	3.32***
4	2.19	1.67	2.43	1.93	1.84	1.31	6.23***	4.33***	4.41***
5	1.70	1.12	2.06	1.52	1.32	0.74	7.79***	4.70***	6.47***

Table 7.6: OLS Regression results of child's schooling gap on parental education

	Dependent variable: Child's schooling gap							
	IHDS-I				IHDS-II, 20011-12			
	Sons	Daughters	Sons	Daughters	Sons	Daughters	Sons	Daughters
	Specification 1		*Specification 2*		*Specification 1*		*Specification 2*	
Father Education	-0.214***	-0.170***	-	-	-0.180***	-0.184***	-	-
	(0.005)	(0.007)			(0.004)	(0.005)		
Mother Education	-	-	-0.202***	-0.187***	-	-	-0.177***	-0.191***
			(0.006)	(0.008)			(0.005)	(0.006)
Rural	0.422***	0.720***	0.382***	0.571***	0.348***	0.489***	0.237***	0.355***
	(0.052)	(0.066)	(0.054)	(0.067)	(0.044)	(0.049)	(0.046)	(0.050)
Household Income	-0.109***	-0.165***	-0.174***	-0.211***	-0.085***	-0.110***	-0.115***	-0.132***
	(0.018)	(0.023)	(0.018)	(0.023)	(0.015)	(0.016)	(0.015)	(0.016)
Forward Caste	-0.351***	-0.232***	-0.473***	-0.217***	-0.395***	-0.461***	-0.358***	-0.415***
	(0.061)	(0.079)	(0.062)	(0.079)	(0.047)	(0.052)	(0.048)	(0.052)
Muslim	0.794***	0.819***	0.831***	0.768***	1.040***	0.982***	1.149***	1.014***
	(0.065)	(0.081)	(0.066)	(0.081)	(0.056)	(0.061)	(0.056)	(0.061)
Public School	-1.342***	-1.299***	-1.379***	-1.354***	-1.319***	-1.364***	-1.307***	-1.382***
	(0.051)	(0.063)	(0.052)	(0.063)	(0.044)	(0.048)	(0.044)	(0.048)
Elderly	0.009	0.020	(0.023)	0.021	(0.035)	(0.043)	(0.048)	(0.026)
	(0.054)	(0.073)	(0.055)	(0.073)	(0.044)	(0.049)	(0.045)	(0.049)
Children Below 15	0.021	0.024	0.019	0.032*	0.005	-0.042***	0.015	-0.029**
	(0.014)	(0.018)	(0.014)	(0.018)	(0.013)	(0.013)	(0.013)	(0.013)
Father Age	-0.070***	(0.050)			-0.075***	-0.108***		
	(0.025)	(0.034)			(0.023)	(0.026)		
Father Age^2	0.001***	0.001			0.001***	0.001***		
	(0.000)	(0.000)			(0.000)	(0.000)		

Table 7.6 (continued).

Mother Age			-0.115^{***}	-0.066^{*}			-0.097^{***}	-0.095^{***}
			(0.028)	(0.038)			(0.025)	(0.028)
Mother Age^2			0.002^{***}	0.001^{*}			0.001^{***}	0.002^{***}
			(0.000)	(0.001)			(0.000)	(0.000)
Child Age	-0.704^{***}	-0.534^{***}	-0.708^{***}	-0.554^{***}	-0.790^{***}	-0.400^{***}	-0.795^{***}	-0.430^{***}
	(0.067)	(0.094)	(0.069)	(0.094)	(0.059)	(0.067)	(0.060)	(0.067)
Child Age^2	0.037^{***}	0.030^{***}	0.036^{***}	0.030^{***}	0.037^{***}	0.022^{***}	0.037^{***}	0.022^{***}
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	7.612^{***}	5.943^{***}	8.002^{***}	6.031^{***}	7.904^{***}	6.289^{***}	7.873^{***}	5.598^{***}
	(0.721)	(0.943)	(0.710)	(0.930)	(0.649)	(0.726)	(0.631)	(0.702)
Observations	1,454	6667	13454	6667	15958	13189	15958	13189
R^2	0.42	0.34	0.403	0.34	0.397	0.341	0.391	0.34
F Statistic	812.630^{***}	285.106^{***}	897.128^{***}	566.484^{***}	736.001^{***}	286.856^{***}	854.213^{***}	564.728^{***}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 7.7: Interaction inclusive models (OLS Estimates)

	Dependent variable: Child's schooling gap							
	IHDS-I				IHDS-II, 20011-12			
	Sons	Daughters	Sons	Daughters	Sons	Daughters	Sons	Daughters
Father Education	-0.147***	-0.148***	-	-	-0.131***	-0.161***	-	-
	(0.017)	(0.021)			(0.014)	(0.015)		
Mother Education	-	-	-0.189***	-0.233***	-	-	-0.160***	-0.184***
			(0.021)	(0.024)			(0.015)	(0.017)
Rural	-0.032	-0.145***	-0.175***	-0.266***	-0.018	-0.083***	-0.105***	-0.142***
	(0.026)	(0.037)	(0.021)	(0.028)	(0.023)	(0.026)	(0.019)	(0.021)
Household Income	0.546***	0.753***	0.520***	0.661***	0.409***	0.505***	0.305***	0.497***
	(0.084)	(0.112)	(0.068)	(0.088)	(0.074)	(0.085)	(0.061)	(0.068)
Forward Caste	-0.656***	-0.601***	-0.733***	-0.389***	-0.644***	-0.706***	-0.617***	-0.769***
	(0.109)	(0.157)	(0.082)	(0.115)	(0.081)	(0.093)	(0.066)	(0.074)
Muslim	1.159***	1.283***	1.031***	0.968***	1.492***	1.474***	1.471***	1.493***
	(0.090)	(0.118)	(0.076)	(0.095)	(0.081)	(0.091)	(0.069)	(0.076)
Public School	-1.335***	-1.282***	-1.355***	-1.326***	-1.304***	-1.349***	-1.290***	-1.339***
	(0.051)	(0.064)	(0.052)	(0.063)	(0.044)	(0.048)	(0.044)	(0.048)
Elderly	0.003	0.023	-0.025	0.021	-0.033	-0.047	-0.047	-0.031
	(0.054)	(0.072)	(0.055)	(0.072)	(0.044)	(0.049)	(0.045)	(0.049)

Children Below 15	0.021	0.024	0.019	0.030*	0.006	-0.042***	0.015	-0.030**
	(0.014)	(0.018)	(0.014)	(0.018)	(0.013)	(0.013)	(0.013)	(0.013)
Father Age	-0.072***	-0.046	-	-	-0.073***	-0.107***	-	-
	(0.025)	(0.034)			(0.023)	(0.026)		
Father Age^2	0.001***	0.0004			0.001***	0.001***	-	-
	(0.000)	(0.000)			(0.000)	(0.000)		
Mother Age	-	-	-0.122***	-0.076**	-	-	-0.102***	-0.105***
			(0.028)	(0.038)			(0.025)	(0.028)
Mother Age^2	-	-	0.002***	0.001**	-	-	0.001***	0.002***
			(0.000)	(0.001)			(0.000)	(0.000)
Child Age	-0.701***	-0.545***	-0.700***	-0.548***	-0.790***	-0.400***	-0.794***	-0.429***
	(0.067)	(0.094)	(0.068)	(0.094)	(0.059)	(0.067)	(0.060)	(0.067)
Child Age^2	0.037***	0.031***	0.036***	0.030***	0.037***	0.022***	0.037***	0.022***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Father Education × Household Income	-0.015***	-0.003	-	-	-0.012***	-0.005	-	-
	(0.004)	(0.005)			(0.003)	(0.003)		
Father Education × Rural	-0.021*	-0.002	-	-	-0.009	-0.001	-	-
	(0.011)	(0.014)			(0.009)	(0.010)		
Father Education × Forward Caste	0.040***	0.040**	-	-	0.033***	0.032***	-	-
	(0.013)	(0.017)			(0.010)	(0.011)		
Father Education × Muslim	-0.086***	-0.095***	-	-	-0.094***	-0.096***	-	-

	(0.015)	(0.017)			(0.012)	(0.013)		
Mother Education × Household Income	-	-	0.001	0.017***	-	-	-0.004	0.002
			(0.005)	(0.005)			(0.003)	(0.004)
Mother Education × Rural	-	-	-0.042***	-0.021	-	-	-0.017*	-0.030***
			(0.013)	(0.015)			(0.010)	(0.011)
Mother Education × Forward Caste	-	-	0.059***	0.026	-	-	0.052***	0.068***
			(0.014)	(0.017)			(0.010)	(0.011)
Mother Education × Muslim	-	-	-0.095***	-0.082***	-	-	-0.101***	-0.143***
			(0.019)	(0.021)			(0.014)	(0.015)
Constant	7.296***	5.798***	7.987***	6.290***	7.563***	6.142***	7.878***	5.709***
	(0.723)	(0.947)	(0.711)	(0.930)	(0.652)	(0.730)	(0.633)	(0.702)
Observations	13,454	6,667	13,454	6,667	15,958	13,189	15,958	13,189
R^2	0.423	0.344	0.4	0.345	0.406	0.343	0.394	0.346
F Statistic	616.263***	217.883***	559.099***	218.929***	680.744***	430.201***	647.915***	435.087***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7.8: High parental education and child's schooling gap (High education defined as primary or above)

Parent	Child	Estimate		Gamma (Γ)					
				1.0	1.2	1.4	1.6	1.8	2.0
Panel A: IHDS-I 2004-05									
Father	Son	-1.684***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-2.0, -0.9]	[-2.5, -0.4]	[-2.5, -0.4]	[-2.5, 0.1]
		(0.086)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.180***	HL estimate bounds	[-1.0, -1.0]	[-1.1, -0.4]	[-1.5, -0.4]	[-1.5, -0.4]	[-2.0, 0.1]	[-2.0, 0.1]
		(0.105)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-1.436***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-2.0, -0.9]	[-2.0, -0.4]	[-2.5, -0.4]	[-2.5, -0.4]
		(0.061)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.308***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-1.6, -0.9]	[-2.0, -0.4]	[-2.0, -0.4]	[-2.0, -0.4]
		(0.077)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Panel B: IHDS-II 2011-12									
Father	Son	-1.286***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -1.0]	[-1.6, -0.5]	[-2.1, -0.5]	[-2.1, 0.0]	[-2.6, 0.0]
		(0.064)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.217***	HL estimate bounds	[-1.0, -1.0]	[-1.1, -0.4]	[-1.5, -0.4]	[-1.5, -0.4]	[-2.0, 0.1]	[-2.0, 0.1]
		(0.076)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-1.331***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-1.6, -0.4]	[-2.0, -0.4]	[-2.0, -0.4]	[-2.0, -0.4]
		(0.057)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.336***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-1.6, -0.4]	[-2.0, -0.4]	[-2.0, -0.4]	[-2.0, -0.4]
		(0.060)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)

Notes: The Abadie-Imbens standard errors are reported below the estimates in parentheses. Gamma (Γ) is the odds ratio of receiving treatment due to unobserved factors such as genetic endowments. "HL estimate bounds" are the Hodges-Lehmann bounds for the ATT estimate. "WSR p-value bounds" are the lower and upper bounds for the p-values from Wilcoxon's signed rank test. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.9: High parental education and child's schooling gap (High education defined as middle or above)

Parent	Child	Estimate		Gamma (Γ) 1.0	1.2	1.4	1.6	1.8	2.0
Panel A: IHDS-I 2004-05									
Father	Son	-1.642***	HL estimate bounds	[-2.0, -2.0]	[-2.1, -1.0]	[-2.1, -1.0]	[-2.6, -0.5]	[-2.6, -0.5]	[-2.6, -0.5]
		(0.069)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.002***	HL estimate bounds	[-1.0, -1.0]	[-1.1, -0.4]	[-1.5, -0.4]	[-1.5, 0.1]	[-1.5, 0.1]	[-2.0, 0.1]
		(0.089)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.87)
Mother	Son	-1.428***	HL estimate bounds	[-2.0, -2.0]	[-2.1, -1.0]	[-2.1, -1.0]	[-2.1, -1.0]	[-2.6, -0.5]	[-2.6, -0.5]
		(0.067)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.027***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.5]	[-1.6, -0.5]	[-1.6, -0.5]	[-1.6, -0.5]	[-2.1, 0.0]
		(0.072)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Panel B: IHDS-II 2011-12									
Father	Son	-1.321***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -1.0]	[-1.6, -0.5]	[-2.1, -0.5]	[-2.1, -0.5]	[-2.1, 0.0]
		(0.054)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.170***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.5]	[-1.6, -0.5]	[-1.6, -0.5]	[-2.1, 0.0]	[-2.1, 0.0]
		(0.059)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-1.345***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-1.6, -0.9]	[-2.0, -0.4]	[-2.0, -0.4]	[-2.0, -0.4]
		(0.052)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.196***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-1.6, -0.9]	[-2.0, -0.4]	[-2.0, -0.4]	[-2.0, -0.4]
		(0.055)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)

Notes: The Abadie-Imbens standard errors are reported below the estimates in parentheses. Gamma (Γ) is the odds ratio of receiving treatment due to unobserved factors such as genetic endowments. "HL estimate bounds" are the Hodges-Lehmann bounds for the ATT estimate. "WSR p-value bounds" are the lower and upper bounds for the p-values from Wilcoxon's signed rank test. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.10: High parental education and child's schooling gap (High education defined as secondary or above)

Parent	Child	Estimate		Gamma (Γ)					
				1.0	1.2	1.4	1.6	1.8	2.0
Panel A: IHDS-I 2004-05									
Father	Son	-1.391***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-1.6, -0.4]	[-2.0, -0.4]	[-2.0, -0.4]	[-2.5, 0.1]
		(0.075)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-0.588***	HL estimate bounds	[-1.0, -1.0]	[-1.1, -0.5]	[-1.1, 0.0]	[-1.1, 0.0]	[-1.6, 0.0]	[-1.6, 0.0]
		(0.088)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.76)	(0.00, 1.00)
Mother	Son	-1.433***	HL estimate bounds	[-2.0, -2.0]	[-2.1, -1.4]	[-2.1, -0.9]	[-2.5, -0.9]	[-2.5, -0.9]	[-2.5, -0.9]
		(0.101)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-0.844***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -0.9]	[-1.6, -0.9]	[-2.0, -0.9]	[-2.0, -0.4]	[-2.0, -0.4]
		(0.106)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Panel B: IHDS-II 2011-12									
Father	Son	-1.213***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -1.0]	[-1.6, -0.5]	[-2.1, -0.5]	[-2.1, -0.5]	[-2.1, 0.0]
		(0.055)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-0.995***	HL estimate bounds	[-1.0, -1.0]	[-1.1, -0.4]	[-1.1, -0.4]	[-1.5, -0.4]	[-1.5, 0.1]	[-1.5, 0.1]
		(0.062)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-1.377***	HL estimate bounds	[-2.0, -2.0]	[-2.1, -1.0]	[-2.1, -1.0]	[-2.1, -1.0]	[-2.1, -0.5]	[-2.6, -0.5]
		(0.068)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.060***	HL estimate bounds	[-1.5, -1.5]	[-1.6, -1.0]	[-1.6, -0.5]	[-1.6, -0.5]	[-1.6, -0.5]	[-2.1, -0.5]
		(0.067)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)

Notes: The Abadie-Imbens standard errors are reported below the estimates in parentheses. Gamma (Γ) is the odds ratio of receiving treatment due to unobserved factors such as genetic endowments. "HL estimate bounds" are the Hodges-Lehmann bounds for the ATT estimate. "WSR p-value bounds" are the lower and upper bounds for the p-values from Wilcoxon's signed rank test. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.11: High parental education and child's schooling gap (High education defined as primary or above)

Parent	Child	Estimate		Gamma (Γ)					
				1.0	1.2	1.4	1.6	1.8	2.0
Panel A: IHDS-I 2004-05									
Father	Son	-2.023***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.5, -1.0]	[-2.5, -1.0]	[-3.0, -1.0]	[-3.0, -0.5]
		(0.050)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.773***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.0, -1.0]	[-2.5, -1.0]	[-2.5, -0.5]	[-2.5, -0.5]
		(0.062)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-1.944***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.0, -1.0]	[-2.5, -1.0]	[-2.5, -1.0]	[-3.0, -0.5]
		(0.062)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.892***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.0, -1.0]	[-2.5, -1.0]	[-2.5, -1.0]	[-2.5, -0.5]
		(0.069)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Panel B: IHDS-II 2011-12									
Father	Son	-1.471***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.5, 0.0]
		(0.041)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.621***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.0, -0.5]	[-2.5, -0.5]	[-2.5, -0.5]
		(0.044)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-1.598***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-2.0, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.5, -0.5]
		(0.047)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.791***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.5, -0.5]	[-2.5, -0.5]
		(0.049)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)

Notes: The Abadie-Imbens standard errors are reported below the estimates in parentheses. Gamma (Γ) is the odds ratio of receiving treatment due to unobserved factors such as genetic endowments. "HL estimate bounds" are the Hodges-Lehmann bounds for the ATT estimate. "WSR p-value bounds" are the lower and upper bounds for the p-values from Wilcoxon's signed rank test. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.12: High parental education and child's schooling gap (High education defined as middle or above)

Parent	Child	Estimate		Gamma (Γ)					
				1.0	1.2	1.4	1.6	1.8	2.0
Panel A: IHDS-I 2004-05									
Father	Son	-2.070***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.5, -1.5]	[-2.5, -1.0]	[-3.0, -1.0]	[-3.0, -0.5]
		(0.059)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.537***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-2.0, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.5, -0.5]
		(0.070)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-2.031***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.5, -1.5]	[-2.5, -1.0]	[-2.5, -1.0]	[-3.0, -1.0]
		(0.080)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.718***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.0, -0.5]	[-2.5, -0.5]
		(0.091)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Panel B: IHDS-II 2011-12									
Father	Son	-1.587***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-2.0, -1.0]	[-2.0, -0.5]	[-2.5, -0.5]	[-2.5, -0.5]
		(0.047)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.521***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-1.5, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.5, -0.5]
		(0.050)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-1.640***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-2.0, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.5, -0.5]
		(0.059)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.652***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-1.5, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.0, -0.5]
		(0.059)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)

Notes: The Abadie-Imbens standard errors are reported below the estimates in parentheses. Gamma (Γ) is the odds ratio of receiving treatment due to unobserved factors such as genetic endowments. "HL estimate bounds" are the Hodges-Lehmann bounds for the ATT estimate. "WSR p-value bounds" are the lower and upper bounds for the p-values from Wilcoxon's signed rank test. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.13: High parental education and child's schooling gap (High education defined as secondary or above)

Parent	Child	Estimate		Gamma (Γ)					
				1.0	1.2	1.4	1.6	1.8	2.0
Panel A: IHDS-I 2004-05									
Father	Son	-2.021***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.0, -1.0]	[-2.5, -1.0]	[-2.5, -1.0]	[-3.0, -0.5]
		(0.110)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.523***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-1.5, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.5, -0.5]
		(0.118)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-2.305***	HL estimate bounds	[-2.5, -2.5]	[-2.5, -1.5]	[-2.5, -1.5]	[-2.5, -1.5]	[-3.0, -1.0]	[-3.0, -1.0]
		(0.168)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.912***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.5]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.5, -1.0]	[-2.5, -0.5]
		(0.180)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Panel B: IHDS-II 2011-12									
Father	Son	-1.581***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-1.5, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.5, -0.5]
		(0.079)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.554***	HL estimate bounds	[-1.5, -1.5]	[-1.5, -1.0]	[-1.5, -1.0]	[-2.0, -0.5]	[-2.0, -0.5]	[-2.0, -0.5]
		(0.085)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Mother	Son	-1.846***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.5, -1.0]	[-2.5, -0.5]
		(0.115)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	Daughter	-1.833***	HL estimate bounds	[-2.0, -2.0]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.0, -1.0]	[-2.0, -0.5]	[-2.5, -0.5]
		(0.121)	WSR p-value bounds	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)

Notes: The Abadie-Imbens standard errors are reported below the estimates in parentheses. Gamma (Γ) is the odds ratio of receiving treatment due to unobserved factors such as genetic endowments. "HL estimate bounds" are the Hodges-Lehmann bounds for the ATT estimate. "WSR p-value bounds" are the lower and upper bounds for the p-values from Wilcoxon's signed rank test. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.14: Descriptive statistics

	Sons					Daughters				
	No. of obs.	Mean	Std. Dev.	Min	Max	No. of obs.	Mean	Std. Dev.	Min	Max
Dependent Variable										
Schooling Gap (Years)	9,031	5.093	2.095	0	8	6,966	5.447	1.922	0	8
Treatment Variables										
Mid-Day Meal	9,031	0.427	0.495	0	1	6,966	0.472	0.499	0	1
Free Books	9,031	0.504	0.500	0	1	6,966	0.583	0.493	0	1
Free Uniform	9,031	0.130	0.337	0	1	6,966	0.200	0.400	0	1
School Fees	9,031	0.159	0.366	0	1	6,966	0.172	0.377	0	1
Scholarship	9,031	0.115	0.319	0	1	6,966	0.149	0.356	0	1
Covariates used for matching										
Child Age (Years)	9,031	10.360	2.754	6	15	6,966	10.002	2.676	6	15
Father Age (Years)	9,031	40.519	6.788	22	82	6,966	40.245	6.884	21	88
Mother Age (Years)	9,031	35.326	6.130	18	68	6,966	34.923	6.129	18	73
Father Education (Years)	9,031	5.560	4.623	0	15	6,966	6.165	4.654	0	15
Mother Education (Years)	9,031	3.138	4.155	0	15	6,966	3.597	4.354	0	15
Father with Nonfarm Occupation	9,031	0.676	0.468	0	1	6,966	0.689	0.463	0	1
Mother with Nonfarm Occupation	9,031	0.748	0.434	0	1	6,966	0.759	0.428	0	1
Rural	9,031	0.686	0.464	0	1	6,966	0.664	0.472	0	1
Forward Caste	9,031	0.189	0.392	0	1	6,966	0.193	0.395	0	1
Muslim	9,031	0.134	0.341	0	1	6,966	0.139	0.346	0	1
Public School	9,031	0.707	0.455	0	1	6,966	0.730	0.444	0	1
Grade Repeated	9,031	0.074	0.262	0	1	6,966	0.060	0.237	0	1
Household Income	9,031	2.791	1.359	1	5	6,966	3.068	1.401	1	5

Notes: This table reports the mean values and other summary statistics for the variables used in the matching analysis.

Table 7.15: Effect of encouragement on schooling gap of Sons using Mahalanobis Matching

Mode of Encouragement	Scholarship	MidDayMeal	FreeBooks	FreeUniform	SchoolFees
	Average treatment effect of treated (ATT Estimates)				
Panel A: Child Age at the time of IHDS-I (in years)					
All Sons	-0.066	0.038	-0.045	0.114	-0.052
Sons aged 6 - 9	-0.320**	0.031	-0.056	0.276**	0.047
Sons aged 10 - 12	0.034	-0.110	-0.153	-0.056	-0.062
Sons aged 13 - 15	0.261	-0.163	0.031	0.058	-0.085
Panel B: Father's education (years of schooling)					
Non-literate	0.142	0.118	0.202	0.228	0.026
1-5 years	-0.045	0.177	-0.036	0.216	-0.005
6-10 years	-0.160	0.169	-0.106	0.216	-0.063
11 years & above	0.027	0.268	-0.056	0.247	0.264
Panel C: Mother's education (years of schooling)					
Non-literate	-0.048	0.109	0.035	0.264**	-0.035
1-5 years	-0.432*	0.384**	0.048	-0.148	-0.613***
6-10 years	-0.185	0.155	0.143	0.142	-0.090
11 years & above	0.143	0.031	-0.733***	-0.167	-0.295

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.16: Effect of encouragement on schooling gap of Daughters using Mahalanobis Matching

Mode of Encouragement	Scholarship	MidDayMeal	FreeBooks	FreeUniform	SchoolFees
	Average treatment effect of treated (ATT Estimates)				
Panel A: Child Age at the time of IHDS-I (in years)					
All Daughters	-0.397***	0.112	-0.080	0.118	-0.056
Daughters aged 6 - 9	-0.385***	0.068	-0.073	0.125	-0.097
Daughters aged 10 - 12	-0.436***	0.006	-0.249*	-0.097	0.040
Daughters aged 13 - 15	-0.407*	-0.098	-0.017	-0.252	-0.072
Panel B: Father's education (years of schooling)					
Non-literate	-0.375**	0.492***	0.127	0.381**	0.028
1-5 years	-0.518**	0.279	-0.215	-0.129	-0.246
6-10 years	-0.382**	0.104	-0.019	-0.046	0.115
11 years & above	-0.516*	0.310	-0.121	0.179	-0.248
Panel C: Mother's education (years of schooling)					
Non-literate	-0.297**	0.305**	-0.167	0.084	0.055
1-5 years	0.014	-0.056	0.391*	0.307	0.067
6-10 years	-0.331	0.154	0.033	0.213	-0.136
11 years & above	-0.615	0.016	0.293	-0.786*	-0.395

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.17: Effect of encouragement on schooling gap of Sons using Kernel Matching

Mode of Encouragement	Scholarship	MidDayMeal	FreeBooks	FreeUniform	SchoolFees
	Average treatment effect of treated (ATT Estimates)				
Panel A: Child Age at the time of IHDS-I (in years)					
All Sons	-0.179**	0.064	0.002	0.199***	0.030
Sons aged 6 - 9	-0.358***	0.113	-0.039	0.261***	0.042
Sons aged 10 - 12	-0.114	-0.052	-0.132	-0.015	-0.017
Sons aged 13 - 15	-0.023	0.159	0.033	0.147	-0.001
Panel B: Father's education (years of schooling)					
Non-literate	-0.093	0.129	0.254**	0.223*	0.057
1-5 years	-0.157	0.182	-0.051	0.323**	0.052
6-10 years	-0.221*	0.053	-0.092	0.255**	-0.025
11 years & above	-0.109	0.161	-0.042	0.323*	0.218
Panel C: Mother's education (years of schooling)					
Non-literate	-0.126	0.069	0.071	0.286***	0.078
1-5 years	-0.143	0.252	0.009	0.060	-0.221
6-10 years	0.005	0.098	0.104	0.123	0.092
11 years & above	-0.082	0.443*	-0.479**	0.285	-0.002

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.18: Effect of encouragement on schooling gap of Daughters using Kernel Matching

Mode of Encouragement	Scholarship	MidDayMeal	FreeBooks	FreeUniform	SchoolFees
	Average treatment effect of treated (ATT Estimates)				
Panel A: Child Age at the time of IHDS-I (in years)					
All Daughters	-0.379***	0.100	-0.018	0.150**	-0.020
Daughters aged 6 - 9	-0.349***	0.119	0.003	0.344***	0.042
Daughters aged 10 - 12	-0.378***	0.109	-0.182*	-0.061	-0.108
Daughters aged 13 - 15	-0.364**	-0.280	-0.226*	-0.388*	-0.127
Panel B: Father's education (years of schooling)					
Non-literate	-0.173	0.321*	0.096	0.332***	0.130
1-5 years	-0.252	0.094	-0.193	0.136	-0.152
6-10 years	-0.431***	-0.042	0.041	0.066	0.047
11 years & above	-0.540***	0.564***	0.021	0.243	-0.258
Panel C: Mother's education (years of schooling)					
Non-literate	-0.256***	0.264**	-0.109	0.195**	0.040
1-5 years	-0.029	-0.047	0.409**	0.250*	-0.069
6-10 years	-0.586***	0.141	0.045	0.113	-0.099
11 years & above	-0.496	0.483**	0.369	-0.383	-0.406*

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7.19: Match quality for Sons

		Pseudo-R^2	LR test (p-value)	Mean Bias	Median Bias	Rubin's R	Rubin's B
		Treatment Variable: Scholarship					
	Unmatched	0.08	0.00	23.25	18.89	0.52	78.11
Mahalnobis Matching	Matched	0.00	1.00	0.71	0.22	1.06	3.17
PSM-Kernel Matching	Matched	0.00	0.93	2.33	2.33	0.92	9.85
		Treatment Variable: Mid-day Meal					
	Unmatched	0.34	0.00	39.14	25.59	0.45	160.62
Mahalnobis Matching	Matched	0.01	0.00	2.21	0.23	1.00	17.01
PSM-Kernel Matching	Matched	0.00	0.00	4.72	2.65	0.87	15.64
		Treatment Variable: Free Books					
	Unmatched	0.29	0.00	34.16	23.39	0.28	145.96
Mahalnobis Matching	Matched	0.00	0.00	1.27	0.20	1.12	10.96
PSM-Kernel Matching	Matched	0.00	0.80	1.34	1.25	1.05	5.52
		Treatment Variable: Free Uniform					
	Unmatched	0.13	0.00	21.30	15.11	0.31	100.68
Mahalnobis Matching	Matched	0.00	1.00	0.47	0.00	1.16	5.19
PSM-Kernel Matching	Matched	0.00	0.99	0.97	0.81	0.85	6.83
		Treatment Variable: School Fees					
	Unmatched	0.06	0.00	13.30	8.02	0.58	63.86
Mahalnobis Matching	Matched	0.00	1.00	0.38	0.19	1.18	4.89
PSM-Kernel Matching	Matched	0.00	0.91	1.26	0.82	0.99	8.67

Table 7.20: Match quality for Daughters

		Pseudo-R^2	LR test (p-value)	Mean Bias	Median Bias	Rubin's R	Rubin's B
		Treatment Variable: Scholarship					
	Unmatched	0.10	0.00	24.86	18.87	0.42	85.26
Mahalnobis Matching	Matched	0.00	1.00	0.69	0.00	1.12	4.62
PSM-Kernel Matching	Matched	0.00	0.99	1.39	1.20	0.95	7.55
		Treatment Variable: Mid-day Meal					
	Unmatched	0.37	0.00	43.32	32.30	0.39	167.09
Mahalnobis Matching	Matched	0.00	0.00	1.53	0.27	0.97	13.45
PSM-Kernel Matching	Matched	0.00	0.00	4.78	4.80	0.83	16.03
		Treatment Variable: Free Books					
	Unmatched	0.35	0.00	41.39	32.91	0.28	165.61
Mahalnobis Matching	Matched	0.00	0.04	1.23	0.54	1.18	10.04
PSM-Kernel Matching	Matched	0.00	0.49	1.96	1.64	1.25	7.19
		Treatment Variable: Free Uniform					
	Unmatched	0.15	0.00	26.96	21.79	0.17	104.92
Mahalnobis Matching	Matched	0.00	1.00	0.56	0.17	1.18	4.69
PSM-Kernel Matching	Matched	0.00	0.99	0.87	0.82	1.06	6.21
		Treatment Variable: School Fees					
	Unmatched	0.07	0.00	15.96	9.16	0.52	68.80
Mahalnobis Matching	Matched	0.00	0.99	1.31	0.19	1.25	7.38
PSM-Kernel Matching	Matched	0.00	0.90	1.60	1.55	0.96	9.62

Table 7.21: Effect of encouragement on schooling gap of Sons (Rosenbaum sensitivity analysis)

	Gamma (Γ)					
	1	1.2	1.4	1.6	1.8	2
Panel A: Scholarship						
All Sons	**0.040**	**0.066**	**0.082**	0.122	0.147	0.295
Sons aged 6 - 9	**0.002**	**0.003**	**0.004**	**0.005**	**0.009**	**0.014**
Sons aged 10 – 12	0.500	0.600	0.708	0.970	1.000	1.000
Sons aged 13 – 15	0.270	0.486	0.627	1.000	1.000	1.000
Panel B: MidDayMeal						
All Sons	0.250	0.425	0.442	0.751	1.000	1.000
Sons aged 6 - 9	0.265	0.472	0.807	1.000	1.000	1.000
Sons aged 10 – 12	0.230	0.361	0.708	0.722	1.000	1.000
Sons aged 13 – 15	0.185	0.315	0.450	0.751	1.000	1.000
Panel C: Free Books						
All Sons	0.440	0.585	0.737	1.000	1.000	1.000
Sons aged 6 - 9	0.400	0.536	0.858	1.000	1.000	1.000
Sons aged 10 – 12	0.405	0.620	0.954	1.000	1.000	1.000
Sons aged 13 – 15	0.335	0.412	0.552	0.701	0.771	1.000
Panel D: Free Uniform						
All Sons	**0.000**	**0.001**	**0.001**	**0.001**	**0.002**	**0.003**
Sons aged 6 - 9	**0.008**	**0.015**	**0.029**	**0.042**	**0.053**	**0.100**
Sons aged 10 – 12	0.585	0.761	1.000	1.000	1.000	1.000
Sons aged 13 – 15	0.420	0.433	0.653	1.000	1.000	1.000
Panel E: School Fees						
All Sons	0.555	0.627	1.000	1.000	1.000	1.000
Sons aged 6 - 9	0.395	0.679	1.000	1.000	1.000	1.000
Sons aged 10 – 12	0.435	0.683	1.000	1.000	1.000	1.000
Sons aged 13 – 15	0.440	0.678	0.772	0.842	1.000	1.000

Notes: This table presents the upper bound of the p-value for the ATT estimate obtained using Kernel matching under different levels of unobserved heterogeneity. The level of unobserved heterogeneity is represented by Gamma (Γ). p-values lower than 0.1 are highlighted in bold.

Table 7.22: Effect of encouragement on schooling gap of Daughters (Rosenbaum sensitivity analysis)

	Gamma (Γ)					
	1	1.2	1.4	1.6	1.8	2
Panel A: Scholarship						
All Daughters	**0.001**	**0.001**	**0.002**	**0.003**	**0.004**	**0.004**
Daughters aged 6 - 9	**0.006**	**0.011**	**0.022**	**0.031**	**0.034**	**0.035**
Daughters aged 10 – 12	**0.007**	**0.008**	**0.013**	**0.018**	**0.032**	**0.049**
Daughters aged 13 – 15	**0.023**	**0.043**	**0.062**	0.105	0.137	0.225
Panel B: MidDayMeal						
All Daughters	0.315	0.627	0.821	1.000	1.000	1.000
Daughters aged 6 - 9	0.370	0.670	1.000	1.000	1.000	1.000
Daughters aged 10 – 12	0.135	0.194	0.301	0.464	0.905	1.000
Daughters aged 13 – 15	0.240	0.430	0.632	0.669	0.870	1.000
Panel C: Free Books						
All Daughters	0.480	0.586	0.826	1.000	1.000	1.000
Daughters aged 6 - 9	0.585	0.942	1.000	1.000	1.000	1.000
Daughters aged 10 – 12	**0.057**	**0.085**	0.136	0.260	0.504	0.765
Daughters aged 13 – 15	**0.081**	0.134	0.259	0.392	0.705	1.000
Panel D: Free Uniform						
All Daughters	**0.015**	**0.018**	**0.032**	**0.057**	**0.087**	0.121
Daughters aged 6 - 9	**0.006**	**0.009**	**0.016**	**0.025**	**0.049**	**0.057**
Daughters aged 10 – 12	0.195	0.333	0.647	0.867	1.000	1.000
Daughters aged 13 – 15	**0.078**	**0.086**	0.105	0.118	0.160	0.198
Panel E: School Fees						
All Daughters	0.445	0.538	0.845	1.000	1.000	1.000
Daughters aged 6 - 9	0.575	0.633	0.664	1.000	1.000	1.000
Daughters aged 10 – 12	0.135	0.162	0.173	0.213	0.369	0.701
Daughters aged 13 – 15	0.520	1.000	1.000	1.000	1.000	1.000

Notes: This table presents the upper bound of the p-value for the ATT estimate obtained using Kernel matching under different levels of unobserved heterogeneity. The level of unobserved heterogeneity is represented by Gamma (Γ). p-values lower than 0.1 are highlighted in bold

Chapter 8

The Macroeconomic Dynamics of Intergenerational Educational Mobility: Evidence from State-level Analysis

8.1 Introduction

The inequality in access to education leads to unequal distribution of labour market opportunities for workers (Black & Devereux, 2011). Many recent studies have shown that the degree of IEM varies substantially between countries and (e.g., Causa & Johansson, 2010; Hertz et al., 2007), but only limited studies have attempted to determine the factors that have been associated with these differences. Theoretical literature on IEM suggests that factors like genetics, family characteristics and institutional framework are likely to affect the differences in IEM across countries. But it is assumed that if the effect of genetic transmission and family characteristics remain similar across countries, the differences in degree of IEM could be attributed to institutional differences across countries.

However, there is paucity of studies that have determined the factors affecting IEM within a country. The goal of this chapter is to measure the degree of IEM across Indian states and to analyse the changes over time. This chapter takes advantage of cross-state differences in the degree of IEM and institutional framework to determine the macroeconomic factors that affect the level of IEM.

This chapter has been planned as follows. In Section 8.2, we review literature related to the topic. Section 8.3 explains the data and empirical framework used for the analysis. In Section 8.5, we present the results. Further, Section 8.6 concludes.

8.2 Background: Relevant literature

There exist innumerable studies which have analysed the cross-country differences in degree of IEM (Causa & Johansson, 2011; Hertz et al., 2007; Narayan et al., 2018; Torul & Öztunalı, 2018).

There is plethora of studies that have analysed the causal effect of genetic transmission (nature effect) and parental education (nurture effect) on IEM. Many other past studies have tried to analyse the impact of institutional features on IEM, but most of them have analysed this association between institutional features and IEM, typically by taking different institutions in separation. We briefly review prior literature analysing the impact of different factors on IEM and expectations regarding the correlation between different institutional factors and IEM.

The close association between inequality and IEM has been widely studied. It is been argued that parent's investment patterns shape the income earning potential of their children, generating substantial wage differential between children belonging to affluent and poor families, which eventually leads to widening income disparities (Neckerman & Torche, 2007). One possible explanation to this relationship could be that rich parents are more capable in providing access to increased educational opportunities to their children than poor parents (Burtless & Jencks, 2003). Second, higher inequality may increase the political influence of affluent parties who are likely to provide political parties with huge funding. Therefore, there contribution towards the political parties is likely to narrow down the scope of implementing progressive policies (Durlauf, 1996). However, the exact link between inequality and IEM bank on the complexity of interaction between families, market structure and policy responsiveness of state institutions. For example, if the economy is facing downward mobility due to increasing inequality, efficient redistribution of resources could be a way out to tackle the negative impact of increasing inequality (Solon, 2002).

Solon (2002) shows that intergenerational earning mobility is likely to be affected by the magnitude of private rate of return to education. If there is a high degree of persistence, children of highly (low) educated parents are likely to be highly (low) educated. This could be attributed to the lack of progressive policies acting as a barrier in equal access of educational opportunities or increasing returns to higher education (assuming that affluent families are more capable of affording higher education for their children).

It is also believed that IEM is quite low in less-developed countries, but lack of reliable data in less-developed countries makes it difficult to provide internationally comparable estimates of IEM. Therefore, we will also include other economic development variables like Gross Domestic Product (GDP) per capita and poverty reduction rate.

Some past studies have tried to gauge the impact of government spending on IEM. (Hanushek, 1996, 2001) shows that public spending has a very little effect on the test scores of the children. However, results obtained by some other empirical show that public spending plays a vital role in affecting the degree of IEM (Greenwald *et al.*, 1994; Harknett *et al.*, 2003). Behrman *et al.* (2001) used data from 16 Latin American countries to find that higher state-level expenditure on primary education coupled with teachers who are well-educated works in reducing intergenerational persistence overtime. Mayer & Lopoo (2008) finds that public spending on elementary and secondary education favours children belonging to low-income families, whereas public spending on tertiary education favours children belonging to high-income families.

8.3 Data and Methodology

8.3.1 Data Source

The analysis in this chapter has been split into two parts. In the first part, we discuss the trend and pattern of IEM across major Indian states. This part of the analysis uses information provided in IHDS-II survey. The second part of the analysis tries to analyse the determinants of IEM. To analyse the macroeconomic determinants of IEM requires much more information. Therefore, we use two rounds of the NSS. The information on key policy variables has been collected from different sources (Asadullah & Yalonetzky, 2012;Timothy Besley et al., 2007).

8.3.2 Methodology

8.3.2.1 Trend and Pattern in Mobility: A State-Level Analysis

We study the trend in IEM by measuring different mobility indicators, which have been discussed in detail in Chapter 3. Our data come from the latest round of IHDS-II (2004-05). The analysis focuses on the father-son pairs regardless of whether the father coresides with their child or not. Additional question on the years of completed education by the father of the head of the household allows us to avoid any kind of sample selection bias due to coresidency restriction. Following prior literature on IEM, we restrict our sample size to children aged between 25 to 65 years old. To study the trend in mobility estimates, we divide the children into four consecutive birth cohorts.

8.3.2.2 Macroeconomic Determinants of Intergenerational Educational Mobility

To empirically analyse the macroeconomic determinants of IEM, we utilize information from different rounds of NSS. We use EUS rounds of NSS as they are thick rounds which cover different dimensions of labour market outcomes. The NSS offers information on completed years of education as level of education rather than continuous years of education as in IHDS dataset. This could lead to biased estimates of mobility due to measurement error. Therefore, it becomes inevitable to compare the estimates obtained by using both datasets (NSS-68th round and IHDS-II) to check robustness of our findings.

8.4 Results

8.4.1 Inter-State Variations in Mobility Estimates

Tables 8.1 to 8.3 show estimates for different mobility indicators for major Indian states. The national level measures of IEM do not indicate the inter-regional disparity among the various states, which is a major issue of economic, social and political significance in India. In a geographically large country like India, different regions are endowed with different resource base and demographics, which in turn results in dissimilar patterns of growth across the different regions. An in-depth analysis of state -level variations will prove useful in framing policies to achieve balanced regional development across Indian states. The subsequent analysis examines whether the inter-regional variations in terms of IEM have increased or declined over time.

Tables 8.1 to 8.3 provide the state-wise estimates of the measures of IEM. Overall, Karnataka, Tamil Nadu and Kerala are among the top states that have shown significant increase in IEM over time. In terms of improvement in upper class persistence, Rajasthan, Maharashtra, and Karnataka are the worst performing states, which have shown increase in persistence at the higher end of the educational level. Results show that downward mobility has increased overtime across all states, except Andhra Pradesh, Kerala and Tamil Nadu which have witnessed decrease in downward mobility. The next section examines how differences across states, in terms of growth and state government policies, have influenced the change in inequality of income opportunity over time.

8.4.2 Macroeconomic Determinants of IEM

For the current analysis, we have used NSS dataset. However, the mobility estimates obtained from NSS are likely to be biased as we are bound to conduct the analysis for father-son pairs who coreside in the same household. The sample truncation due to coresidency could challenge the robustness of our estimates. Therefore, we use IHDS- II dataset and estimate the IEM estimates for Indian states. Both NSS (68^{th} round) and IHDS-II surveys were carried out during 2011-12. Further, IHDS data allows us to estimate the IEM without imposition of coresidency restriction (discussed in detail in the data and methodology section). Therefore, comparison of mobility estimates for different Indian states based on NSS and IHDS dataset allows us to analyse to what extent our results could be biased due to coresidency restriction.

Table 8.4 provides the ranking of states based on the estimates of IEM. As discussed above, we have used two datasets: NSS 68^{th} round and IHDS-II to estimates the IEM estimates. The states have been ranked according to the value of the estimates, where the highest and the lowest ranks correspond to the highest and lowest IEM in terms of educational attainment, respectively. The rank correlations between the state rankings based on NSS and IHDS indicate that there is high and significant pair-wise rank correlation between states ranking. Moreover, the Friedman test of rank independence (shown in Table 8.5) rejects the null hypothesis that the rankings based on the different inequality measures are significantly different from each other. It indicates that our results are less likely to be biased.

Table 8.7 examines the relationship between the changes in the IEM estimates and performance of policies in India. It is expected that most of the pro-poor policies have a delayed or lingering effect on the welfare outcomes; therefore, the variables corresponding to the performance of different policies across different states have been taken in the lagged form. The description of all the policy variables used in the analysis has been shown in Table 8.6. We have ranked all the fifteen states based on the percentage change in the inequality of opportunity index. We build on Besley etal.(2007), in order to rank the Indian states in terms of their economic performance and pro-poor policies. The states with highest percentage reduction in IEM estimates is ranked highest, i.e., rank=1. However, ranking of Indian states based on the capability of states to reduce inequality of educational opportunity has been taken form Assadullah & Yalonetzky (2012).

States with higher poverty reduction rate were more capable in decreasing the downward mobility where children are moving down the education ladder as compared to the education level of their parents. Results also suggest significant negative correlation between labour regulations and bottom upward mobility. In India, where each state has the freedom to draft its own labour law; multiplicity of labour laws challenge investors. States with more pro-worker policies were more successful in reducing increasing bottom upward mobility. In case of gender variable, results suggest that states with higher female representation at workplace were ranked higher in terms of reduction of downward mobility as well as increasing the bottom upward mobility. Significant results indicate that progress towards gender parity can be an important tool to increase bottom upward mobility and reduce the downward mobility. Another important policy variable is 'inequality of educational opportunity'. Our results show that states that were more capable in reducing inequality of educational opportunity were found to be more efficient in lowering downward mobility and increasing the upward mobility at the bottom end of the educational distribution.

Another interesting result is that higher voice accountability is found to be positively correlated to the degree of upper-class persistence. However, states with higher voice accountability are found to be less capable in increasing upward mobility and decreasing downward mobility (see, Table 8.7). This raises concern about the role played by mass media in perpetuating mobility across generations.

The correlation of human capital variable which is proxied as education expenditure per-capita is found to be insignificant, irrespective of the mobility indicators used. This points out the fact that education may strongly determines the employment status of an individual but may not be a determining factor in reducing the income opportunity. All other policy indicators have insignificant effect on IEM estimates.

In short, States witnessing poverty reduction, increase in pro-worker policies and greater gender parity at workplace were successfully able to increase bottom upward mobility. However, this analysis gives an informal look at the impact of policy variables on the inequality of opportunity which has provided us a backdrop for further research work necessary to explore the relationship between the performance of various policy indicators and the inequality of opportunity.

8.5 Conclusion

This chapter studies the inter-state variations in IEM estimates and how they have changed over the years. Results shows that there is considerable variations in mobility estimates across the major Indian states. We have used six different mobility indicators to get a better understanding of the mobility pattern.

To study the correlates of IEM, we build on data available from previous studies (Besley, Burgess, and Esteve-Volart, 2007; Assadullah & Yalonetzky, 2012). We have used the ranking of Indian states based on their performance regarding different policy variables over the years. Further, we rank the Indian states according to the rate of change in the mobility estimates. Results from correlation analysis show that inequality of educational opportunity and rigidity of the labour market are the main obstacle in increasing mobility. We find that flexible labour laws along with policies which promote equality of educational opportunities are effective in increasing Bottom Upward Mobility (BUM). Additionally, results show that reduction in poverty rate is an effective tool to increase mobility by preventing the younger generation from moving down the education ladder as compared to the education level of their parents. In terms of policy implications, state governments need to focus on bringing more flexibility to the labour laws. We need to introduce labour reforms which can encourage equality of opportunity. Despite implementation of various laws, the caste and gender inequality in India adversely impacts the economic outcomes of the society. For example, Dalit women face double plight of gender and caste. The oppression of Dalit women by the upper caste members of the society makes it harder for them to strive to put an end to their misery. They are denied of basic amenities which are controlled by the upper sections of the society and are more prone to physical and sexual violence due to suppression by men. Therefore, policies targeting at increasing basic amenities and reducing drug abuse are unlikely to change the future of Dalit women. The only possibility is to rely on labour laws which can give them voice and an option to break out from the trap of casteism. However, the complexity of labour laws, dis-incentivizing formal labour market, stringent regulatory framework and biased hiring process are likely to deter the upward mobility of Dalit women. Thus, introduction of sensible labour reforms is the need of the hour which can give voice to the backward and oppressed people of the society.

However, the findings have its own shortcomings. Future research should use richer datasets to establish a causal link between state- level policy variables and IEM. Larger dataset will allow to conduct cohort-level study.

Table 8.1: Trends in IEM across Indian States (IRC and ICC estimates)

States	*IRC*					*ICC*				
	All	1947-56	1957-66	1967-76	1977-86	All	1947-56	1957-66	1967-76	1977-86
AP	0.652	0.695	0.681	0.612	0.532	0.484	0.538	0.485	0.469	0.438
AS	0.581	0.724	0.646	0.481	0.503	0.525	0.488	0.486	0.427	0.578
BI	0.678	0.776	0.731	0.605	0.618	0.594	0.564	0.608	0.554	0.609
GU	0.602	0.727	0.609	0.529	0.553	0.569	0.559	0.497	0.541	0.606
HR	0.533	0.688	0.558	0.469	0.422	0.497	0.463	0.428	0.467	0.486
KA	0.611	0.793	0.644	0.527	0.474	0.466	0.497	0.439	0.411	0.451
KE	0.491	0.511	0.407	0.417	0.417	0.508	0.465	0.401	0.511	0.504
MP	0.632	0.793	0.711	0.583	0.505	0.513	0.473	0.479	0.494	0.496
MH	0.536	0.645	0.594	0.467	0.398	0.481	0.455	0.444	0.413	0.473
OR	0.702	0.862	0.794	0.629	0.541	0.541	0.539	0.562	0.491	0.514
PU	0.543	0.709	0.505	0.484	0.486	0.507	0.488	0.411	0.495	0.544
RJ	0.623	0.671	0.728	0.537	0.564	0.516	0.429	0.524	0.462	0.546
TN	0.598	0.637	0.621	0.583	0.456	0.568	0.582	0.599	0.598	0.462
UP	0.621	0.812	0.647	0.542	0.565	0.528	0.544	0.503	0.486	0.543
WB	0.741	0.818	0.745	0.746	0.672	0.665	0.681	0.633	0.661	0.677

Table 8.2: Trends in transition probabilities across Indian states (secondary education as the minimum threshold)

States	*BUM*					*UCP*				
	All	1947-56	1957-66	1967-76	1977-86	All	1947-56	1957-66	1967-76	1977-86
AP	0.267	0.113	0.187	0.281	0.423	0.837	0.857	0.744	0.881	0.847
AS	0.336	0.312	0.237	0.396	0.395	0.872	0.846	0.895	0.806	0.895
BI	0.256	0.222	0.227	0.307	0.264	0.828	0.854	0.802	0.864	0.805
GU	0.245	0.243	0.212	0.246	0.275	0.801	0.788	0.766	0.821	0.804
HR	0.346	0.271	0.284	0.371	0.436	0.833	0.952	0.818	0.854	0.811
KA	0.266	0.184	0.214	0.303	0.337	0.781	0.765	0.762	0.751	0.813
KE	0.396	0.277	0.368	0.418	0.567	0.912	0.917	0.759	0.959	0.947
MP	0.181	0.129	0.175	0.197	0.205	0.734	0.846	0.872	0.688	0.708
MH	0.398	0.273	0.358	0.445	0.482	0.861	0.756	0.911	0.821	0.887
OR	0.208	0.151	0.192	0.212	0.271	0.874	0.876	0.895	0.807	0.882
PU	0.338	0.285	0.345	0.321	0.395	0.811	0.793	0.786	0.797	0.832
RJ	0.219	0.171	0.243	0.215	0.236	0.781	0.742	0.827	0.734	0.795
TN	0.258	0.193	0.171	0.243	0.421	0.881	0.911	0.893	0.843	0.899
UP	0.251	0.198	0.263	0.282	0.251	0.79	0.769	0.796	0.823	0.766
WB	0.197	0.231	0.155	0.184	0.234	0.814	0.805	0.795	0.827	0.822

Notes: All the estimates are significant at 1% level of significance.

Table 8.3: Trends in absolute and directional mobility across Indian states

	Absolute mobility (M1)					*Directional mobility (M2)*				
States	All	1947-56	1957-66	1967-76	1977-86	All	1947-56	1957-66	1967-76	1977-86
AP	4.174	2.678	3.242	4.392	5.651	3.651	2.178	2.665	3.861	5.173
AS	4.646	4.453	4.478	5.348	4.287	4.134	3.972	4.048	4.957	3.577
BI	3.772	3.453	3.533	4.141	3.871	2.949	2.844	2.687	3.264	2.946
GU	4.047	3.791	3.908	4.311	4.086	3.453	3.364	3.407	3.667	3.359
HR	4.911	4.311	4.665	5.329	5.081	4.419	4.044	4.399	4.758	4.356
KA	4.558	3.666	4.063	4.907	5.268	4.072	3.358	3.611	4.342	4.703
KE	4.794	4.309	4.823	5.159	4.865	4.578	4.017	4.528	5.042	4.709
MP	4.028	3.407	3.746	4.298	4.385	3.425	3.035	3.327	3.644	3.544
MH	5.191	4.514	5.224	5.651	5.142	4.678	4.068	4.771	5.139	4.541
OR	4.059	3.425	3.803	4.507	4.339	3.473	2.911	3.121	3.964	3.751
PU	4.587	4.463	4.845	4.518	4.521	4.032	4.111	4.326	3.835	3.913
RJ	4.335	3.653	4.412	4.665	4.441	3.695	3.134	4.003	3.976	3.601
TN	3.591	3.265	2.985	3.474	4.794	3.129	2.475	2.425	3.037	4.459
UP	4.369	3.768	4.505	4.672	4.403	3.653	3.327	4.012	3.803	3.482
WB	3.443	3.413	3.307	3.516	3.526	2.691	2.682	2.373	2.706	2.669

Notes: All the estimates are significant at 1% level of significance.

Table 8.4: Ranking of Indian states based on the estimated IEM estimates

States	ICC	
	IHDS	NSS
AP	3	1
AS	9	13
BI	14	15
GU	13	9
HR	4	4
KA	1	2
KE	6	6
MP	7	5
MH	2	3
OR	11	12
PU	5	7
RJ	8	8
TN	12	11
UP	10	10
WB	15	14
rank correlation	0.911***	

*Note:**** ,significant at 1% level of significance

Table 8.5: Friedman test of rank independence

2011-12
Friedman=0.216
Kendall=0.004
p-value=0.041

Notes: The null hypothesis of Friedman's test is that the rankings based on the different datasets are independent of each other which gets rejected at 95 per cent.

Table 8.6: Description of Policy Variables

Policy Variables	Description	Definition
P1	Poverty Reduction	Poverty headcount ratio
P2	Growth rate	Real income per capita
P3	Voice accountability	Newspaper circulation per-capita
P4	Regulation	Labour regulations
P5	Access to finance	Total credit per capita (per-capita credit extended by the commercial banks of the state)
P6	Human capital investment	State education expenditures per capita
P7	Gender	female-to-male workers
P8	Inequality of educational opportunity	inequality of educational opportunity based on religion and gender

Source: Besley et al. (2007); Asadullah & Yalonetzky (2012)

Table 8.7: Changes in IEM estimates across major Indian states and it's policy implications

States	% change in mobility estimates				Ranking of growth progress and other policies of Indian states							
	M1-M2	*ICC*	*BUM*	*UCP*	*P1*	*P2*	*P3*	*P4*	*P5*	*P6*	*P7*	*P8*
AP	3	2	15	12	4	5	10	1	9	12	4	8
AS	7	14	4	4	16	10	15	7	15	4	2	7
BI	8	10	3	13	14	16	14	7	16	15	12	12
GU	11	12	2	8	6	6	6	9	6	7	8	13
HR	15	9	9	14	9	1	13	7	7	2	14	11
KA	12	3	12	3	8	8	7	5	5	8	3	9
KE	2	11	13	6	1	7	1	4	8	5	1	1
MP	14	8	8	15	13	9	12	6	11	14	9	10
MH	6	7	10	1	11	4	2	10	1	13	7	6
OR	4	4	11	9	7	12	16	8	14	6	6	3
PU	10	13	6	5	3	2	5	7	3	3	15	2
RJ	9	15	7	2	12	14	9	3	13	9	11	15
TN	1	1	14	10	5	3	3	2	2	10	5	4
UP	13	6	5	11	10	15	8	7	12	16	13	14
WB	5	5	1	7	2	11	4	11	4	11	10	5
rank correlations												
M1-M2					**0.453**	0.039	0.314	0.185	0.117	-0.004	**0.582**	**0.646**
					(0.08)	(0.88)	(0.25)	(0.51)	(0.67)	(0.98)	**(0.02)**	**(0.01)**
ICC					0.317	0.117	0.111	0.149	0.314	-0.382	0.271	0.232
					(0.24)	(0.67)	(0.69)	(0.59)	(0.25)	(0.15)	(0.32)	(0.40)
BUM					-0.289	-0.435	-0.164	**-0.654**	-0.225	-0.092	**-0.517**	**-0.432**
					(0.29)	(0.11)	(0.55)	**(0.00)**	(0.42)	(0.74)	**(0.04)**	**(0.09)**
UCP					0.057	0.025	**0.435**	-0.141	0.31	0.271	0.296	0.225
					(0.83)	(0.92)	**(0.08)**	(0.61)	(0.27)	(0.32)	(0.28)	(0.42)

Notes: (a) Figures of rankings of Growth Elasticities of Poverty, Growth Rates, and Policies of Indian States correspond to the period 1958–2000 and are obtained from Besley, Burgess and Esteve-Volart (2007). “Voice and accountability” is measured by newspaper circulation per capita; Regulation refers to the labour laws ranking highest for state with the most restrictive labour laws; Finance=Access to finance proxied by total credit per capita;(b) Rankings are based on the average variable of interest over the period (1 = highest). (c) Significant levels for correlations are in parentheses. (d) Ranks in terms of changes in inequality indices are used. (e) Significant coefficients are highlighted in bold.

Chapter 9

Summary and Conclusions

9.1 Introduction

IEM is considered as a fundamental issue which helps in understanding the dynamics behind the transformation of the society across generations. Analysing the level of social mobility across generations help us to evaluate the extent of equality of opportunity prevailing in the society. In the last few decades, the topic of inequality has been of great interest to researchers. This could be because high growth rate is generally accompanied by higher inequality. Zhuang (2010) argues that higher inequality act as a barrier in achieving the idea of 'inclusive society'. Some recent studies have investigated the impact of economic reforms on inequality, but most of these studies focus on a single economic outcome (Krishna & Sethupathy, 2011; Motiram & Sarma, 2014; Sarkar & Mehta, 2010; Vakulabharanam, 2010). In contrast, there are only a few studies that have analysed the inequality across generations. Difference in social mobility across generations could lead to different economic outcomes. It is argued that given the equal distribution of genetic ability and talent across different socioeconomic groups, a highly mobile society will be able to develop faster by optimum utilization of it's human capital. In addition, in highly mobile society there is less probability of societal conflict for redistributive policies, which affects economic growth of the society. Third, societies experiencing high social mobility across generations is more likely to provide equal access to education, regardless of the socioeconomic position of the household.

Past studies have used different datasets to measure the degree of IEM in India. But, no other study has attempted to identify the determinants of IEM in India. The present study has focused on measuring robust estimates of IEM in India. In addition, the study also identifies the correlates and determinants of IEM. This chapter has been divided into three sections. In the following section, we present the summary of the finding of our study. Section 9.3 discusses the policy implications of our findings. Last section concludes by discussing the limitations of the present work and scope for further research.

9.2 Summary of the findings

The issue of IEM has garnered a lot of attention recently, especially in the last few years. The various aspects of IEM and its strong relationship with economic variables makes it an important topic. Considering the importance of this issue, this study has measured robust estimates of IEM in India. We have also analysed the data to identify the correlates and determinants of IEM in India. The study has focused on educational mobility rather than occupational mobility or income mobility. The main reason for this is attributed to lack of reliable long-term data on income and occupation of the individuals. In addition, an individual is likely to complete his/her education by 25 years of age, therefore there is a very less chance of biasness due to measurement error when measuring educational mobility.

The broad objectives of the study were as follows:

(i) To measure and analyse the trend in IEM in India and across major Indian states.

(ii) To measure the robust estimates of IEM by employing various empirical methodologies.

(iii) To assess the degree of heterogeneity in mobility estimates among migrants (children who migrated with their family during their schooling years) and non-migrants.

(iv) To examine the impact of return migration on schooling progression of the child and its impact on IEM.

(v) To examine the impact of monetary/non-monetary aid on schooling outcome (schooling gap) of the child and the role of parental education in mediating this relationship.

(vi) Finally, to identify the macroeconomic correlates of IEM in India.

The primary purpose of Chapter 4 was to measure the degree of IEM in India. The literature on IEM in the developing countries remains sparse, and it is frequently plagued with methodological issues such as the endogeneity bias. Different approaches have been employed in the literature to account for the endogeneity bias by using data on adopted children, twins, or instrumental variable approach. However, all these approaches have their own limitations, and are therefore, unable to eliminate the endogeneity problem. In this study, we mitigated the bias resulting from potential endogeneity of parental education by using a novel two stage estimation strategy proposed by Lewbel (2012). Our two-stage least squares (2SLS) results are based on a new identification strategy that does not require additional data, such as instrumental variables that are uncorrelated to the error term.

We drew some major conclusions from our analyses. First, we found that although the degree of persistence in terms of education has declined steadily, implying increasing mobility; parental education even now plays a very significant role in affecting child's education level. Second, decomposing the correlation coefficient between father's years of schooling and child's years of schooling revealed that although the intergenerational persistence has declined over time, implying an increasing mobility; it is still significant. The positive persistence at the lower end of the education distribution has increased from 43% to 61% over time, while negative persistence at the lower end has soared from 5% to 22%. In other words, the proportion of highly educated fathers with sons who are equally well-off in terms of education has risen over time and the proportion of illiterate fathers with sons who are highly educated has declined over time.

Third, the traditional instrumental variable approach is of limited use with secondary databases, as it is difficult to identify variables which are independent of error term and which do not affect the dependant variable when independent variable is held constant (exclusion restriction). We use an alternative identification strategy proposed by Lewbel (2012) which replaces endogenous regressors, such as parental education, with synthetic instrumental variables constructed using linear combinations of exogenous regressors. The major advantage of this identification strategy is that it does not depend on the standard exclusion restriction. In general, the OLS estimates can be upward biased, as opposed to the 2SLS estimates, due to unobserved nature effects. For instance, unobserved natural ability which is positively related to the education level of both parents and their children can induce an upward bias in the OLS estimates. However, after controlling for potential endogeneity of parental education level, we find that the 2SLS estimates are considerably smaller than the corresponding OLS estimates. This suggests that regression-based measures may potentially overestimate intergenerational educational persistence due to heterogeneity in unobserved characteristics such as ability and preference. Notwithstanding, the Lewbel IV models, which address the omitted ability bias and potential measurement error in parental education, also show a substantial positive effect of parental education on child's education.

In Chapter 5, we analysed the impact of migration on child's education. We defined the migration variable as a dummy which was 1 if the child experienced migration during his/her schooling days, 0 otherwise. In addition, we also calculated the age at the time of migration to study its effect on child's educational level. Our results showed that migrants are more mobile than their non-migrant counterparts. However, decomposing the correlation coefficient among the migrants revealed that higher mobility in terms of educational attainment observed among migrant population is primarily due to the migrant children moving down the education ladder relative to their parents. Another useful finding was the negative association between the age at the time of migration and child's educational outcomes. However, this negative association was found to be greater for daughter than sons. This shows that daughters are at double disadvantage.

Chapter 6 focused on the impact of return migration on IEM. IHDS-2 provides us with an additional information on return migrants who migrated to other areas but returned home to live there. For our analysis, we identified migrant households who had at least one return migrant parent. Our analysis in this chapter was focused only on father-son pairs. This chapter attempted to fill a void in the literature by examining the link between parental return migration and IEM in India. We used a dynamic sequential framework that allowed the effects of the determinants of child's education to vary across different educational transitions. Examining determinants schooling progression is more informative than a static analysis of education attainment, as Indian educational system is characterized by large difference in enrolment and dropout rates at different levels of education. The sequential logit framework used in this study models the conditional sequence of educational transitions, considering both the successful completion of the previous education level as well as any self-selection into the next education level. Therefore, it provides a rich description of the nature of the selection process at each transition, and allows policymakers to identify which children progress less than others, and also to locate the education level at which they are likely to dropout. This is useful for framing and assessing targeted policies that are aimed at improving enrolment or reducing dropouts at specific levels of schooling.

Further, we conducted sensitivity analysis proposed by Buis (2011) to examine the robustness of our results under a wide range of scenarios for unobserved heterogeneity. These scenarios vary both in terms of the magnitude of unobserved heterogeneity and its correlation with our main explanatory variables (parental education and return migration). We found that an increase in father's education improves the odds of primary to secondary and secondary to post-secondary transitions of sons. Father's migration experience reduces IEM (increases the effect of father's education) for primary to secondary transition of sons, and it increases IEM for the secondary to post-secondary transition. The odds of educational transitions by daughters improve with an increase in mother's education, but they are unrelated to father's education and migration status. This suggests that programs to encourage maternal education can have significant effect on the schooling progression of daughters. The literature on the impact of maternal literacy programs on their child' schooling in India is sparse. The only significant work that we are aware of is Banerji et al. (2017), who used a randomized evaluation to examine the causal effect of maternal literacy programs in two Indian states with relatively low literacy levels – Rajasthan and Bihar. They found that mother's participation in literacy programs improved the math and language scores of their children. Our results are consistent with their results, and suggest the positive effect of maternal education on child's schooling progression is not an artefact of the endogeneity bias induced by unobserved factors.

We also found that household assets have a positive effect on the odds of educational transitions for the children in the household, and they have a larger effect on daughters' transitions than on sons. The presence of younger siblings lowers the odds of transition to higher levels of education for both sons and daughters. These effects are large and robust to unobserved heterogeneity. In economically constrained households, children compete to the household resources, and there is a greater likelihood that the children will be engaged in some part-time employment to supplement household income. Policy initiatives that provides some form of financial assistance, such as free books, school uniform, school meals, can encourage child schooling economically constrained families[8]. More targeted initiatives like a scholarship for girl child, or a scholarship for a girl child who reaches a particular level of education can reduce the difference between schooling progression of sons and daughters.

Despite having a patchwork of programs to subsidize higher education, the cost to access higher education has been increasing over time in India. These costs include both high school fees, as well as the opportunity cost of lost child earnings, which is higher for older children as they are more suited to participate in economic activities and domestic farm work than younger children who are enrolled in primary education. The problem of high costs of higher education level can be mitigated through the following policy prescriptions. First, a substantial proportion of the tuition fees could be deferred and paid after completion of the course. This ensures that the incidence of dropouts due to financial constraints would reduce. Also, after clearing a higher education level the child is likely to have better employment opportunities, which would make the costs of education more affordable. Second, each registered institute of higher education should be required to have a tie-up with one or more lending institutions that provide education loans to finance costs of higher education. Indian banks are required to reserve some proportion of there overall lending for priority sector lending which includes education loans. Unfortunately, the top-down policy approach to allocate some proportion of overall credit in the banking system towards priority sectors does not guide the distribution of this credit. By ensuring that each institute of higher

[8] There is some evidence in India that found that such targeted policy initiatives improve child schooling outcomes. For example, Afridi (2011) found that transitioning from providing free monthly ration to a daily provision of free cooked meals to school children increased the attendance of young girl child by 12 percentage points.

education has some provision of providing credit for its students, policymakers could increase both the awareness and access to education loans. Third, the fee structure of all higher education programs should be required meet a minimum threshold level of cross subsidization, wherein some portion of overall the fee income is dedicated towards subsidizing education of children from low-income households[9].

In Chapter 7 we analysed the IEM by using a different indicator. In Chapter 5, we regressed the years of completed education of the child on the completed years of education of the parent to obtain the IEM estimates. In this chapter, we regressed the years of schooling gap of the child on the education level of the parent. In short, we used schooling gap as an indicator of child's educational progress. We measured years of schooling gap of the child and studied the trend and pattern in schooling gap. The analysis in this chapter was based on panel information provided by two consecutive rounds of IHDS: IHDS-I and IHDS-II. The analysis was also done separately for four different age-groups to avoid any kind of sample selection bias. Results showed that the size of schooling gap reduced between the two survey rounds of IHDS, across all age-groups. In addition, we also found that the average schooling gap increased with the age of the child but remained on an average higher for sons than daughters. This suggests that gender inequality in higher education is still prevalent in the society. In addition, average schooling gap was found to be more for children residing in rural areas than their counterparts.

Results regarding the average schooling gap by household income quintiles showed that the gaps are larger for the household who belong to the lower income quintile. The average schooling gap declines with increasing household income quintile, regardless of child's gender. This suggests

[9] A similar threshold-based policy has been implemented for Indian corporates to encourage expenditure in corporate social responsibility (CSR) activities. All Indian companies (above a certain threshold of profits and annual turnover) are required to spend at least 2 percent of their average net profits made in the preceding three years on CSR activities.

that household income is a crucial factor which affects the schooling gap in children. Statistics showed that the difference between the mean schooling gap between children of bottom and top income quintile households has fell between two rounds of IHDS, both for sons and daughters. However, the reduction in difference was larger for sons than daughters.

In this chapter, we also examined the impact of encouragement on child's educational outcomes. The educational outcome was defined as the years of schooling gap for the child which occurred between the two consecutive survey rounds of IHDS i.e., IHDS-I and IHDS-II. We defined five modes of encouragement, viz., scholarship (fellowship), mid-day meal, free books, free uniform, and school fees which a child received at the time of IHDS-I survey. Our results showed that among all the different modes of providing encouragement to the students, scholarship was found to be the most effective method of reducing the schooling gap. Second, the benefits of providing encouragement in reducing the schooling gap were found to be considerably larger for girls that those for boys.

We also find that for sons having low-educated mothers (1 to 5 years), the most effective modes of reducing the schooling gap are scholarship and school fees. For sons having high-educated mothers (11 years and above), the most effective modes of reducing the schooling gap is free books. We also observed that regardless of the age of the daughter, scholarship reduces the schooling gap by around 0.4 year. The estimated ATT for daughters is more significant than those estimated for sons, for whom only the youngest cohort benefits from scholarship. In addition, we found that daughters of non-literate mothers that receive the scholarship have significantly lower schooling gap than those who do not receive scholarship. For all other forms of encouragement, we find weak evidence that they reduce schooling gap of children.

In chapter 8, we examined the inter-state variations in IEM estimates and identify the macroeconomic correlates of IEM in India. Results shows that there is sizable variations in mobility estimates among the major Indian states. We have used six different mobility indicators to get a better understanding of the mobility pattern.

To study the correlates of IEM, we built on data available from previous studies (Besley, Burgess, and Esteve-Volart, 2007; Assadullah & Yalonetzky, 2012). We used the ranking of Indian states based on their performance regarding different policy variables over the years. Further, we ranked the Indian states according to the rate of change in the mobility estimates. Results from correlation analysis showed that inequality of educational opportunity and rigidity of the labour market are the main obstacle in increasing mobility. We found that flexible labour laws along with policies which promote equality of educational opportunities can be effective tools to encourage Bottom Upward Mobility (BUM). Additionally, results also showed that reduction in poverty rate is an effective tool to increase mobility by preventing the younger generation from moving down the education ladder as compared to the education level of their parents.

9.3 Conclusion and Policy Implications

In Chapter 4, we found that the proportion of highly educated fathers with sons who are equally well-off in terms of education has increased over time and the proportion of illiterate fathers with sons who are highly educated has declined over time.

Therefore, we try understand on how the distribution of public expenditure on education may affect the IEM. Specifically, we measure the tertiary tilt in public expenditure on education, which is the ratio of per-student spending on higher education to the per-student spending on primary education. Our findings reveal that tertiary tilt in education expenditure in India has steadily declined over time. We show how this bias in education expenditure towards primary education could lead to increasing positive persistence at the upper end of the education distribution, and negative persistence at the lower end of the educational distribution. To this extent, the excessive emphasis primary education at the cost of all other types of education may be leading to undesirable outcomes in terms of IEM. Although, the enrolment in higher education and technical institutions in India has been increasing at a very fast pace and most of these students are enrolled in private institutions. Nonetheless, education loan facilities are still a cause for concern in India which compels majority of students to rely on private investment rather than public investment. Therefore, there is an urgent need to introduce reforms at the institutional level.

For instance, the proportion of expenditure on elementary education by the central government has increased from 13.74% in 1990-91 to 61.19% in 2005-06 (refer Fig.9.1). On the other hand, the proportion of expenditure on higher education by central government has declined from 28.94% in 1990-91 to 11.50% in 2005-06. However, the sectoral composition of education expenditure by state government remains stable over time (Refer Table 9.1).

We measure the tertiary tilt in public expenditure on education, which is the ratio of per-student spending on higher education to the per-student spending on primary education. Our findings reveal that tertiary tilt in education expenditure in India has steadily declined over time. The excessive emphasis on primary education at the cost of all other types of education may be leading to undesirable outcomes in terms of IEM. Although, the enrolment in higher education and technical institutions in India has been increasing at a very fast pace and most of these students are enrolled in private institutions. Nonetheless, education loan facilities are still a cause for concern in India which compels majority of students to rely on private investment rather than public investment. Therefore, there is an urgent need to introduce reforms at the institutional level.

Table 9.2 reveals that education expenditure on primary education per student has seen major decline, while education expenditure on higher education per student has increased drastically. This spending bias is called the “tertiary tilt”, and it is measured as the ratio of per-student spending on higher education to the per-student spending on primary education. This bias is well-documented (Addison & Rahman, 2001; Bourguignon et al., 2003; González Rozada & Menendez, 2002; Psacharopoulos, 1977, 1994; Stasavage, 2005).

Gruber & Kosack (2014) have found that most of the country’s government who tilt their education spending towards higher education witness rising primary enrolment, but higher inequality later. Table 9.2 shows a steady decline in the tertiary tilt in the public expenditure on education in India. It may be argued that a declining tertiary tilt is favouring parents who are struggling to afford their child’s elementary education. The concentration of financial resources on primary education might help the less-resourceful children to gain primary education, but at the same time it inhibits other to pursue higher education due to high cost associated with higher education. This implies that it would be easier for the children of non-literate parents to gain primary education, whereas, rising cost of higher education may deter some children of highly educated parents to attain higher education. Therefore, unobserved institutional environment may such as the allocation of public expenditure on education, may result in a negative association between the education levels of the parents and their children.

In 1947 when India got independence, only one out of six Indians was literate. Since the country's independence, the Indian government has sponsored an array of ambitious programs to address the problem of illiteracy. The Sarva Sikhsha Abhiyan (meaning: Education for all movement) is aimed at universalization of primary education giving free and compulsory education to all children in the age group 6-14. The government has also established several important institutions with a dual objective of modernising India's education system and ensuring uniform access to education to all Indians, chief amongst them are the University Education Commission, the Secondary Education Commission, and the University Grants Commission. The challenge of illiteracy in India is immense, not just because of the scale of the problem but because of the vast disparities in literacy that exist along the lines of gender, caste, religion, geographical region, socioeconomic class, and migration status. While the overall literacy rate has witnessed a marked improvement from 16.1 percent in 1941 to 74.04 percent in 2011, these disparities still persist. This study aimed to motivate targeted policy action to tackle the problems faced by female and migrant children in Indian education system.

We also highlight how the intersection of gender identity and migration status affects IEM. The effect of gender on IEM has attracted some research attention, however, the effect of migration status remains largely unexplored. This is surprising given the vast population of migrant workers in India. According to the Census of 2011, there are 139 million internal migrants in the country. Most of the migrant workers tend to enter the job market at a very early age, experience no upward mobility and are involved in low-paying, unskilled, informal sector jobs for their entire work-life (Sharma, 2017). With manual labour taking its toll, and poor access to public health services, migrant workers are often forced to go back to their hometown due to health problems. This lowers their household income, forcing their children to start migrating for work at a relatively young age. The vicious cycle has perverse intergenerational consequences, transferring poor wellbeing and low level of education from parents to children.

Our results highlight how the general improvement in IEM over time can mask latent disparities, which can only be addressed through targeted policy measures. We find that IEM has increased over time; however, parental education remains a significant predictor of their child's education attainment. The decomposition of ICC suggests that most of the persistence emanates from tails of the educational distribution. More specifically, from pairs where both generations have low education or from pairs where both generations have high education. This is indicative of poor prospects of upward mobility for children having the least educated parents. In addition, we find that gender and migration status have a significant effect on IEM. The education level of father is more strongly correlated to the education level of sons than that of daughters, whereas maternal education is more strongly correlated to the education level of daughters than that of sons. The most concerning finding, however, was the incidence of downward mobility among migrant children. Children who have had some migration experience during their schooling years are more likely to be less educated than their parents than those that did not experience migration. The effect of "age at the time of migration" is also informative, with the youngest children faring worse in terms of their eventual educational attainment. This implies that the environment at host areas impedes the educational progress of children of migrant families. Thus, a suitable policy response to address the educational disadvantage of migrant children should focus on providing equal opportunities to migrant households.

We recommend some policy measures to address the disadvantage faced by migrant children. Due to a lack of local proofs of identify and residence, migrant families are often unable to receive social welfare entitlement such as subsidized food under public distribution system. These requirements also hinder access to banking services, and connections for utility services such as cooking gas. A government initiative that integrate access to all public services in a single platform that uses the Aadhaar information (12-digit national identification number linked to basic demographic and basic biometric information) for authentication will obviate the need for local proofs of residence and identity. Second, policy response needs to address the vicious cycle where poor occupational health of migrant parents eventually forces their children to migrate at a young age, usually working in similar unskilled, low-paying jobs as their parents with poor prospects of upward mobility. Government investment in affordable public health services, especially focused in urban areas with the highest concentration of migrant population would be well placed. Third, there is a need to improve the effectiveness of the existing legal framework in resolving informal sector disputes. Migrant workers routinely face workplace disputes related to non-payment of wages, compensation for workplace accidents and even deaths. A significant institutional reform could be establishing a National Commission for Migrant workers that represents the right of migrant workers, and provides advisory to state and central governments on all policy matters that affect migrant workers. Fourth, the government education initiatives should aim to sensitize school authorities about the various disadvantages faced migrant children. Migrant children face added challenges of adapting to a new learning environment with different linguistic and academic practices, which leads to an increase in dropout rates among them. Dropout rates among migrant children can be reduced if schools make concerted efforts to improve awareness of migrant parents regarding the economic benefits of educating their children and the different support schemes available to them. For example, schools may have policies that provide financial support to children belonging to low-income families, but often the migrant families are not able to avail them due to language barriers, lack of familiarity with the administration process (Ainscow & Hargreaves, 2016) or due to the stigma attached to claiming financial support (Baumberg Geiger, 2016). Simple measures such as translating standard school textbooks in all major regional languages and providing an open access to them through knowledge portals can help migrant children overcome the linguistic barriers. School administration should also focus on preventing discrimination against migrant children and make efforts towards promoting community cohesion to enable better integration of students from different backgrounds.

This study also offers some implications for research in IEM. First, intuition suggests that in the presence of unobserved confounders, the IRC and ICC coefficients are likely to present a conservative estimate of IEM. This is because the most widely documented unobserved confounders such as social connections, parental attitude towards education, or genetically transmitted ability are likely to have directionally similar effects on the education attainment of both parents and children. We believe that researchers should be cautious regarding the direction of the potential endogeneity bias in mobility measures. Indeed, in our analysis, the IV-Lewbel estimates of intergenerational educational persistence are substantially larger than the corresponding OLS estimates, suggesting that not accounting for unobserved heterogeneity leads to overestimation, not underestimation, of mobility measures. This suggests the presence of some unobserved confounders that display opposite relation with the educational levels of the parents and their children. In the Indian context, this could be attributed to the fact that the proportion of public expenditure on education towards primary education has been increasing as compared to that on higher education. As discussed earlier, a declining tertiary tilt favors parents who would otherwise struggle to afford their child’s elementary education. The concentration of financial resources on primary education makes it easier for the children of non-literate parents to gain primary education, whereas, rising cost of higher education may deter some children of highly educated parents to obtain higher education themselves. Therefore, unobserved institutional environment such as the allocation of public expenditure on education, may result in a negative association between the education level of the parents and their children.

Research on IEM in India has provided us with many descriptive and causal evidence. The descriptive results from our research on IEM in India include (i) There has been an upward trend in the degree of IEM; (ii) mobility appears higher for individuals belonging to the marginalized section of the society than for non-marginalized individuals; (iii) the extent of IEM varies across Indian states, and (iv) that the mobility levels are determined by different macroeconomic variables.

An important focus of this research is identifying the correlates and determinants of IEM – what factors play important role in determining the degree of mobility across generation? Our results indicate that migration status of an individual affect the opportunities available for the children. For example, children who had some migration experience at the time of schooling were found to be more mobile than children who did not experience any kind of migration. Decomposition of mobility estimates show that migrants being more mobile than their non-migrant counterparts is attributed to increasing proportion of children moving down the education ladder relative to their parent's educational level. However, age at the time of migration plays a very important role in child's educational attainment. Children who experience migration at an early age are more likely to move up the social ladder in terms of educational attainment. It is therefore likely that policies that may provide equal opportunities to children of all sections of the society can increase mobility. Accordingly, comprehensive schooling system and better environment for migrants at the destination places could be beneficial for children from migrant families.

It is also plausible that discrimination based on race and ethnicity may affect mobility. Our results indicate that children from marginalized sections of the society are more likely to fall behind in terms of education than their non-marginalized groups. Therefore, if some groups are consistently privileged then it will lead to persistent inequality across generations. Therefore, policies that may help in combating discrimination which may reduce the prevalent friction in the labour market are likely to improve mobility.

Household composition also plays an important role in affecting mobility. Our results suggest that children who have more siblings are more less likely to move up the education ladder. In order to keep the doors to higher education open for all, it becomes crucial to provide free quality education to all. In addition, provision of employment opportunities that may absorb different types of labour (skilled and unskilled) may help in improving IEM.

In Chapter 8, results from correlation analysis showed that states with flexible labour laws, greater gender parity and higher equality of opportunity were more capable in encouraging bottom upward mobility. In terms of policy implications, state governments need to focus on bringing more flexibility to the labour laws. We need to introduce labour reforms which can encourage equality of opportunity. Despite implementation of various laws, the caste and gender inequality in India adversely impacts the economic outcomes of the society. For example, Dalit women face double plight of gender and caste. The oppression of Dalit women by the upper caste members of the society makes it harder for them to strive to put an end to their misery. They are denied of basic amenities which are controlled by the upper sections of the society and are more prone to physical and sexual violence due to suppression by men. Therefore, policies targeting at increasing basic amenities and reducing drug abuse are unlikely to change the future of Dalit women. The only possibility is to rely on labour laws which can give them voice and an option to break out from the trap of casteism. However, the complexity of labour laws, dis-incentivizing formal labour market, stringent regulatory framework and biased hiring process are likely to deter the upward mobility of Dalit women. Thus, introduction of sensible labour reforms is the need of the hour which can give voice to the backward and oppressed people of the society.

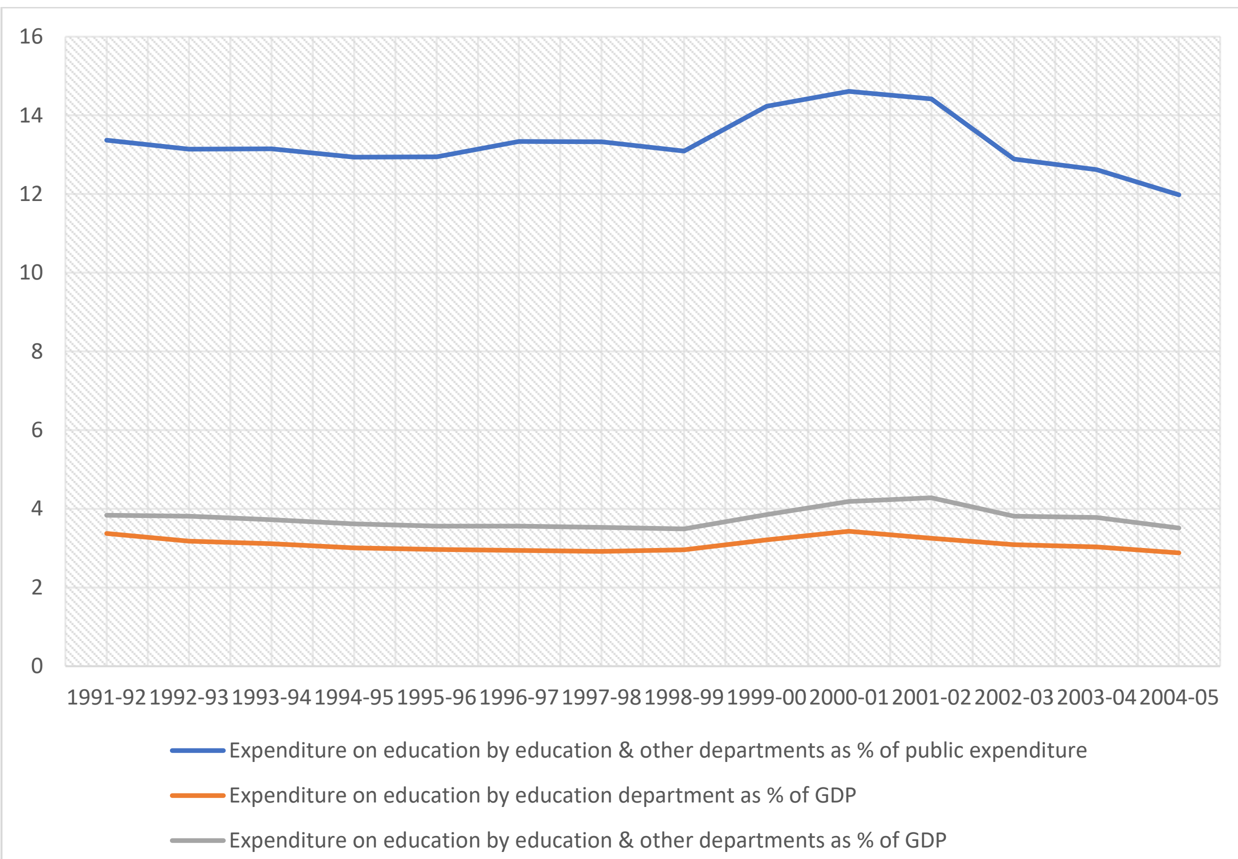
In terms of government accountability, we found that states with higher circulation of newspapers per capita were more capable in Upper Class Persistence (UCP). However, the results come out to be insignificant for Bottom Upward Mobility (BUM). This raises a serious concern pertaining to the role of mass media in our society. The change in the society can be brought through media by transmitting vital information regarding the importance of education and by being more informative about the changes across distributions. A lot of perceptions in our society are based on the information which we receive through different modes of media. Thus, media needs to focus on transformative role by challenging traditional gender stereotypes and social and cultural norms which will give voice to women to fight for equal rights.

9.4 Limitations and scope for future research

This study examined the correlates and determinants of IEM in India. The issue analysed this topic theoretically and empirically. The focus of the study remained on educational mobility. But the work can also be extended by analysing income or occupational mobility in India. In case of developing countries like India, there is lack of secondary datasets which can provide long-term information on income and occupation of an individual. Therefore, future research can focus on identifying the correlates and determinants of occupational and income mobility by conducting primary surveys. Also, future studies can also incorporate qualitative factors, like parental attitude, total time spent with kids per day, etc. which the researchers feel could be important determinants of IEM. The study can also be extended to other developing countries. Additionally, a cross-country analysis of developing countries can help us in understanding the dynamics of IEM.

Another limitation of the present study is lack of causal evidence regarding the impact of different policy variables on the extent of IEM. This is due to the unavailability of long-term panel information on Indian households. Also, the analysis could be extended to other states, which remain otherwise out of focus due to their negligible contribution to the economy or due to lack of reliable information

Figure 9.1: Trends in education expenditure (1991-92 to 2005-06)



Source : - (1) GDP figures are taken from National Accounts Statistics published by C S O.
(2) Expenditure on Education Figures are taken from Budgeted Expenditure on Education published by D/o Higher Education

Table 9.1: Sectoral Composition of Expenditure by Education Departments, %

Year	All States and Union Territories					
	Elementary	Secondary	Higher	Technical	Other	Total
1990-91	49.71	33.08	11.81	2.86	2.54	100
1991-92	49.29	33.98	11.43	2.90	2.40	100
1992-93	45.23	34.26	12.89	4.33	3.30	100
1993-94	46.22	33.14	13.26	4.35	3.04	100
1994-95	49.01	34.28	11.52	2.94	2.26	100
1995-96	49.62	33.51	11.41	3.03	2.42	100
1996-97	49.81	33.69	10.92	2.83	2.74	100
1997-98	49.76	34.44	10.75	2.83	2.23	100
1998-99	49.80	35.05	10.06	2.81	2.27	100
1999-00	46.45	36.66	11.21	2.67	3.01	100
2000-01	48.82	34.05	12.66	2.61	1.86	100
2001-02	50.91	33.80	11.34	2.32	1.64	100
2002-03	49.12	34.91	11.95	2.42	1.59	100
2003-04	49.57	34.95	11.61	2.28	1.59	100
2004-05	50.86	33.75	11.04	2.53	1.81	100
2005-06	51.01	33.53	11.02	2.72	1.73	100
			Centre			
1990-91	13.74	23.52	28.94	18.74	15.07	100
1991-92	16.50	23.77	28.92	18.43	12.38	100
1992-93	17.60	24.86	28.09	18.52	10.93	100
1993-94	18.59	26.72	24.53	19.33	10.83	100
1994-95	21.39	23.15	26.70	18.13	10.64	100
1995-96	39.55	19.93	19.89	14.04	6.59	100
1996-97	42.53	19.55	19.51	14.28	4.13	100
1997-98	48.37	15.08	20.29	12.77	3.49	100
1998-99	43.51	15.54	25.30	12.81	2.83	100
1999-00	38.85	14.53	30.02	13.89	2.72	100
2000-01	39.35	14.63	28.84	13.94	3.25	100
2001-02	44.44	15.32	20.50	15.45	4.29	100
2002-03	46.87	14.17	19.27	15.19	4.50	100
2003-04	51.13	13.53	17.31	13.76	4.28	100
2004-05	54.64	10.28	15.13	10.90	9.06	100
2005-06	61.19	8.68	11.50	8.73	9.91	100

Source: Calculated from various issues of the "Analysis of Budgeted Expenditure on Education", published by Ministry of HRD, Government of India.

Table 9.2: Tertiary tilt in education expenditure in India, (1991-92 to 2005-06)

Year	Education expenditure on primary education per student	Education expenditure on higher education per student	Tertiary tilt in education expenditure
1991-92	1917.861	10341.892	5.392
1992-93	1947.709	9985.141	5.127
1993-94	2076.422	9929.202	4.782
1994-95	2038.753	9775.280	4.795
1995-96	2207.413	9148.649	4.145
1996-97	2377.811	9031.436	3.798
1997-98	2499.364	9048.281	3.620
1998-99	2812.658	9931.554	3.531
1999-00	2980.887	12433.577	4.171
2000-01	3069.426	12849.637	4.186
2001-02	3245.913	10265.029	3.162
2002-03	2999.749	10210.846	3.404
2003-04	2996.293	9467.373	3.160
2004-05	3201.376	8829.000	2.758
2005-06	3644.809	9183.051	2.519

Source: Author's calculations based on various issues of the "Analysis of Budgeted Expenditure on Education", published by Ministry of HRD, Government of India.

References

Abada, T., & Tenkorang, E. Y. (2009). Gender differences in educational attainment among the children of Canadian immigrants. *International Sociology*, *24*(4), 580–608.

Ackerson, L. K., & Subramanian, S. V. (2008). State gender inequality, socioeconomic status and intimate partner violence (IPV) in India: A multilevel analysis. *Australian Journal of Social Issues*, *43*(1), 81.

Addison, T., & Rahman, A. (2001). *Why is so little spent on educating the poor?* (Working Paper No. 2001/29). WIDER Discussion Paper. https://www.econstor.eu/handle/10419/52953

Afridi, F. (2011). The impact of school meals on school participation: Evidence from rural India. *Journal of Development Studies*, *47*(11), 1636–1656.

Ainscow, M., & Hargreaves, A. (2016). Collaboration as a strategy for promoting equity in education: Possibilities and barriers. *Journal of Professional Capital and Community*.

Alcott, B. (2017). Does teacher encouragement influence students' educational progress? A propensity-score matching analysis. *Research in Higher Education*, *58*(7), 773–804.

Aldous, J., & Hill, R. (1965). Social Cohesion, Lineage Type, and Intergenerational Transmission. *Social Forces*, *43*(4), 471–482. https://doi.org/10.2307/2574453

Ammermueller, A. (2007). Poor Background or Low Returns? Why Immigrant Students in Germany Perform so Poorly in the Programme for International Student Assessment. *Education Economics*, *15*(2), 215–230. https://doi.org/10.1080/09645290701263161

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arora, R. U. (2012). Gender Inequality, Economic Development, and Globalization: A State Level Analysis of India. *The Journal of Developing Areas*, *46*(1), 147–164. https://doi.org/10.1353/jda.2012.0019

Asadullah, M. N., & Yalonetzky, G. (2012). Inequality of Educational Opportunity in India: Changes Over Time and Across States. *World Development*, *40*(6), 1151–1163. https://doi.org/10.1016/j.worlddev.2011.11.008

Atkinson, A. (1981). On Intergenerational Income Mobility in Britain. *Journal of Post Keynesian Economics*, *3*(2), 194–218.

Aydemir, A., Chen, W.-H., & Corak, M. (2009). Intergenerational Earnings Mobility among the Children of Canadian Immigrants. *Review of Economics and Statistics*, *91*(2), 377–397. https://doi.org/10.1162/rest.91.2.377

Azam, M., & Bhatt, V. (2015a). Like Father, Like Son? IEM in India. *Demography*, *52*(6), 1929–1959. https://doi.org/10.1007/s13524-015-0428-8

Azam, M., & Bhatt, V. (2015b). Like Father, Like Son? IEM in India. *Demography*, *52*(6), 1929–1959. https://doi.org/10.1007/s13524-015-0428-8

Baker, D. P., & Stevenson, D. L. (1986). Mothers' Strategies for Children's School Achievement: Managing the Transition to High School. *Sociology of Education*, *59*(3), 156–166. https://doi.org/10.2307/2112340

Banerji, R., Berry, J., & Shotland, M. (2017). The Impact of Maternal Literacy and Participation Programs: Evidence from a Randomized Evaluation in India. *American Economic Journal: Applied Economics*, *9*(4), 303–337. https://doi.org/10.1257/app.20150390

BATTISTELLA, G., & CONACO, Ma. C. G. (1998). The Impact of Labour Migration on the Children Left Behind: A Study of Elementary School Children in the Philippines. *Sojourn: Journal of Social Issues in Southeast Asia*, *13*(2), 220–241. JSTOR.

Bauer, P., & Riphahn, R. T. (2006). Timing of school tracking as a determinant of IEM. *Economics Letters*, *91*(1), 90–97. https://doi.org/10.1016/j.econlet.2005.11.003

Bauer, P. W., Schweitzer, M. E., & Shane, S. A. (2012). Knowledge Matters: The Long-Run Determinants of State Income Growth*. *Journal of Regional Science*, *52*(2), 240–255. https://doi.org/10.1111/j.1467-9787.2011.00729.x

Baumberg Geiger, B. (2016). The stigma of claiming benefits: A quantitative study. *Journal of Social Policy*, *45*(2), 181–199.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, *70*(5, Part 2), 9–49. https://doi.org/10.1086/258724

Becker, G. S. (1981). Altruism in the Family and Selfishness in the Market Place. *Economica*, *48*(189), 1–15. JSTOR. https://doi.org/10.2307/2552939

Becker, G. S., Murphy, K. M., & Tamura, R. (1990). Human Capital, Fertility, and Economic Growth. *Journal of Political Economy*, *98*(5, Part 2), S12–S37. https://doi.org/10.1086/261723

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, *4*(3, Part 2), S1–S39. https://doi.org/10.1086/298118

Behrman, J. R. (1988). Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion? *Oxford Economic Papers*, *40*(1), 32–54.

Behrman, J. R., Gaviria, A., Székely, M., Birdsall, N., & Galiani, S. (2001). IEM in Latin America [with Comments]. *Economía*, *2*(1), 1–44.

Beller, E. (2009). Bringing Intergenerational Social Mobility Research into the Twenty-first Century: Why Mothers Matter. *American Sociological Review*, *74*(4), 507–528. https://doi.org/10.1177/000312240907400401

Beller, E., & Hout, M. (2006). Intergenerational Social Mobility: The United States in Comparative Perspective. *The Future of Children*, *16*(2), 19–36.

Besley, T., & Burgess, R. (2000). Land Reform, Poverty Reduction, and Growth: Evidence from India. *The Quarterly Journal of Economics*, *115*(2), 389–430. https://doi.org/10.1162/003355300554809

Besley, Timothy, Burgess, R., & Esteve-Volart, B. (2007). The Policy Origins of Growth and Poverty in India. *Delivering on the Promise of Pro-Poor Growth, Palgrave Macmillan and the World Bank, New York*, 59–78.

Bhattacharya, P. C. (2006). Economic Development, Gender Inequality, and Demographic Outcomes: Evidence from India. *Population and Development Review*, *32*(2), 263–292. https://doi.org/10.1111/j.1728-4457.2006.00118.x

Bhaumik, S. K., & Chakrabarty, M. (2009). Is education the panacea for economic deprivation of Muslims?: Evidence from wage earners in India, 1987–2005. *Journal of Asian Economics*, *20*(2), 137–149. https://doi.org/10.1016/j.asieco.2008.09.014

Bijwaard, G. E., Schluter, C., & Wahba, J. (2014). The impact of labor market dynamics on the return migration of immigrants. *Review of Economics and Statistics*, *96*(3), 483–494.

Björklund, A., & Jäntti, M. (1997). Intergenerational income mobility in Sweden compared to the United States. *The American Economic Review*, *87*(5), 1009–1018.

Björklund, A., Lindahl, M., & Plug, E. (2006). The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data. *The Quarterly Journal of Economics*, *121*(3), 999–1028. https://doi.org/10.1162/qjec.121.3.999

Black, S. E., & Devereux, P. J. (2010). *Recent Developments in IEM* (Working Paper No. 15889). National Bureau of Economic Research. https://doi.org/10.3386/w15889

Black, S. E., & Devereux, P. J. (2011). Recent Developments in IEM. In D. C. and O. Ashenfelter (Ed.), *Handbook of Labor Economics* (Vols. 4, Part B, pp. 1487–1541). Elsevier. http://www.sciencedirect.com/science/article/pii/S0169721811024142

Blane, D., Smith, G. D., & Hart, C. (1999). Some Social and Physical Correlates of Intergenerational Social Mobility: Evidence from the West of Scotland Collaborative Study. *Sociology*, *33*(1), 169–183. https://doi.org/10.1177/S0038038599000097

Borjas, G. J. (1985). Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics*, *3*(4), 463–489. https://doi.org/10.1086/298065

Borjas, G. J. (1992). Ethnic Capital and IEM. *The Quarterly Journal of Economics*, *107*(1), 123–150. https://doi.org/10.2307/2118325

Borooah, V. (2010). On the Risk Associated With Belonging to Disadvantaged Groups: A Bayesian Analysis with an Application to Labour Market Outcomes in India. In R. Basant & A. Shariff (Eds.), *Handbook of Muslims in India* (pp. 199–220). Oxford University Press. http://uir.ulster.ac.uk/9066/

Borooah, V. K. (2004). Gender bias among children in India in their diet and immunisation against disease. *Social Science & Medicine*, *58*(9), 1719–1731. https://doi.org/10.1016/S0277-9536(03)00342-3

Bound, J., & Solon, G. (1999). Double trouble: On the value of twins-based estimation of the return to schooling. *Economics of Education Review*, *18*(2), 169–182. https://doi.org/10.1016/S0272-7757(98)00048-X

Bourguignon, F., Ferreira, F. H. G., & Menendez, M. (2003). *Inequality of Outcomes and Inequality of Opportunities in Brazil* (SSRN Scholarly Paper ID 497242). Social Science Research Network. https://papers.ssrn.com/abstract=497242

Bowles, S., & Gintis, H. (2002). The Inheritance of Inequality. *Journal of Economic Perspectives*, *16*(3), 3–30. https://doi.org/10.1257/089533002760278686

Buis, M. L. (2011). The consequences of unobserved heterogeneity in a sequential logit model. *Research in Social Stratification and Mobility*, *29*(3), 247–262. https://doi.org/10.1016/j.rssm.2010.12.006

Burtless, G., & Jencks, C. (2003). *American inequality and its consequences* (Working Paper No. 339). LIS Working Paper Series. https://www.econstor.eu/handle/10419/95399

Cameron, S. V., & Heckman, J. J. (1998). Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. *Journal of Political Economy*, *106*(2), 262–333. https://doi.org/10.1086/250010

Causa, O., & Johansson, A. (2010). Intergenerational Social Mobility in OECD Countries. *OECD Journal: Economic Studies*, *2010*(1), 1–44. https://doi.org/10.1787/eco_studies-2010-5km33scz5rjj

Causa, O., & Johansson, Å. (2011). Intergenerational Social Mobility in OECD Countries. *OECD Journal: Economic Studies*, *2010*(1), 1–44. https://doi.org/10.1787/eco_studies-2010-5km33scz5rjj

Checchi, D., Fiorio, C. V., & Leonardi, M. (2013). Intergenerational persistence of educational attainment in Italy. *Economics Letters*, *118*(1), 229–232. https://doi.org/10.1016/j.econlet.2012.10.033

Checchi, D., & Flabbi, L. (2007). *IEM and Schooling Decisions in Germany and Italy: The Impact of Secondary School Tracks* (IZA Discussion Paper No. 2876). Institute for the Study of Labor (IZA). https://ideas.repec.org/p/iza/izadps/dp2876.html

Chen, H., & Wang, X. (2019). Exploring the relationship between rural village characteristics and Chinese return migrants' participation in farming: Path dependence in rural employment. *Cities*, *88*, 136–143. https://doi.org/10.1016/j.cities.2019.01.016

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of Opportunity? The Geography of IEM in the United States. *The Quarterly Journal of Economics*, *129*(4), 1553–1623. https://doi.org/10.1093/qje/qju022

Chusseau, N., Hellier, J., & Ben-Halima, B. (2013). Education, IEM and Inequality. In J. Hellier & N. Chusseau (Eds.), *Growing Income Inequalities: Economic Analyses* (pp. 227–273). Palgrave Macmillan UK. https://doi.org/10.1057/9781137283306_9

Coffey, D. (2013). Children's Welfare and Short-term Migration from Rural India. *The Journal of Development Studies*, *49*(8), 1101–1117. https://doi.org/10.1080/00220388.2013.794934

Coffey, D., Papp, J., & Spears, D. (2015). Short-Term Labor Migration from Rural North India: Evidence from New Survey Data. *Population Research and Policy Review*, *34*(3), 361–380. https://doi.org/10.1007/s11113-014-9349-2

Craig, L. (2006). Parental education, time in paid work and time with children: An Australian time-diary analysis. *The British Journal of Sociology*, *57*(4), 553–575. https://doi.org/10.1111/j.1468-4446.2006.00125.x

Currie, J., & Moretti, E. (2003). Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. *The Quarterly Journal of Economics*, *118*(4), 1495–1532. https://doi.org/10.1162/003355303322552856

DAS, P. (2012). Wage Inequality in India: Decomposition by Sector, Gender and Activity Status. *Economic and Political Weekly*, *47*(50), 58–64. JSTOR.

Daude, C. (2011). Ascendance by descendants? On IEM in Latin America. *OECD Development Centre Working Papers*, *297*, 1.

Dearden, L., Machin, S., & Reed, H. (1997). IEM in Britain. *The Economic Journal*, 47–66.

Démurger, S., & Xu, H. (2011). Return migrants: The rise of new entrepreneurs in rural China. *World Development*, *39*(10), 1847–1861.

Deshingkar, P. (2015). *Internal Migration, Poverty and Development in Asia*. 19.

Dickson, M., & Harmon, C. (2011). Economic returns to education: What We Know, What We Don't Know, and Where We Are Going—Some brief pointers. *Economics of Education Review*, *30*(6), 1118–1122. https://doi.org/10.1016/j.econedurev.2011.08.003

Djajić, S. (2014). Temporary Emigration and Welfare: The Case of Low-Skilled Labor. *International Economic Review*, *55*(2), 551–574. https://doi.org/10.1111/iere.12061

Dodd, W., Humphries, S., Patel, K., Majowicz, S., & Dewey, C. (2016). Determinants of temporary labour migration in southern India. *Asian Population Studies*, *12*(3), 294–311. https://doi.org/10.1080/17441730.2016.1207929

Dubey, A., Jones, R. P., & Sen, K. (2004). Surplus Labor. *Social Structure and Rural to Urban Migration: Evidence from Indian Data, School of Development Studies, University of East Anglia, Norwich*.

Duncan, B. (1967). Education and Social Background. *American Journal of Sociology*, *72*(4), 363–372. https://doi.org/10.1086/224335

Dunn, D. (1993). Gender inequality in education and employment in the scheduled castes and tribes of India. *Population Research and Policy Review*, *12*(1), 53–70. https://doi.org/10.1007/BF01074509

Duraisamy, P. (2002). Changes in returns to education in India, 1983–94: By gender, age-cohort and location. *Economics of Education Review*, *21*(6), 609–622. https://doi.org/10.1016/S0272-7757(01)00047-4

Durlauf, S. N. (1996). A theory of persistent income inequality. *Journal of Economic Growth*, *1*(1), 75–93. https://doi.org/10.1007/BF00163343

Dutta, P. V. (n.d.). *Accounting for Wage Inequality in India*. 31.

Dynarski, S. (2008). Building the Stock of College-Educated Labor. *Journal of Human Resources*, *43*(3), 576–610. https://doi.org/10.3368/jhr.43.3.576

Emran, M. S., Greene, W., Shilpi, F. J., & others. (2016). *When measure matters: Coresidency, truncation bias, and IEM in developing countries*. The World Bank. https://ideas.repec.org/p/wbk/wbrwps/7608.html

Emran, M. S., & Shilpi, F. (2015a). Gender, Geography, and Generations: IEM in Post-Reform India. *World Development*, *72*, 362–380. https://doi.org/10.1016/j.worlddev.2015.03.009

Emran, M. S., & Shilpi, F. (2015b). Gender, Geography, and Generations: IEM in Post-Reform India. *World Development*, *72*, 362–380. https://doi.org/10.1016/j.worlddev.2015.03.009

Erikson, R., Goldthorpe, J. H., Jackson, M., Yaish, M., & Cox, D. R. (2005). On class differentials in educational attainment. *Proceedings of the National Academy of Sciences*, *102*(27), 9730–9733.

Faggian, A., & Mccann, P. (2009). Universities, Agglomerations and Graduate Human Capital Mobility. *Tijdschrift Voor Economische En Sociale Geografie*, *100*(2), 210–223. https://doi.org/10.1111/j.1467-9663.2009.00530.x

Farré, L., & Vella, F. (2013). The Intergenerational Transmission of Gender Role Attitudes and its Implications for Female Labour Force Participation. *Economica*, *80*(318), 219–247. https://doi.org/10.1111/ecca.12008

Frölich, M. (2004). Finite-sample properties of propensity-score matching and weighting estimators. *Review of Economics and Statistics*, *86*(1), 77–90.

Gaviria, A. (2002). IEM, sibling inequality and borrowing constraints. *Economics of Education Review*, *21*(4), 331–340. https://doi.org/10.1016/S0272-7757(01)00031-0

Gibson, J., McKenzie, D., & Stillman, S. (2010). The Impacts of International Migration on Remaining Household Members: Omnibus Results from a Migration Lottery Program. *The Review of Economics and Statistics*, *93*(4), 1297–1318. https://doi.org/10.1162/REST_a_00129

Glaeser, E. L., & Resseger, M. G. (2010). The Complementarity Between Cities and Skills*. *Journal of Regional Science*, *50*(1), 221–244. https://doi.org/10.1111/j.1467-9787.2009.00635.x

González Rozada, M., & Menendez, A. (2002). Public university in Argentina: Subsidizing the rich? *Economics of Education Review*, *21*(4), 341–351. https://doi.org/10.1016/S0272-7757(01)00030-9

Grantham-McGregor, S. M., Walker, S. P., Chang, S. M., & Powell, C. A. (1997). Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children. *The American Journal of Clinical Nutrition*, *66*(2), 247–253. https://doi.org/10.1093/ajcn/66.2.247

Greenwald, R., Laine, R., & Hedges, L. (1994). A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, *23*(3), 5–14.

Griliches, Z. (1979). Sibling Models and Data in Economics: Beginnings of a Survey. *Journal of Political Economy*, *87*(5), S37–S64.

Groen, J. A. (2011). Building Knowledge Stocks Locally: Consequences of Geographic Mobility for the Effectiveness of State Higher Education Policies. *Economic Development Quarterly*, *25*(4), 316–329. https://doi.org/10.1177/0891242410388934

Gruber, L., & Kosack, S. (2014). The Tertiary Tilt: Education and Inequality in the Developing World. *World Development*, *54*, 253–272. https://doi.org/10.1016/j.worlddev.2013.08.002

Guryan, J., Hurst, E., & Kearney, M. (2008). Parental Education and Parental Time with Children. *The Journal of Economic Perspectives*, *22*(3), 23. https://doi.org/10.1257/089533008785539791

Haberfeld, Y., Menaria, R. K., Sahoo, B. B., & Vyas, R. N. (1999). Seasonal migration of rural labor in India. *Population Research and Policy Review*, *18*(5), 471–487.

Hanushek, E. A. (1996). Measuring Investment in Education. *Journal of Economic Perspectives*, *10*(4), 9–30. https://doi.org/10.1257/jep.10.4.9

Hanushek, E. A. (2001). Black-White Achievement Differences and Governmental Interventions. *American Economic Review*, *91*(2), 24–28. https://doi.org/10.1257/aer.91.2.24

Hanushek, E. A., Lavy, V., & Hitomi, K. (2008). Do Students Care about School Quality? Determinants of Dropout Behavior in Developing Countries. *Journal of Human Capital*, *2*(1), 69–105. https://doi.org/10.1086/529446

Harknett, K., Garfinkel, I., Bainbridge, J., Smeeding, T. M., McLanahan, S., & Folbre, N. (2003). *Do Public Expenditures Improve Child Outcomes in the U.S.? A Comparison Across Fifty States* (SSRN Scholarly Paper ID 396102). Social Science Research Network. https://papers.ssrn.com/abstract=396102

Hauser, R. M., & Featherman, D. L. (1976). Equality of Schooling: Trends and Prospects. *Sociology of Education*, *49*(2), 99–120. JSTOR. https://doi.org/10.2307/2112516

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, *65*(2), 261–294.

Heineck, G., & Riphahn, R. T. (2009). IEMal Attainment in Germany—The Last Five Decades. *Jahrbücher Für Nationalökonomie Und Statistik / Journal of Economics and Statistics*, *229*(1), 36–60.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., & Verashchagina, A. (2007). Intergenerational economic mobility around the world. *The BE Journal of Economic Analysis & Policy*, *7*(2), 1–46.

Hnatkovska, V., Lahiri, A., & Paul, S. B. (2013a). Breaking the Caste Barrier IEM in India. *Journal of Human Resources*, *48*(2), 435–473. https://doi.org/10.3368/jhr.48.2.435

Hnatkovska, V., Lahiri, A., & Paul, S. B. (2013b). Breaking the Caste Barrier IEM in India. *Journal of Human Resources*, *48*(2), 435–473. https://doi.org/10.3368/jhr.48.2.435

Holm, A., & Jaeger, M. M. (2011). Dealing with selection bias in educational transition models: The bivariate probit selection model. *Research in Social Stratification and Mobility*, *29*(3), 311–322.

Holmlund, H., Lindahl, M., & Plug, E. (2011a). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. *Journal of Economic Literature*, *49*(3), 615–651. https://doi.org/10.1257/jel.49.3.615

Holmlund, H., Lindahl, M., & Plug, E. (2011b). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. *Journal of Economic Literature*, *49*(3), 615–651. https://doi.org/10.1257/jel.49.3.615

Hu, B. Y., & Szente, J. (2010). Education of Young Chinese Migrant Children: Challenges and Prospects. *Early Childhood Education Journal*, *37*(6), 477–482. https://doi.org/10.1007/s10643-009-0362-8

Ilahi, N. (1999). Return migration and occupational change. *Review of Development Economics*, *3*(2), 170–186.

Iranzo, S., & Peri, G. (2009). Schooling Externalities, Technology, and Productivity: Theory and Evidence from U.S. States. *The Review of Economics and Statistics*, *91*(2), 420–431. https://doi.org/10.1162/rest.91.2.420

Jacobs, J. A. (1996). Gender Inequality and Higher Education. *Annual Review of Sociology*, *22*(1), 153–185. https://doi.org/10.1146/annurev.soc.22.1.153

Jacoby, B., & And Others. (1996). *Service-Learning in Higher Education: Concepts and Practices. The Jossey-Bass Higher and Adult Education Series*. Jossey-Bass Publishers, 350 Sansome St.

Jalan, J., & Murgai, R. (2008). *IEM in Education in India'. Paper presented at the 5th annual conference of the Indian Statistical Institute, Delhi, December*.

Jencks, C., Crouse, J., & Mueser, P. (1983). The Wisconsin Model of Status Attainment: A National Replication with Improved Measures of Ability and Aspiration. *Sociology of Education*, *56*(1), 3–19. https://doi.org/10.2307/2112300

Jerrim, J., & Macmillan, L. (2015). Income Inequality, IEM, and the Great Gatsby Curve: Is Education the Key? *Social Forces*, *94*(2), 505–533. https://doi.org/10.1093/sf/sov075

Judson, R. (1998). Economic Growth and Investment in Education: How Allocation Matters. *Journal of Economic Growth*, *3*(4), 337–359. https://doi.org/10.1023/A:1009763906365

Junge, V., Revilla Diez, J., & Schätzl, L. (2015). Determinants and Consequences of Internal Return Migration in Thailand and Vietnam. *World Development*, *71*, 94–106. https://doi.org/10.1016/j.worlddev.2013.11.007

Keshri, K., & Bhagat, R. B. (2013). Socioeconomic Determinants of Temporary Labour Migration in India. *Asian Population Studies*, *9*(2), 175–195. https://doi.org/10.1080/17441730.2013.797294

Kijima, Y. (2006). Why did wage inequality increase? Evidence from urban India 1983–99. *Journal of Development Economics*, *81*(1), 97–117. https://doi.org/10.1016/j.jdeveco.2005.04.008

Kingdon, G. G. (1998). Does the labour market explain lower female schooling in India? *The Journal of Development Studies*, *35*(1), 39–65. https://doi.org/10.1080/00220389808422554

Kingdon, G. G. (2007). The progress of school education in India. *Oxford Review of Economic Policy*, *23*(2), 168–195. https://doi.org/10.1093/oxrep/grm015

Kingdon, G. G., & Unni, J. (2001). Education and women’s labour market outcomes in India. *Education Economics*, *9*(2), 173–195.

Kishor, S. (1993). “May God Give Sons to All”: Gender and Child Mortality in India. *American Sociological Review*, *58*(2), 247–265. https://doi.org/10.2307/2095969

Kırdar, M. G. (2009). Labor market outcomes, savings accumulation, and return migration. *Labour Economics*, *16*(4), 418–428.

Kloosterman, R., Ruiter, S., Graaf, P. M. D., & Kraaykamp, G. (2009). Parental education, children’s performance and the transition to higher secondary education: Trends in primary and secondary effects over five Dutch school cohorts (1965–99). *The British Journal of Sociology*, *60*(2), 377–398. https://doi.org/10.1111/j.1468-4446.2009.01235.x

Kone, Z. L., Liu, M. Y., Mattoo, A., Ozden, C., & Sharma, S. (2018). Internal borders and migration in India. *Journal of Economic Geography*, *18*(4), 729–759.

Krishna, P., & Sethupathy, G. (2011). *Trade and Inequality in India* (Working Paper No. 17257; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w17257

Lall, M. (2005). *The challenges for India's education system*. Chatham House, Royal Institute of International Affairs. http://www.chathamhouse.org.uk/files/3233_bpindiaeducation.pdf

Levy, M. B. (1971). Determinants of Primary School Dropouts in Developing Countries. *Comparative Education Review*, *15*(1), 44–58. https://doi.org/10.1086/445512

Lewbel, A. (2012). Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models. *Journal of Business & Economic Statistics*, *30*(1), 67–80. https://doi.org/10.1080/07350015.2012.643126

Lillard, L. A., & Willis, R. J. (1994). IEM: Effects of Family and State in Malaysia. *The Journal of Human Resources*, *29*(4), 1126–1166. https://doi.org/10.2307/146136

Lindahl, M., Palme, M., Sandgren Massih, S., & Sjögren, A. (2012). *The Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations* (SSRN Scholarly Paper ID 2039649). Social Science Research Network. https://papers.ssrn.com/abstract=2039649

Lopreato, J., & Hazelrigg, L. E. (1970). Intragenerational Versus IEM in Relation to Sociopolitical Attitudes. *Social Forces*, *49*(2), 200–210. https://doi.org/10.1093/sf/49.2.200

Lucas, E. (2015). Making Sense of Migrants. *New Eastern Europe*, *19*(6), 7–11.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3–42. https://doi.org/10.1016/0304-3932(88)90168-7

Lucas, S. R. (2001). Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects. *American Journal of Sociology*, *106*(6), 1642–1690. https://doi.org/10.1086/321300

Lucas, S. R., Fucella, P. N., & Berends, M. (2011). A neo-classical education transitions approach: A corrected tale for three cohorts. *Research in Social Stratification and Mobility*, *3*(29), 263–285. https://doi.org/10.1016/j.rssm.2011.03.003

Maitra, P., & Sharma, A. (2009). Parents and children: Education across generations in India. *5th Annual Conference on Economic Growth and Development, ISI Delhi, Delhi*. http://www.isid.ac.in/~pu/conference/dec_09_conf/Papers/PushkarMaitra.pdf

Mare, R. D. (1980). Social Background and School Continuation Decisions. *Journal of the American Statistical Association*, *75*(370), 295–305. https://doi.org/10.1080/01621459.1980.10477466

Mare, R. D. (1993). *Educational Stratification on Observed and Unobserved Components of Familiy Background. Aus: Blossfeld, Hans-Peter; Shavit, Yossi (Eds.)(Hrsg.): Persistent Inequality. Changing Educational Attainment in Thirteen Countries*. Oxford et. alii (Westview Press).

Mare, R. D. (1997). Differential fertility, IEM, and racial inequality. *Social Science Research*, *26*(3), 263–291.

Massey, D. S. (1987). Understanding Mexican Migration to the United States. *American Journal of Sociology*, *92*(6), 1372–1403. https://doi.org/10.1086/228669

Maurin, E., & McNally, S. (2008). Vive la Révolution! Long-Term Educational Returns of 1968 to the Angry Students. *Journal of Labor Economics*, *26*(1), 1–33. https://doi.org/10.1086/522071

Mayer, S. E., & Lopoo, L. M. (2008). Government spending and IEM. *Journal of Public Economics*, *92*(1), 139–158. https://doi.org/10.1016/j.jpubeco.2007.04.003

Mood, C. (2010). Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It. *European Sociological Review*, *26*(1), 67–82. https://doi.org/10.1093/esr/jcp006

Morgan, S. L., & Winship, C. (2015). *Counterfactuals and causal inference*. Cambridge University Press.

Motiram, S., & Sarma, N. (2014). Polarization, Inequality, and Growth: The Indian Experience. *Oxford Development Studies*, *42*(3), 297–318. https://doi.org/10.1080/13600818.2014.897319

MOTIRAM, S., & SINGH, A. (2012). How Close Does the Apple Fall to the Tree? Some Evidence from India on Intergenerational Occupational Mobility. *Economic and Political Weekly*, *47*(40), 56–65. JSTOR.

Murthi, M., Guio, A.-C., & Drèze, J. (1995). Mortality, Fertility, and Gender Bias in India: A District-Level Analysis. *Population and Development Review*, *21*(4), 745–782. https://doi.org/10.2307/2137773

Nabi, I. (1984). Village-end considerations in rural-urban migration. *Journal of Development Economics*, *14*(1), 129–145. https://doi.org/10.1016/0304-3878(84)90046-4

Narayan, A., Weide, R. V. der, Cojocaru, A., Lakner, C., Redaelli, S., Mahler, D. G., Ramasubbaiah, R. G. N., & Thewissen, S. (2018). *Fair Progress?: Economic Mobility across Generations around the World*. World Bank Publications.

Nayyar, G., & Kim, K. Y. (2018). *India's internal labor migration paradox: The statistical and the real*. The World Bank.

Neckerman, K. M., & Torche, F. (2007). Inequality: Causes and Consequences. *Annual Review of Sociology*, *33*(1), 335–357. https://doi.org/10.1146/annurev.soc.33.040406.131755

Newbold, K. B. (2001). Counting migrants and migrations: Comparing lifetime and fixed-interval return and onward migration. *Economic Geography*, *77*(1), 23–40.

Nielsen, F. (2006). Achievement and Ascription in Educational Attainment: Genetic and Environmental Influences on Adolescent Schooling. *Social Forces*, *85*(1), 193–216. https://doi.org/10.1353/sof.2006.0135

Nimubona, A.-D., & Vencatachellum, D. (2007). IEM of black and white South Africans. *Journal of Population Economics*, *20*(1), 149–182. https://doi.org/10.1007/s00148-006-0120-9

O'Donovan, T. (1962). IEM. *Sociology and Social Research*, *47*(1), 57–67.

Olshansky, S. J., Antonucci, T., Berkman, L., Binstock, R. H., Boersch-Supan, A., Cacioppo, J. T., Carnes, B. A., Carstensen, L. L., Fried, L. P., Goldman, D. P., Jackson, J., Kohli, M., Rother, J., Zheng, Y., & Rowe, J. (2012). Differences In Life Expectancy Due To Race And Educational

Differences Are Widening, And Many May Not Catch Up. *Health Affairs*, *31*(8), 1803–1813. https://doi.org/10.1377/hlthaff.2011.0746

Oreopoulos, P., Page, M. E., & Stevens, A. H. (2006). The Intergenerational Effects of Compulsory Schooling. *Journal of Labor Economics*, *24*(4), 729–760. https://doi.org/10.1086/506484

Parida, J. K., & Madheswaran, S. (2011). *Determinants of migration and remittance in India: Empirical evidence*. Institute for Social and Economic Change.

Piketty, T. (2000). Chapter 8 Theories of persistent inequality and IEM. In *Handbook of Income Distribution* (Vol. 1, pp. 429–476). Elsevier. https://doi.org/10.1016/S1574-0056(00)80011-1

Piotrowski, M., & Tong, Y. (2010). Economic and non-economic determinants of return migration: Evidence from rural Thailand. *Population*, *65*(2), 333–348.

Platt, L. (2005). The Intergenerational Social Mobility of Minority Ethnic Groups. *Sociology*, *39*(3), 445–461. https://doi.org/10.1177/0038038505052494

Psacharopoulos, G. (1977). Unequal access to education and income distribution. *De Economist*, *125*(3), 383–392. https://doi.org/10.1007/BF01479118

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, *22*(9), 1325–1343. https://doi.org/10.1016/0305-750X(94)90007-8

Psacharopoulos, G., & Patrinos *, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, *12*(2), 111–134. https://doi.org/10.1080/0964529042000239140

Rauch, J. E. (1993). Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities. *Journal of Urban Economics*, *34*(3), 380–400. https://doi.org/10.1006/juec.1993.1042

Ray, R. (2002). The Determinants of Child Labour and Child Schooling in Ghana. *Journal of African Economies*, *11*(4), 561–590. https://doi.org/10.1093/jae/11.4.561

Reddy, A. B. (2015). Changes in Intergenerational Occupational Mobility in India: Evidence from National Sample Surveys, 1983–2012. *World Development*, *76*, 329–343. https://doi.org/10.1016/j.worlddev.2015.07.012

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688.

SARKAR, S., & MEHTA, B. S. (2010). Income Inequality in India: Pre- and Post-Reform Periods. *Economic and Political Weekly*, *45*(37), 45–55. JSTOR.

Schneebaum, A., Rumplmaier, B., & Altzinger, W. (2015). Gender and migration background in IEM. *Education Economics*, *0*(0), 1–23. https://doi.org/10.1080/09645292.2015.1006181

Schneeweis, N. (2011). Educational institutions and the integration of migrants. *Journal of Population Economics*, *24*(4), 1281–1308. https://doi.org/10.1007/s00148-009-0271-6

Schnepf, S. V. (2002). *A Sorting Hat that Fails? The Transition from Primary to Secondary School in Germany, UNICEF Innocenti Research Centre, Florenz*.

Sewell, W. H., & Hauser, R. M. (1975). *Education, Occupation, and Earnings. Achievement in the Early Career.* Academic Press. https://eric.ed.gov/?id=ED116041

Sharda, B. D. (1984). Return Mobility and the Process of Stratification in Rural India. *Journal of Asian and African Studies*, *19*(1–2), 64–72. https://doi.org/10.1163/156852184X00064

Sharma, K. (2017). India has 139 million internal migrants They must not be forgotten. *World Economic Forum*.

SIKDAR, S., & MUKHERJEE, A. N. (2012). Enrolment and Dropout Rate in School Education. *Economic and Political Weekly*, *47*(1), 27–31. JSTOR.

Sjoquist, D. L., & Winters, J. V. (2012). Building the Stock of College-Educated Labor Revisited. *Journal of Human Resources*, *47*(1), 270–285. https://doi.org/10.3368/jhr.47.1.270

Solon, G. (1992a). Intergenerational Income Mobility in the United States. *The American Economic Review*, *82*(3), 393–408.

Solon, G. (1992b). Intergenerational Income Mobility in the United States. *The American Economic Review*, *82*(3), 393–408. JSTOR.

Solon, G. (2002). Cross-Country Differences in Intergenerational Earnings Mobility. *Journal of Economic Perspectives*, *16*(3), 59–66. https://doi.org/10.1257/089533002760278712

Spady, W. G. (1967). Educational Mobility and Access: Growth and Paradoxes. *American Journal of Sociology*, *73*(3), 273–286. https://doi.org/10.1086/224475

Stasavage, D. (2005). Democracy and Education Spending in Africa. *American Journal of Political Science*, *49*(2), 343–358. https://doi.org/10.1111/j.0092-5853.2005.00127.x

Sundaram, A., & Vanneman, R. (2008). Gender Differentials in Literacy in India: The Intriguing Relationship with Women's Labor Force Participation. *World Development*, *36*(1), 128–143. https://doi.org/10.1016/j.worlddev.2007.02.017

Thanh, N. T., Lebailly, P., & Dien, N. T. (2019). Internal Return Migration in Rural of Vietnam: Reasons and Consequences. *Mediterranean Journal of Social Sciences*, *10*(1), 27.

Tilak, J. B. G. (2007). Post-elementary education, poverty and development in India. *International Journal of Educational Development*, *27*(4), 435–445. https://doi.org/10.1016/j.ijedudev.2006.09.018

Torul, O., & Öztunalı, O. (2018). On income and wealth inequality in Turkey. *Central Bank Review*, *18*(3), 95–106. https://doi.org/10.1016/j.cbrev.2018.06.002

Treiman, D. J., & Ganzeboom, H. B. G. (1990). Cross-national comparative status attainment research. *In: Research in Social Stratification and Mobility*, 105–127.

Trostel, P. A. (2010). The Fiscal Impacts of College Attainment. *Research in Higher Education*, *51*(3), 220–247. https://doi.org/10.1007/s11162-009-9156-5

Unni, J. (2010). Informality and Gender in the Labour Market for Muslims: Has Education been a Route out of Poverty? *Handbook of Muslims in India*, 221–234.

VAKULABHARANAM, V. (2010). Does Class Matter? Class Structure and Worsening Inequality in India. *Economic and Political Weekly*, *45*(29), 67–76. JSTOR.

Wang, W. W., & Fan, C. C. (2006). Success or Failure: Selectivity and Reasons of Return Migration in Sichuan and Anhui, China. *Environment and Planning A: Economy and Space*, *38*(5), 939–958. https://doi.org/10.1068/a37428

Ware, H. (1984). Effects of Maternal Education, Women's Roles, and Child Care on Child Mortality. *Population and Development Review*, *10*, 191–214. https://doi.org/10.2307/2807961

Yang, D. (2006). Why Do Migrants Return to Poor Countries? Evidence from Philippine Migrants' Responses to Exchange Rate Shocks. *The Review of Economics and Statistics*, *88*(4), 715–735. https://doi.org/10.1162/rest.88.4.715

Zhao, Y. (1999). Leaving the Countryside: Rural-to-Urban Migration Decisions in China. *American Economic Review*, *89*(2), 281–286. https://doi.org/10.1257/aer.89.2.281

Zhao, Y. (2002). Causes and Consequences of Return Migration: Recent Evidence from China. *Journal of Comparative Economics*, *30*(2), 376–394. https://doi.org/10.1006/jcec.2002.1781

Zhuang, J. (2010). *Poverty, Inequality, and Inclusive Growth in Asia: Measurement, Policy Issues, and Country Studies*. Anthem Press.

Zimmerman, D. J. (1992). Regression Toward Mediocrity in Economic Stature. *The American Economic Review*, *82*(3), 409–429.

Appendix A: ATT estimates and Rosenbaum's sensitivity analysis

Table A.1: Effect of high father education (threshold primary & above) on son's educational attainment using Mahalanobis matching estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.23	8.60	2.62	32.58***
	ATT		11.23	9.12	2.10	17.11***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)		T-Stat
Variable	Matched	Treated	Control		Bias reduction (%)	
Child age	U	28.54	28.41	4.90		1.98**
	M	28.54	28.51	1.10	78.00	0.49
Father age	U	59.05	60.06	-13.50		-5.56***
	M	59.05	59.00	0.60	95.60	0.3
Muslim	U	0.11	0.13	-8.70		-3.58***
	M	0.11	0.10	0.10	99.10	0.04
Other religion	U	0.07	0.05	9.70		3.85***
	M	0.07	0.07	0.00	100.00	0
OBC	U	0.40	0.42	-2.60		-1.05
	M	0.40	0.40	0.10	96.20	0.04
SC	U	0.15	0.24	-22.60		-9.34***
	M	0.15	0.15	-0.20	99.20	-0.09
ST	U	0.05	0.10	-22.60		-9.51***
	M	0.05	0.05	0.00	100.00	0
Others	U	0.01	0.02	-6.60		-2.75***
	M	0.01	0.01	0.00	100.00	0
Rural	U	0.51	0.74	-48.90		-19.52***
	M	0.51	0.52	-1.90	96.10	-0.81
Non metro	U	0.87	0.95	-25.70		-10.02***
	M	0.87	0.88	-0.70	97.30	-0.27

Table A.1 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.09	796.84***		16.50	13.50
Matched		0.00	0.96		0.40	0.20
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.01	2.01	2.01	2.53
1.20	0.00	0.00	2.01	2.53	1.48	2.53
1.40	0.00	0.00	1.48	2.99	1.48	2.99
1.60	0.00	0.00	1.02	2.99	1.02	2.99
1.80	0.00	0.00	1.02	3.52	1.02	3.52
2.00	0.00	0.00	1.02	3.52	0.49	3.52

Table A.2: Effect of high father education (threshold middle & above) on son's educational attainment using Mahalanobis matching estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.98	8.92	3.06	39.43***
	ATT		11.98	9.43	2.54	21.67***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.62	28.40	8.00		3.28***
	M	28.62	28.57	1.90	76.60	0.71
Father age	U	59.19	59.63	-6.00		-2.42**
	M	59.19	59.06	1.90	69.30	0.78
Muslim	U	0.09	0.13	-13.00		-5.21***
	M	0.09	0.09	0.10	99.10	0.05
Other religion	U	0.08	0.06	8.80		3.64***
	M	0.08	0.08	0.00	100.00	0
OBC	U	0.37	0.44	-13.20		-5.39***
	M	0.37	0.37	0.10	98.90	0.05
SC	U	0.13	0.22	-21.60		-8.66***
	M	0.13	0.14	-0.30	98.70	-0.12
ST	U	0.04	0.09	-20.60		-8.14***
	M	0.04	0.04	0.00	100.00	0
Others	U	0.01	0.02	-5.90		-2.34**
	M	0.01	0.01	0.00	100.00	0
Rural	U	0.46	0.71	-51.90		-21.32***
	M	0.46	0.46	-1.20	97.70	-0.43
Non metro	U	0.85	0.94	-28.20		-11.86***
	M	0.85	0.85	-0.30	98.80	-0.11

Table A.2 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	901.34***		17.80	13.20
Matched		0.00	1.1		0.60	0.30

Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.01	2.01	2.00	2.01
1.20	0.00	0.00	1.52	2.51	1.52	2.51
1.40	0.00	0.00	1.52	2.51	1.02	3.00
1.60	0.00	0.00	1.02	3.00	1.02	3.00
1.80	0.00	0.00	1.02	3.00	0.53	3.50
2.00	0.00	0.00	0.53	3.50	0.53	3.50

Table A.3: Effect of high father education (threshold secondary & above) on son's educational attainment using Mahalanobis matching estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		13.33	9.75	3.58	29.51***
	ATT		13.33	10.64	2.69	15.24***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction	T-Stat
Variable	Matched	Treated	Control		(%)	
Child age	U	28.90	28.43	17.30		4.68***
	M	28.90	28.86	1.50	91.10	0.32
Father age	U	59.70	59.41	4.10		1.04
	M	59.70	59.58	1.70	59.30	0.38
Muslim	U	0.09	0.12	-11.10		-2.86***
	M	0.09	0.09	0.00	100.00	0
Other religion	U	0.07	0.06	1.80		0.49
	M	0.07	0.07	0.00	100.00	0
OBC	U	0.31	0.42	-22.40		-5.95***
	M	0.31	0.31	0.00	100.00	0
SC	U	0.11	0.19	-22.50		-5.66***
	M	0.11	0.11	0.00	100.00	0
ST	U	0.04	0.07	-12.20		-3.06***
	M	0.04	0.04	0.00	100.00	0
Others	U	0.01	0.02	-5.30		-1.32
	M	0.01	0.01	0.00	100.00	0
Rural	U	0.32	0.64	-68.90		-18.55***
	M	0.32	0.32	0.00	100.00	0
Non metro	U	0.81	0.92	-31.90		-10.06***
	M	0.81	0.81	0.00	100.00	0

Table A.3 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	520.9***		20.20	17.30
Matched		0.00	0.18		0.30	0.00
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.02	2.02	1.52	2.02
1.20	0.00	0.00	1.52	2.02	1.52	2.52
1.40	0.00	0.00	1.02	2.52	1.02	2.52
1.60	0.00	0.00	1.02	2.52	0.53	3.01
1.80	0.00	0.00	0.53	3.01	0.53	3.01
2.00	0.00	0.00	0.53	3.01	0.53	3.51

Table A.4: Effect of high father education (threshold primary & above) on son's educational attainment using Nearest Neighbour PSM estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.23	8.60	2.62	32.58***
	ATT		11.23	9.07	2.16	17.37***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)		T-Stat
Variable	Matched	Treated	Control		Bias reduction (%)	
Child age	U	28.54	28.41	4.90		1.98**
	M	28.54	28.48	2.40	50.20	1.1
Father age	U	59.05	60.06	-13.50		-5.56***
	M	59.05	58.50	7.30	45.90	3.18***
Muslim	U	0.11	0.13	-8.70		-3.58***
	M	0.11	0.10	1.50	82.90	0.72
Other religion	U	0.07	0.05	9.70		3.85***
	M	0.07	0.08	-0.80	91.70	-0.33
OBC	U	0.40	0.42	-2.60		-1.05
	M	0.40	0.41	-1.70	33.60	-0.78
SC	U	0.15	0.24	-22.60		-9.34***
	M	0.15	0.15	-0.90	96.20	-0.43
ST	U	0.05	0.10	-22.60		-9.51***
	M	0.05	0.04	1.80	92.20	1.03
Others	U	0.01	0.02	-6.60		-2.75***
	M	0.01	0.01	2.40	64.40	1.33
Rural	U	0.51	0.74	-48.90		-19.52***
	M	0.51	0.53	-3.10	93.60	-1.34
Non metro	U	0.87	0.95	-25.70		-10.02***
	M	0.87	0.87	0.60	97.70	0.23

Table A.4 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.09	796.84***		16.50	13.50
Matched		0.00	17.84*		2.30	1.80

Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	1.99	1.99	1.99	2.53
1.20	0.00	0.00	1.99	2.53	1.52	2.53
1.40	0.00	0.00	1.52	3.00	1.52	3.00
1.60	0.00	0.00	0.98	3.00	0.98	3.47
1.80	0.00	0.00	0.98	3.47	0.98	3.47
2.00	0.00	0.00	0.98	3.47	0.51	4.01

Table A.5: Effect of high father education (threshold middle & above) on son's educational attainment using Nearest Neighbour PSM estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.98	8.92	3.06	39.43***
	ATT		11.98	9.47	2.51	20.48***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)		T-Stat
Variable	Matched	Treated	Control		Bias reduction (%)	
Child age	U	28.62	28.40	8.00		3.28***
	M	28.62	28.56	2.00	75.70	0.73
Father age	U	59.19	59.63	-6.00		-2.42**
	M	59.19	58.82	5.10	14.90	2.03**
Muslim	U	0.09	0.13	-13.00		-5.21***
	M	0.09	0.09	1.10	91.40	0.46
Other religion	U	0.08	0.06	8.80		3.64***
	M	0.08	0.06	8.50	3.80	3.18***
OBC	U	0.37	0.44	-13.20		-5.39***
	M	0.37	0.39	-4.80	63.60	-1.83*
SC	U	0.13	0.22	-21.60		-8.66***
	M	0.13	0.11	6.40	70.20	2.8***
ST	U	0.04	0.09	-20.60		-8.14***
	M	0.04	0.03	1.90	90.90	0.91
Others	U	0.01	0.02	-5.90		-2.34**
	M	0.01	0.01	-1.20	79.60	-0.5
Rural	U	0.46	0.71	-51.90		-21.32***
	M	0.46	0.47	-2.60	95.00	-0.93
Non metro	U	0.85	0.94	-28.20		-11.86***
	M	0.85	0.86	-2.40	91.40	-0.79

Table A.5 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	901.34***		17.80	13.20
Matched		0.00	25.5***		3.30	2.40
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.00	2.00	2.02	2.00
1.20	0.00	0.00	1.52	2.48	1.52	2.48
1.40	0.00	0.00	1.52	2.48	0.99	3.02
1.60	0.00	0.00	0.99	3.02	0.99	3.02
1.80	0.00	0.00	0.99	3.02	0.51	3.49
2.00	0.00	0.00	0.51	3.49	0.51	3.49

Table A.6: Effect of high father education (threshold secondary & above) on son's educational attainment using Nearest Neighbour estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		13.33	9.75	3.58	29.51***
	ATT		13.33	10.43	2.90	15.64***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)		T-Stat
Variable	Matched	Treated	Control		Bias reduction (%)	
Child age	U	28.90	28.43	17.30		4.68***
	M	28.90	28.82	3.00	82.60	0.61
Father age	U	59.70	59.41	4.10		1.04
	M	59.70	59.66	0.60	85.30	0.13
Muslim	U	0.09	0.12	-11.10		-2.86***
	M	0.09	0.07	4.70	57.70	1.08
Other religion	U	0.07	0.06	1.80		0.49
	M	0.07	0.06	2.40	-33.50	0.49
OBC	U	0.31	0.42	-22.40		-5.95***
	M	0.31	0.32	-2.00	91.10	-0.42
SC	U	0.11	0.19	-22.50		-5.66***
	M	0.11	0.11	-0.70	97.00	-0.15
ST	U	0.04	0.07	-12.20		-3.06***
	M	0.04	0.04	2.00	83.30	0.49
Others	U	0.01	0.02	-5.30		-1.32
	M	0.01	0.01	2.10	59.20	0.54
Rural	U	0.32	0.64	-68.90		-18.55***
	M	0.32	0.32	-0.30	99.60	-0.05
Non metro	U	0.81	0.92	-31.90		-10.06***
	M	0.81	0.81	-1.40	95.60	-0.25

Table A.6 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	520.9***		20.20	17.30
Matched		0.00	2.79		2.00	2.00

Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.01	2.01	1.51	2.01
1.20	0.00	0.00	1.51	2.01	1.51	2.50
1.40	0.00	0.00	1.02	2.50	1.02	2.50
1.60	0.00	0.00	1.02	2.50	0.52	3.00
1.80	0.00	0.00	0.52	3.00	0.52	3.00
2.00	0.00	0.00	0.52	3.00	0.03	3.50

Table A.7: Effect of high father education (threshold primary & above) on son's educational attainment using Epanechnikov kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.23	8.60	2.62	32.58***
	ATT		11.23	8.96	2.27	23.71***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.54	28.41	4.90		1.98**
	M	28.54	28.52	0.60	87.30	0.28
Father age	U	59.05	60.06	-13.50		-5.56***
	M	59.05	58.68	4.80	64.30	2.11**
Muslim	U	0.11	0.13	-8.70		-3.58***
	M	0.11	0.11	-1.80	79.70	-0.84
Other religion	U	0.07	0.05	9.70		3.85***
	M	0.07	0.07	1.30	86.40	0.56
OBC	U	0.40	0.42	-2.60		-1.05
	M	0.40	0.41	-1.70	33.50	-0.78
SC	U	0.15	0.24	-22.60		-9.34***
	M	0.15	0.15	-0.60	97.40	-0.29
ST	U	0.05	0.10	-22.60		-9.51***
	M	0.05	0.04	1.10	95.00	0.65
Others	U	0.01	0.02	-6.60		-2.75***
	M	0.01	0.01	-1.40	78.40	-0.73
Rural	U	0.51	0.74	-48.90		-19.52***
	M	0.51	0.53	-3.10	93.70	-1.32
Non metro	U	0.87	0.95	-25.70		-10.02***
	M	0.87	0.87	0.80	96.80	0.32

Table A.7 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.09	796.84***		16.50	13.50
Matched		0.00	13.78		2.00	1.40
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.38	2.38	2.24	2.45
1.20	0.00	0.00	2.02	2.59	1.95	2.73
1.40	0.00	0.00	1.88	2.87	1.74	2.94
1.60	0.00	0.00	1.60	3.02	1.45	3.16
1.80	0.00	0.00	1.45	3.23	1.38	3.30
2.00	0.00	0.00	1.31	3.37	1.17	3.44

Table A.8: Effect of high father education (threshold middle & above) on son's educational attainment using Epanechnikov kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.98	8.92	3.06	39.43***
	ATT		11.98	9.37	2.61	28.7***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.62	28.40	8.00		3.28***
	M	28.62	28.56	1.90	75.90	0.73
Father age	U	59.19	59.63	-6.00		-2.42**
	M	59.19	58.96	3.20	47.60	1.21
Muslim	U	0.09	0.13	-13.00		-5.21***
	M	0.09	0.09	-0.10	99.10	-0.05
Other religion	U	0.08	0.06	8.80		3.64***
	M	0.08	0.07	2.20	74.60	0.8
OBC	U	0.37	0.44	-13.20		-5.39***
	M	0.37	0.37	0.20	98.50	0.07
SC	U	0.13	0.22	-21.60		-8.66***
	M	0.13	0.13	2.00	90.70	0.85
ST	U	0.04	0.09	-20.60		-8.14***
	M	0.04	0.04	0.10	99.50	0.05
Others	U	0.01	0.02	-5.90		-2.34**
	M	0.01	0.01	-0.70	87.70	-0.31
Rural	U	0.46	0.71	-51.90		-21.32***
	M	0.46	0.47	-3.10	94.00	-1.12
Non metro	U	0.85	0.94	-28.20		-11.86***
	M	0.85	0.86	-2.50	91.10	-0.82

Table A.8 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	901.34***		17.80	13.20
Matched		0.00	5.95		1.70	2.00
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.08	2.08	2.01	2.21
1.20	0.00	0.00	1.82	2.40	1.70	2.46
1.40	0.00	0.00	1.63	2.59	1.50	2.72
1.60	0.00	0.00	1.44	2.78	1.31	2.91
1.80	0.00	0.00	1.25	2.97	1.12	3.04
2.00	0.00	0.00	1.06	3.10	0.99	3.17

Table A.9: Effect of high father education (threshold secondary & above) on son's educational attainment using Epanechnikov kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		13.33	9.75	3.58	29.51***
	ATT		13.33	10.46	2.87	25.79***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.90	28.43	17.30		4.68***
	M	28.90	28.84	2.00	88.30	0.41
Father age	U	59.70	59.41	4.10		1.04
	M	59.70	59.51	2.70	33.00	0.57
Muslim	U	0.09	0.12	-11.10		-2.86***
	M	0.09	0.09	-1.40	87.40	-0.31
Other religion	U	0.07	0.06	1.80		0.49
	M	0.07	0.07	0.90	47.80	0.19
OBC	U	0.31	0.42	-22.40		-5.95***
	M	0.31	0.31	0.60	97.40	0.12
SC	U	0.11	0.19	-22.50		-5.66***
	M	0.11	0.12	-1.80	92.20	-0.4
ST	U	0.04	0.07	-12.20		-3.06***
	M	0.04	0.04	0.40	97.00	
Others	U	0.01	0.02	-5.30		-1.32
	M	0.01	0.01	-0.60	88.60	-0.14
Rural	U	0.32	0.64	-68.90		-18.55***
	M	0.32	0.34	-3.70	94.60	-0.77
Non metro	U	0.81	0.92	-31.90		-10.06***
	M	0.81	0.81	0.40	98.70	0.07

Table A.9 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	520.9***		20.20	17.30
Matched		0.00	2.18		1.70	1.40

Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	1.89	1.89	1.83	2.01
1.20	0.00	0.00	1.66	2.18	1.54	2.30
1.40	0.00	0.00	1.43	2.42	1.31	2.48
1.60	0.00	0.00	1.25	2.59	1.14	2.65
1.80	0.00	0.00	1.08	2.77	0.96	2.83
2.00	0.00	0.00	0.96	2.88	0.84	3.00

Table A.10: Effect of high father education (threshold primary & above) on son's educational attainment using biweight kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.23	8.60	2.62	32.58***
	ATT		11.23	8.96	2.26	23.48***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.54	28.41	4.90		1.98**
	M	28.54	28.53	0.60	88.40	0.26
Father age	U	59.05	60.06	-13.50		-5.56***
	M	59.05	58.67	5.00	63.10	2.18**
Muslim	U	0.11	0.13	-8.70		-3.58***
	M	0.11	0.11	-1.50	82.30	-0.74
Other religion	U	0.07	0.05	9.70		3.85***
	M	0.07	0.07	1.00	89.70	0.42
OBC	U	0.40	0.42	-2.60		-1.05
	M	0.40	0.41	-1.40	46.00	-0.64
SC	U	0.15	0.24	-22.60		-9.34***
	M	0.15	0.15	-0.40	98.30	-0.19
ST	U	0.05	0.10	-22.60		-9.51***
	M	0.05	0.04	1.20	94.70	0.7
Others	U	0.01	0.02	-6.60		-2.75***
	M	0.01	0.01	-1.40	78.80	-0.71
Rural	U	0.51	0.74	-48.90		-19.52***
	M	0.51	0.53	-2.80	94.30	-1.18
Non metro	U	0.87	0.95	-25.70		-10.02***
	M	0.87	0.87	1.40	94.70	0.52

Table A.10 (continued)

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.09	796.84***		16.50	13.50
Matched		0.00	13.22		1.90	1.40

Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.37	2.37	2.23	2.44
1.20	0.00	0.00	2.01	2.58	1.94	2.72
1.40	0.00	0.00	1.87	2.86	1.73	2.93
1.60	0.00	0.00	1.59	3.00	1.45	3.15
1.80	0.00	0.00	1.45	3.22	1.38	3.36
2.00	0.00	0.00	1.31	3.36	1.17	3.43

Table A.11: Effect of high father education (threshold middle & above) on son's educational attainment using biweight kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.98	8.92	3.06	39.43***
	ATT		11.98	9.38	2.60	28.55***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.62	28.40	8.00		3.28***
	M	28.62	28.56	1.90	76.10	0.72
Father age	U	59.19	59.63	-6.00		-2.42**
	M	59.19	58.92	3.70	39.40	1.4
Muslim	U	0.09	0.13	-13.00		-5.21***
	M	0.09	0.09	-0.10	99.00	-0.05
Other religion	U	0.08	0.06	8.80		3.64***
	M	0.08	0.07	2.10	76.40	0.74
OBC	U	0.37	0.44	-13.20		-5.39***
	M	0.37	0.37	0.10	99.00	0.05
SC	U	0.13	0.22	-21.60		-8.66***
	M	0.13	0.13	2.10	90.20	0.9
ST	U	0.04	0.09	-20.60		-8.14***
	M	0.04	0.04	0.20	99.10	0.09
Others	U	0.01	0.02	-5.90		-2.34**
	M	0.01	0.01	-0.60	89.00	-0.28
Rural	U	0.46	0.71	-51.90		-21.32***
	M	0.46	0.47	-2.80	94.70	-0.99
Non metro	U	0.85	0.94	-28.20		-11.86***
	M	0.85	0.86	-2.50	91.10	-0.82

Table A.11 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	901.34***		17.80	13.20
Matched		0.00	5.88		1.60	2.00
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.07	2.07	2.01	2.20
1.20	0.00	0.00	1.82	2.39	1.69	2.52
1.40	0.00	0.00	1.56	2.58	1.50	2.71
1.60	0.00	0.00	1.37	2.77	1.31	2.90
1.80	0.00	0.00	1.24	2.96	1.12	3.03
2.00	0.00	0.00	1.05	3.09	0.99	3.16

Table A.12: Effect of high father education (threshold secondary & above) on son's educational attainment using biweight kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		13.33	9.75	3.58	29.51***
	ATT		13.33	10.46	2.86	25.61***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.90	28.43	17.30		4.68***
	M	28.90	28.85	1.60	90.50	0.33
Father age	U	59.70	59.41	4.10		1.04
	M	59.70	59.51	2.70	33.30	0.57
Muslim	U	0.09	0.12	-11.10		-2.86***
	M	0.09	0.09	-1.00	90.70	-0.23
Other religion	U	0.07	0.06	1.80		0.49
	M	0.07	0.07	1.00	42.80	0.21
OBC	U	0.31	0.42	-22.40		-5.95***
	M	0.31	0.31	0.80	96.40	0.17
SC	U	0.11	0.19	-22.50		-5.66***
	M	0.11	0.12	-1.90	91.40	-0.44
ST	U	0.04	0.07	-12.20		-3.06***
	M	0.04	0.04	0.80	93.80	0.18
Others	U	0.01	0.02	-5.30		-1.32
	M	0.01	0.01	-0.50	90.00	-0.12
Rural	U	0.32	0.64	-68.90		-18.55***
	M	0.32	0.33	-2.50	96.40	-0.51
Non metro	U	0.81	0.92	-31.90		-10.06***
	M	0.81	0.80	1.50	95.30	0.27

Table A.12 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	520.9***		20.20	17.30
Matched		0.00	1.81		1.70	1.50
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	1.89	1.89	1.83	2.00
1.20	0.00	0.00	1.65	2.18	1.54	2.29
1.40	0.00	0.00	1.42	2.41	1.31	2.47
1.60	0.00	0.00	1.25	2.58	1.13	2.70
1.80	0.00	0.00	1.07	2.76	0.96	2.82
2.00	0.00	0.00	0.96	2.87	0.84	2.99

Table A.13: Effect of high father education (threshold primary & above) on son's educational attainment using normal kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		13.33	9.75	3.58	29.51***
	ATT		13.33	10.46	2.86	25.61***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.90	28.43	17.30		4.68***
	M	28.90	28.85	1.60	90.50	0.33
Father age	U	59.70	59.41	4.10		1.04
	M	59.70	59.51	2.70	33.30	0.57
Muslim	U	0.09	0.12	-11.10		-2.86***
	M	0.09	0.09	-1.00	90.70	-0.23
Other religion	U	0.07	0.06	1.80		0.49
	M	0.07	0.07	1.00	42.80	0.21
OBC	U	0.31	0.42	-22.40		-5.95***
	M	0.31	0.31	0.80	96.40	0.17
SC	U	0.11	0.19	-22.50		-5.66***
	M	0.11	0.12	-1.90	91.40	-0.44
ST	U	0.04	0.07	-12.20		-3.06***
	M	0.04	0.04	0.80	93.80	0.18
Others	U	0.01	0.02	-5.30		-1.32
	M	0.01	0.01	-0.50	90.00	-0.12
Rural	U	0.32	0.64	-68.90		-18.55***
	M	0.32	0.33	-2.50	96.40	-0.51
Non metro	U	0.81	0.92	-31.90		-10.06***
	M	0.81	0.80	1.50	95.30	0.27

Table A.13 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	520.9***		20.20	17.30
Matched		0.00	1.81		1.70	1.50
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	1.89	1.89	1.83	2.00
1.20	0.00	0.00	1.65	2.18	1.54	2.29
1.40	0.00	0.00	1.42	2.41	1.31	2.47
1.60	0.00	0.00	1.25	2.58	1.13	2.70
1.80	0.00	0.00	1.07	2.76	0.96	2.82
2.00	0.00	0.00	0.96	2.87	0.84	2.99

Table A.14: Effect of high father education (threshold middle & above) on son's educational attainment using normal kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.98	8.92	3.06	39.43***
	ATT		11.98	9.34	2.64	30.08***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.62	28.40	8.00		3.28***
	M	28.62	28.56	2.20	72.70	0.82
Father age	U	59.19	59.63	-6.00		-2.42**
	M	59.19	59.02	2.30	62.10	0.87
Muslim	U	0.09	0.13	-13.00		-5.21***
	M	0.09	0.10	-1.40	89.10	-0.57
Other religion	U	0.08	0.06	8.80		3.64***
	M	0.08	0.07	2.10	76.30	0.74
OBC	U	0.37	0.44	-13.20		-5.39***
	M	0.37		0.38	-1.50	88.5***
SC	U	0.13	0.22	-21.60		-8.66***
	M	0.13	0.14	-0.80	96.10	-0.35
ST	U	0.04	0.09	-20.60		-8.14***
	M	0.04	0.04	-0.80	96.20	-0.36
Others	U	0.01	0.02	-5.90		-2.34**
	M	0.01	0.01	-1.30	77.70	-0.55
Rural	U	0.46	0.71	-51.90		-21.32***
	M	0.46	0.49	-6.20	88.10	-2.22**
Non metro	U	0.85	0.94	-28.20		-11.86***
	M	0.85	0.87	-5.00	82.10	-1.66*

Table A.14 (continued).

Panel C: Covariate Balance (Joint test)

Sample	Pseudo R2	LR chi2	Mean bias	Median bias
Unmatched	0.10	901.34***	17.80	13.20
Matched	0.00	11.61	2.50	2.10

Panel D: Rosenbaum sensitivity analysis

Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.11	2.11	2.05	2.24
1.20	0.00	0.00	1.85	2.44	1.72	2.57
1.40	0.00	0.00	1.66	2.63	1.53	2.76
1.60	0.00	0.00	1.46	2.83	1.33	2.96
1.80	0.00	0.00	1.27	3.02	1.14	3.09
2.00	0.00	0.00	1.14	3.15	1.01	3.22

Table A.15: Effect of high father education (threshold secondary & above) on son's educational attainment using normal kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		13.33	9.75	3.58	29.51***
	ATT		13.33	10.33	3.00	27.75***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.90	28.43	17.30		4.68***
	M	28.90	28.73	6.20	64.00	1.27
Father age	U	59.70	59.41	4.10		1.04
	M	59.70	59.47	3.30	18.80	0.69
Muslim	U	0.09	0.12	-11.10		-2.86***
	M	0.09	0.10	-3.90	64.60	-0.84
Other religion	U	0.07	0.06	1.80		0.49
	M	0.07	0.07	0.70	58.70	0.15
OBC	U	0.31	0.42	-22.40		-5.95***
	M	0.31	0.34	-4.30	80.70	-0.91
SC	U	0.11	0.19	-22.50		-5.66***
	M	0.11	0.13	-5.00	77.90	-1.12
ST	U	0.04	0.07	-12.20		-3.06***
	M	0.04	0.05	-2.50	79.60	-0.56
Others	U	0.01	0.02	-5.30		-1.32
	M	0.01	0.01	-1.50	71.50	-0.34
Rural	U	0.32	0.64	-68.90		-18.55***
	M	0.32	0.40	-17.30	74.90	-3.51***
Non metro	U	0.81	0.92	-31.90		-10.06***
	M	0.81	0.84	-8.50	73.50	-1.55

Table A.15 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	520.9***		20.20	17.30
Matched		0.01	19.29*		5.60	4.30
Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.00	2.00	1.88	2.06
1.20	0.00	0.00	1.70	2.18	1.59	2.30
1.40	0.00	0.00	1.53	2.48	1.41	2.54
1.60	0.00	0.00	1.29	2.66	1.17	2.72
1.80	0.00	0.00	1.11	2.78	1.05	2.90
2.00	0.00	0.00	0.99	2.96	0.87	3.02

Table A.16: Effect of high father education (threshold primary & above) on son's educational attainment using uniform kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.23	8.60	2.62	32.58***
	ATT		11.23	8.94	2.29	24.2***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.54	28.41	4.90		1.98**
	M	28.54	28.52	0.70	85.20	0.33
Father age	U	59.05	60.06	-13.50		-5.56***
	M	59.05	58.72	4.30	68.30	1.88*
Muslim	U	0.11	0.13	-8.70		-3.58***
	M	0.11	0.11	-2.20	74.50	-1.06
Other religion	U	0.07	0.05	9.70		3.85***
	M	0.07	0.07	1.60	83.60	0.68
OBC	U	0.40	0.42	-2.60		-1.05
	M	0.40	0.42	-2.50	4.20	-1.13
SC	U	0.15	0.24	-22.60		-9.34***
	M	0.15	0.15	-1.10	95.30	-0.53
ST	U	0.05	0.10	-22.60		-9.51***
	M	0.05	0.04	1.00	95.70	0.56
Others	U	0.01	0.02	-6.60		-2.75***
	M	0.01	0.01	-1.50	77.00	-0.77
Rural	U	0.51	0.74	-48.90		-19.52***
	M	0.51	0.53	-3.80	92.20	-1.63
Non metro	U	0.87	0.95	-25.70		-10.02***
	M	0.87	0.87	-0.50	98.20	-0.18

Table A.16 (continued).

Panel C: Covariate Balance (Joint test)					
Sample	Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched	0.09	796.84***		16.50	13.50
Matched	0.00	15.59		2.20	1.60

Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.32	2.32	2.25	2.47
1.20	0.00	0.00	2.04	2.68	1.97	2.75
1.40	0.00	0.00	1.82	2.90	1.75	2.97
1.60	0.00	0.00	1.61	3.04	1.54	3.18
1.80	0.00	0.00	1.47	3.25	1.32	3.33
2.00	0.00	0.00	1.32	3.40	1.18	3.47

Table A.17: Effect of high father education (threshold middle & above) on son's educational attainment using uniform kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		11.98	8.92	3.06	39.43***
	ATT		11.98	9.36	2.62	28.93***
Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.62	28.40	8.00		3.28***
	M	28.62	28.56	1.90	75.90	0.73
Father age	U	59.19	59.63	-6.00		-2.42**
	M	59.19	59.03	2.20	63.80	0.83
Muslim	U	0.09	0.13	-13.00		-5.21***
	M	0.09	0.09	-0.30	97.50	-0.13
Other religion	U	0.08	0.06	8.80		3.64***
	M	0.08	0.07	2.30	73.70	0.83
OBC	U	0.37	0.44	-13.20		-5.39***
	M	0.37	0.37	0.10	99.40	0.03
SC	U	0.13	0.22	-21.60		-8.66***
	M	0.13	0.13	1.60	92.70	0.66
ST	U	0.04	0.09	-20.60		-8.14***
	M	0.04	0.04	-0.10	99.60	-0.04
Others	U	0.01	0.02	-5.90		-2.34**
	M	0.01	0.01	-1.00	83.60	-0.41
Rural	U	0.46	0.71	-51.90		-21.32***
	M	0.46	0.48	-3.70	92.90	-1.32
Non metro	U	0.85	0.94	-28.20		-11.86***
	M	0.85	0.86	-2.30	91.70	-0.76

Table A.17 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	901.34***		17.80	13.20
Matched		0.00	6.17		1.70	1.90

Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	2.09	2.09	2.03	2.22
1.20	0.00	0.00	1.83	2.41	1.70	2.54
1.40	0.00	0.00	1.58	2.60	1.51	2.73
1.60	0.00	0.00	1.45	2.80	1.32	2.93
1.80	0.00	0.00	1.25	2.99	1.13	3.05
2.00	0.00	0.00	1.06	3.12	1.00	3.18

Table A.18 : Effect of high father education (threshold secondary & above) on son's educational attainment using uniform kernel estimator

Panel A: ATT Estimate						
Variable	Sample		Treated	Controls	Difference	T-Stat
Child education	Unmatched		13.33	9.75	3.58	29.51***
	ATT		13.33	10.45	2.88	26.05***

Panel B: Covariate Balance						
	Unmatched	Mean		Bias (%)	Bias reduction (%)	T-Stat
Variable	Matched	Treated	Control			
Child age	U	28.90	28.43	17.30		4.68***
	M	28.90	28.82	2.70	84.40	0.55
Father age	U	59.70	59.41	4.10		1.04
	M	59.70	59.51	2.70	33.10	0.57
Muslim	U	0.09	0.12	-11.10		-2.86***
	M	0.09	0.09	-2.40	78.00	-0.53
Other religion	U	0.07	0.06	1.80		0.49
	M	0.07	0.07	0.60	66.90	0.12
OBC	U	0.31	0.42	-22.40		-5.95***
	M	0.31	0.31	0.10	99.40	0.03
SC	U	0.11	0.19	-22.50		-5.66***
	M	0.11	0.12	-1.50	93.40	-0.34
ST	U	0.04	0.07	-12.20		-3.06***
	M	0.04	0.04	-0.40	97.10	-0.08
Others	U	0.01	0.02	-5.30		-1.32
	M	0.01	0.01	-0.70	85.80	-0.17
Rural	U	0.32	0.64	-68.90		-18.55***
	M	0.32	0.35	-6.30	90.80	-1.3
Non metro	U	0.81	0.92	-31.90		-10.06***
	M	0.81	0.81	-2.00	93.80	-0.36

Table A.18 (continued).

Panel C: Covariate Balance (Joint test)						
Sample		Pseudo R2	LR chi2		Mean bias	Median bias
Unmatched		0.10	520.9***		20.20	17.30
Matched		0.00	3.61		2.30	2.00

Panel D: Rosenbaum sensitivity analysis						
Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1.00	0.00	0.00	1.96	1.96	1.84	2.02
1.20	0.00	0.00	1.67	2.20	1.55	2.25
1.40	0.00	0.00	1.43	2.43	1.38	2.49
1.60	0.00	0.00	1.26	2.61	1.14	2.66
1.80	0.00	0.00	1.08	2.72	1.02	2.84
2.00	0.00	0.00	0.97	2.90	0.85	3.02

Appendix B: Sensitivity analysis for sequential logit estimates

Table B.1: Non-literate to primary transition for sons: Sensitivity analysis for odds ratio estimates

Effect of unobserved heterogeneity βu (LnOdds)	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Father Education	1.066***	1.069***	1.077***	1.087***	1.099***	1.113***	1.127***	1.142***	1.157***	1.174***	1.190***
	(0.009)	(0.009)	(0.010)	(0.011)	(0.013)	(0.015)	(0.017)	(0.019)	(0.021)	(0.023)	(0.026)
Migrant	0.829	0.823	0.806	0.777	0.741	0.700	0.658	0.616	0.575	0.534	0.496
	(0.119)	(0.123)	(0.131)	(0.142)	(0.154)	(0.166)	(0.176)	(0.184)	(0.190)	(0.195)	(0.198)
Father Education × Migrant	1.006	1.007	1.009	1.013	1.017	1.021	1.026	1.030	1.035	1.039	1.043
	(0.025)	(0.025)	(0.028)	(0.032)	(0.036)	(0.041)	(0.047)	(0.053)	(0.059)	(0.065)	(0.072)
Mother Education	1.027**	1.027**	1.030**	1.034**	1.038**	1.042**	1.047**	1.050*	1.054*	1.058*	1.061*
	(0.010)	(0.010)	(0.011)	(0.012)	(0.014)	(0.016)	(0.018)	(0.021)	(0.023)	(0.025)	(0.028)
Remittances	0.974*	0.974*	0.972*	0.969*	0.966*	0.961*	0.957*	0.951*	0.946*	0.941*	0.935*
	(0.011)	(0.011)	(0.012)	(0.014)	(0.015)	(0.017)	(0.019)	(0.022)	(0.024)	(0.026)	(0.028)
Child Age	1.141***	1.145***	1.159***	1.180***	1.207***	1.238***	1.273***	1.311***	1.351***	1.393***	1.437***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.007)	(0.008)	(0.010)	(0.011)	(0.012)	(0.014)
Child Age^2	1.000***	1.000***	1.000***	1.000***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.988	0.988	0.986	0.985	0.983	0.981	0.979	0.977	0.975	0.973	0.971
	(0.009)	(0.009)	(0.010)	(0.011)	(0.013)	(0.014)	(0.016)	(0.018)	(0.020)	(0.022)	(0.024)
Household Assets	1.063***	1.065***	1.071***	1.080***	1.091***	1.102***	1.115***	1.129***	1.143***	1.158***	1.173***
	(0.007)	(0.007)	(0.008)	(0.009)	(0.010)	(0.011)	(0.013)	(0.015)	(0.016)	(0.018)	(0.020)
Non-Hindu	0.615***	0.606***	0.582***	0.547***	0.508***	0.467***	0.426***	0.388***	0.352***	0.318***	0.287***
	(0.041)	(0.042)	(0.044)	(0.046)	(0.048)	(0.051)	(0.052)	(0.053)	(0.053)	(0.053)	(0.052)

Table B.1 (continued).

OBC	0.847*	0.844*	0.837*	0.826*	0.813	0.798	0.781	0.764	0.746	0.728	0.710
	(0.064)	(0.066)	(0.071)	(0.078)	(0.087)	(0.097)	(0.107)	(0.116)	(0.126)	(0.135)	(0.144)
SC & ST	0.950	0.949	0.945	0.942	0.941	0.940	0.937	0.934	0.930	0.925	0.920
	(0.079)	(0.081)	(0.088)	(0.098)	(0.111)	(0.125)	(0.140)	(0.156)	(0.172)	(0.188)	(0.203)
Urban	1.470***	1.484***	1.524***	1.592***	1.687***	1.804***	1.941***	2.097***	2.270***	2.463***	2.675***
	(0.102)	(0.106)	(0.118)	(0.139)	(0.166)	(0.202)	(0.244)	(0.294)	(0.353)	(0.421)	(0.498)
Children under 6	0.754***	0.746***	0.723***	0.690***	0.653***	0.615***	0.577***	0.540***	0.505***	0.472***	0.442***
	(0.045)	(0.046)	(0.049)	(0.052)	(0.056)	(0.060)	(0.064)	(0.067)	(0.069)	(0.071)	(0.073)
6-14 year old	0.651***	0.653***	0.652***	0.642**	0.625**	0.605**	0.583**	0.561**	0.540*	0.519*	0.498*
	(0.077)	(0.079)	(0.083)	(0.090)	(0.098)	(0.106)	(0.114)	(0.122)	(0.129)	(0.136)	(0.141)
Ln (Debt)	1.004	1.005	1.006	1.007	1.008	1.009	1.010	1.012	1.013	1.015	1.016
	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.010)	(0.011)	(0.012)	(0.013)	(0.015)
Only Sons	0.847*	0.839*	0.821**	0.798**	0.774**	0.748**	0.722**	0.695**	0.669**	0.642**	0.616**
	(0.057)	(0.058)	(0.062)	(0.067)	(0.074)	(0.080)	(0.087)	(0.094)	(0.100)	(0.105)	(0.110)
Siblings	0.824***	0.818***	0.801***	0.777***	0.750***	0.720***	0.689***	0.657***	0.626***	0.595***	0.565***
	(0.018)	(0.018)	(0.020)	(0.021)	(0.024)	(0.026)	(0.028)	(0.030)	(0.031)	(0.033)	(0.034)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and it is uncorrelated with any of the observed explanatory variables.

Table B.2: Primary to secondary transition for Sons: Sensitivity analysis for odds ratio estimates

Effect of unobserved heterogeneity βu (LnOdds)	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Father Education	1.044***	1.047***	1.054***	1.064***	1.076***	1.090***	1.104***	1.119***	1.135***	1.151***	1.167***
	(0.007)	(0.007)	(0.008)	(0.009)	(0.010)	(0.012)	(0.013)	(0.015)	(0.017)	(0.019)	(0.021)
Migrant	0.718*	0.705*	0.675*	0.637*	0.596*	0.552*	0.508*	0.465*	0.423*	0.383*	0.345*
	(0.111)	(0.113)	(0.120)	(0.128)	(0.135)	(0.141)	(0.145)	(0.147)	(0.148)	(0.147)	(0.144)
Father Education × Migrant	1.057*	1.060*	1.066*	1.076*	1.088*	1.100*	1.114*	1.129*	1.144*	1.160*	1.177*
	(0.026)	(0.028)	(0.031)	(0.035)	(0.040)	(0.046)	(0.052)	(0.059)	(0.066)	(0.074)	(0.082)
Mother Education	1.046***	1.048***	1.053***	1.060***	1.068***	1.077***	1.087***	1.096***	1.106***	1.117***	1.127***
	(0.008)	(0.008)	(0.009)	(0.011)	(0.012)	(0.014)	(0.016)	(0.018)	(0.020)	(0.022)	(0.024)
Remittances	0.971**	0.970**	0.967**	0.963**	0.958**	0.953**	0.947**	0.942**	0.936**	0.931**	0.925**
	(0.009)	(0.010)	(0.011)	(0.012)	(0.013)	(0.015)	(0.017)	(0.019)	(0.020)	(0.022)	(0.024)
Child Age	1.384***	1.403***	1.458***	1.538***	1.638***	1.755***	1.888***	2.037***	2.202***	2.382***	2.579***
	(0.010)	(0.010)	(0.011)	(0.013)	(0.016)	(0.019)	(0.023)	(0.027)	(0.032)	(0.038)	(0.045)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.998***	0.998***	0.998***	0.998***	0.998***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.993	0.992	0.991	0.990	0.988	0.985	0.983	0.981	0.979	0.978	0.976
	(0.009)	(0.010)	(0.011)	(0.012)	(0.014)	(0.016)	(0.017)	(0.019)	(0.021)	(0.023)	(0.025)
Household Assets	1.088***	1.092***	1.103***	1.119***	1.139***	1.160***	1.184***	1.208***	1.234***	1.260***	1.288***
	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.010)	(0.012)	(0.013)	(0.015)	(0.017)	(0.019)
Non-Hindu	0.586***	0.570***	0.530***	0.477***	0.422***	0.370***	0.321***	0.278***	0.239***	0.206***	0.176***
	(0.035)	(0.035)	(0.036)	(0.036)	(0.036)	(0.036)	(0.035)	(0.034)	(0.032)	(0.030)	(0.028)
OBC	1.151*	1.156*	1.170*	1.188*	1.207*	1.227*	1.247	1.266	1.285	1.304	1.321
	(0.071)	(0.074)	(0.083)	(0.095)	(0.109)	(0.125)	(0.143)	(0.161)	(0.181)	(0.201)	(0.222)
SC & ST	1.021	1.018	1.014	1.011	1.008	1.004	0.998	0.992	0.984	0.976	0.967
	(0.068)	(0.071)	(0.078)	(0.087)	(0.099)	(0.111)	(0.124)	(0.137)	(0.150)	(0.163)	(0.176)

Table B.2 (continued).

Urban	1.655***	1.691***	1.792***	1.946***	2.143***	2.379***	2.655***	2.974***	3.341***	3.761***	4.240***
	(0.097)	(0.103)	(0.119)	(0.146)	(0.182)	(0.228)	(0.285)	(0.355)	(0.440)	(0.543)	(0.667)
Children under 6	0.712***	0.701***	0.671***	0.631***	0.588***	0.546***	0.505***	0.466***	0.430***	0.396***	0.365***
	(0.044)	(0.045)	(0.048)	(0.051)	(0.054)	(0.056)	(0.058)	(0.059)	(0.060)	(0.061)	(0.061)
6-14 year old	0.804*	0.797**	0.776**	0.742**	0.703**	0.663**	0.625**	0.587***	0.551***	0.517***	0.484***
	(0.069)	(0.070)	(0.073)	(0.078)	(0.082)	(0.087)	(0.091)	(0.095)	(0.098)	(0.100)	(0.102)
Ln (Debt)	1.005	1.005	1.005	1.006	1.008	1.009	1.010	1.012	1.014	1.015	1.017
	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)	(0.007)	(0.008)	(0.009)	(0.010)	(0.011)	(0.012)
Only Sons	0.778***	0.768***	0.741***	0.705***	0.666***	0.627***	0.589***	0.552***	0.517***	0.483***	0.452***
	(0.043)	(0.044)	(0.047)	(0.050)	(0.054)	(0.057)	(0.060)	(0.063)	(0.065)	(0.066)	(0.068)
Siblings	0.819***	0.810***	0.788***	0.758***	0.724***	0.690***	0.656***	0.623***	0.591***	0.560***	0.531***
	(0.018)	(0.018)	(0.019)	(0.021)	(0.023)	(0.024)	(0.026)	(0.027)	(0.028)	(0.030)	(0.030)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and it is uncorrelated with any of the observed explanatory variables.

Table B.3: Secondary to post-secondary transition for Sons: Sensitivity analysis for odds ratio estimates

Effect of unobserved heterogeneity βu (LnOdds)	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Father Education	1.084***	1.090***	1.103***	1.121***	1.142***	1.165***	1.189***	1.215***	1.242***	1.270***	1.300***
	(0.008)	(0.009)	(0.010)	(0.012)	(0.013)	(0.015)	(0.018)	(0.020)	(0.023)	(0.026)	(0.029)
Migrant	1.299	1.292	1.284	1.287	1.299	1.318	1.342	1.369	1.399	1.430	1.461
	(0.354)	(0.368)	(0.403)	(0.455)	(0.518)	(0.593)	(0.676)	(0.769)	(0.869)	(0.977)	(1.091)
Father Education × Migrant	0.903*	0.901*	0.895*	0.887*	0.876*	0.864*	0.852*	0.839*	0.826*	0.813*	0.800*
	(0.036)	(0.038)	(0.041)	(0.046)	(0.052)	(0.058)	(0.064)	(0.070)	(0.076)	(0.082)	(0.088)
Mother Education	1.060***	1.063***	1.071***	1.083***	1.096***	1.111***	1.127***	1.143***	1.160***	1.177***	1.195***
	(0.009)	(0.010)	(0.011)	(0.012)	(0.014)	(0.016)	(0.019)	(0.021)	(0.024)	(0.027)	(0.030)
Remittances	1.004	1.003	1.002	1.000	0.998	0.996	0.994	0.992	0.990	0.988	0.986
	(0.012)	(0.013)	(0.014)	(0.016)	(0.018)	(0.020)	(0.023)	(0.025)	(0.028)	(0.031)	(0.034)
Child Age	2.135***	2.209***	2.415***	2.728***	3.146***	3.676***	4.331***	5.126***	6.079***	7.211***	8.547***
	(0.082)	(0.088)	(0.106)	(0.135)	(0.174)	(0.228)	(0.298)	(0.389)	(0.506)	(0.653)	(0.839)
Child Age^2	0.998***	0.998***	0.998***	0.998***	0.997***	0.997***	0.997***	0.996***	0.996***	0.996***	0.995***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	1.005	1.005	1.005	1.005	1.006	1.006	1.007	1.007	1.008	1.009	1.010
	(0.014)	(0.014)	(0.016)	(0.018)	(0.021)	(0.024)	(0.027)	(0.030)	(0.033)	(0.036)	(0.040)
Household Assets	1.093***	1.100***	1.116***	1.139***	1.164***	1.192***	1.223***	1.255***	1.289***	1.324***	1.361***
	(0.007)	(0.008)	(0.009)	(0.010)	(0.012)	(0.014)	(0.016)	(0.018)	(0.021)	(0.023)	(0.026)
Non-Hindu	0.829*	0.811**	0.768**	0.716***	0.664***	0.614***	0.568***	0.524***	0.485***	0.448***	0.415***
	(0.062)	(0.064)	(0.067)	(0.071)	(0.075)	(0.078)	(0.081)	(0.083)	(0.085)	(0.087)	(0.087)
OBC	1.064	1.068	1.079	1.094	1.111	1.129	1.149	1.168	1.187	1.205	1.222
	(0.076)	(0.080)	(0.090)	(0.104)	(0.120)	(0.139)	(0.159)	(0.180)	(0.203)	(0.227)	(0.251)
SC & ST	0.874	0.871	0.862	0.847	0.830	0.810	0.788	0.766	0.743	0.720	0.698
	(0.070)	(0.073)	(0.080)	(0.090)	(0.100)	(0.111)	(0.122)	(0.132)	(0.142)	(0.152)	(0.160)

Table B.3 (continued).

Urban	1.251***	1.276***	1.340***	1.427***	1.529***	1.644***	1.772***	1.913***	2.068***	2.239***	2.426***
	(0.084)	(0.090)	(0.105)	(0.127)	(0.155)	(0.189)	(0.230)	(0.277)	(0.331)	(0.395)	(0.467)
Children under 6	0.648***	0.627***	0.581***	0.528***	0.475***	0.425***	0.380***	0.339***	0.302***	0.270***	0.241***
	(0.060)	(0.061)	(0.063)	(0.065)	(0.066)	(0.067)	(0.067)	(0.067)	(0.066)	(0.065)	(0.063)
6-14 year old	0.886	0.875	0.851*	0.824*	0.796*	0.768*	0.741*	0.714*	0.687*	0.661*	0.636*
	(0.058)	(0.060)	(0.065)	(0.072)	(0.079)	(0.087)	(0.095)	(0.102)	(0.109)	(0.116)	(0.122)
Ln (Debt)	1.021***	1.022***	1.025***	1.028***	1.032***	1.037***	1.042***	1.047***	1.053***	1.058***	1.064***
	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)	(0.010)	(0.011)	(0.012)	(0.014)	(0.015)	(0.016)
Only Sons	0.875*	0.867*	0.845*	0.816*	0.786*	0.754*	0.722*	0.689**	0.657**	0.626**	0.595**
	(0.059)	(0.061)	(0.066)	(0.073)	(0.080)	(0.087)	(0.093)	(0.100)	(0.105)	(0.110)	(0.115)
Siblings	0.907**	0.898***	0.876***	0.850***	0.822***	0.794***	0.767***	0.740***	0.714***	0.689***	0.665***
	(0.028)	(0.029)	(0.031)	(0.034)	(0.038)	(0.041)	(0.045)	(0.048)	(0.051)	(0.054)	(0.057)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and it is uncorrelated with any of the observed explanatory variables.

Table B.4: Non-literate to primary transition for daughters: Sensitivity analysis for odds ratio estimates

Effect of unobserved heterogeneity βu (LnOdds)	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Father Education	1.073***	1.077***	1.085***	1.097***	1.111***	1.126***	1.142***	1.159***	1.176***	1.194***	1.212***
	(0.009)	(0.009)	(0.010)	(0.012)	(0.013)	(0.015)	(0.018)	(0.020)	(0.022)	(0.025)	(0.028)
Migrant	0.834	0.832	0.822	0.808	0.791	0.771	0.747	0.721	0.693	0.664	0.635
	(0.125)	(0.128)	(0.138)	(0.153)	(0.171)	(0.189)	(0.207)	(0.224)	(0.239)	(0.251)	(0.262)
Father Education × Migrant	1.038	1.041	1.047	1.056	1.066	1.078	1.090	1.103	1.116	1.129	1.144
	(0.030)	(0.031)	(0.034)	(0.038)	(0.044)	(0.051)	(0.058)	(0.066)	(0.075)	(0.083)	(0.092)
Mother Education	1.047***	1.048***	1.051***	1.056***	1.062***	1.068***	1.075***	1.083***	1.090***	1.098***	1.106***
	(0.010)	(0.011)	(0.012)	(0.013)	(0.015)	(0.017)	(0.020)	(0.022)	(0.025)	(0.028)	(0.030)
Remittances	0.970**	0.969**	0.966**	0.962**	0.956**	0.949**	0.942**	0.934**	0.926**	0.918**	0.910**
	(0.011)	(0.011)	(0.012)	(0.014)	(0.015)	(0.017)	(0.019)	(0.021)	(0.023)	(0.025)	(0.027)
Child Age	1.130***	1.135***	1.148***	1.168***	1.193***	1.223***	1.256***	1.292***	1.331***	1.371***	1.413***
	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.008)	(0.009)	(0.010)	(0.012)	(0.013)	(0.015)
Child Age^2	1.000***	1.000***	1.000***	1.000***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.997	0.996	0.995	0.994	0.992	0.990	0.987	0.985	0.982	0.980	0.978
	(0.009)	(0.010)	(0.011)	(0.012)	(0.013)	(0.015)	(0.017)	(0.019)	(0.020)	(0.022)	(0.023)
Household Assets	1.075***	1.078***	1.085***	1.096***	1.109***	1.124***	1.140***	1.158***	1.176***	1.196***	1.215***
	(0.007)	(0.007)	(0.008)	(0.009)	(0.011)	(0.012)	(0.014)	(0.016)	(0.018)	(0.020)	(0.022)
Non-Hindu	0.640***	0.631***	0.609***	0.577***	0.540***	0.501***	0.462***	0.425***	0.389***	0.356***	0.325***
	(0.045)	(0.046)	(0.048)	(0.051)	(0.054)	(0.057)	(0.059)	(0.061)	(0.062)	(0.062)	(0.062)
OBC	0.696***	0.691***	0.675***	0.650***	0.617***	0.582***	0.545***	0.510***	0.475***	0.443***	0.412***
	(0.058)	(0.059)	(0.062)	(0.067)	(0.072)	(0.076)	(0.081)	(0.084)	(0.087)	(0.089)	(0.090)
SC & ST	0.838	0.836	0.828	0.815	0.797	0.777	0.756	0.736	0.715	0.695	0.674
	(0.076)	(0.078)	(0.084)	(0.092)	(0.102)	(0.113)	(0.123)	(0.134)	(0.144)	(0.154)	(0.163)

Table B.4 (continued).

Urban	1.486***	1.504***	1.559***	1.649***	1.771***	1.924***	2.107***	2.320***	2.565***	2.844***	3.160***
	(0.111)	(0.116)	(0.130)	(0.154)	(0.187)	(0.230)	(0.284)	(0.349)	(0.428)	(0.521)	(0.632)
Children under 6	0.684***	0.672***	0.643***	0.606***	0.564***	0.521***	0.479***	0.438***	0.398***	0.362***	0.328***
	(0.043)	(0.044)	(0.045)	(0.048)	(0.051)	(0.053)	(0.055)	(0.056)	(0.057)	(0.057)	(0.056)
6-14 year old	0.551***	0.546***	0.531***	0.508***	0.477***	0.441***	0.405***	0.368***	0.333***	0.300***	0.269***
	(0.071)	(0.071)	(0.074)	(0.077)	(0.081)	(0.084)	(0.086)	(0.087)	(0.087)	(0.086)	(0.084)
Ln (Debt)	0.994	0.994	0.994	0.993	0.993	0.992	0.991	0.990	0.989	0.988	0.988
	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.010)	(0.011)	(0.013)	(0.014)	(0.015)
Only Daughters	1.026	1.027	1.032	1.041	1.055	1.073	1.092	1.114	1.137	1.160	1.184
	(0.080)	(0.083)	(0.090)	(0.103)	(0.118)	(0.136)	(0.157)	(0.179)	(0.203)	(0.228)	(0.254)
Siblings	0.984	0.983	0.982	0.980	0.977	0.974	0.970	0.965	0.961	0.956	0.951
	(0.013)	(0.013)	(0.014)	(0.016)	(0.018)	(0.020)	(0.023)	(0.026)	(0.028)	(0.031)	(0.033)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and it is uncorrelated with any of the observed explanatory variables.

Table B.5: Primary to secondary transition for daughters: Sensitivity analysis for odds ratio estimates

Effect of unobserved heterogeneity βu (LnOdds)	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Father Education	1.053***	1.057***	1.066***	1.078***	1.093***	1.109***	1.127***	1.145***	1.164***	1.184***	1.204***
	(0.008)	(0.008)	(0.009)	(0.010)	(0.011)	(0.013)	(0.015)	(0.017)	(0.019)	(0.021)	(0.023)
Migrant	0.706*	0.697*	0.670*	0.632*	0.591*	0.548*	0.505*	0.463*	0.423*	0.386*	0.351*
	(0.109)	(0.112)	(0.119)	(0.128)	(0.135)	(0.142)	(0.147)	(0.150)	(0.152)	(0.152)	(0.151)
Father Education × Migrant	1.055*	1.057*	1.064*	1.075*	1.088*	1.102*	1.118*	1.134*	1.150*	1.167*	1.184*
	(0.028)	(0.029)	(0.033)	(0.037)	(0.043)	(0.049)	(0.056)	(0.064)	(0.071)	(0.080)	(0.088)
Mother Education	1.052***	1.054***	1.060***	1.070***	1.081***	1.094***	1.107***	1.122***	1.137***	1.152***	1.168***
	(0.009)	(0.009)	(0.010)	(0.011)	(0.013)	(0.015)	(0.017)	(0.019)	(0.021)	(0.024)	(0.026)
Remittances	0.967***	0.965***	0.961***	0.956***	0.949***	0.941***	0.934***	0.926***	0.918***	0.910***	0.902***
	(0.010)	(0.010)	(0.011)	(0.012)	(0.014)	(0.016)	(0.017)	(0.019)	(0.021)	(0.022)	(0.024)
Child Age	1.374***	1.393***	1.446***	1.526***	1.625***	1.741***	1.873***	2.020***	2.182***	2.360***	2.555***
	(0.010)	(0.011)	(0.012)	(0.014)	(0.017)	(0.021)	(0.025)	(0.030)	(0.035)	(0.041)	(0.048)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.998***	0.998***	0.998***	0.998***	0.998***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.992	0.992	0.991	0.990	0.988	0.986	0.984	0.982	0.980	0.978	0.976
	(0.010)	(0.010)	(0.011)	(0.013)	(0.014)	(0.016)	(0.018)	(0.020)	(0.022)	(0.025)	(0.027)
Household Assets	1.095***	1.099***	1.112***	1.130***	1.152***	1.176***	1.202***	1.229***	1.258***	1.287***	1.317***
	(0.006)	(0.007)	(0.007)	(0.009)	(0.010)	(0.011)	(0.013)	(0.015)	(0.017)	(0.019)	(0.021)
Non-Hindu	0.648***	0.633***	0.593***	0.543***	0.491***	0.439***	0.391***	0.348***	0.308***	0.273***	0.242***
	(0.041)	(0.042)	(0.043)	(0.044)	(0.045)	(0.046)	(0.046)	(0.045)	(0.044)	(0.043)	(0.041)
OBC	0.947	0.942	0.928	0.908	0.884	0.858	0.833	0.807	0.782	0.757	0.732
	(0.064)	(0.066)	(0.071)	(0.079)	(0.087)	(0.095)	(0.104)	(0.112)	(0.119)	(0.126)	(0.133)
SC & ST	1.024	1.021	1.014	1.006	0.996	0.986	0.975	0.964	0.953	0.942	0.930
	(0.076)	(0.078)	(0.085)	(0.095)	(0.107)	(0.120)	(0.133)	(0.146)	(0.159)	(0.172)	(0.185)

Table B.5 (continued).

Urban	1.744***	1.785***	1.905***	2.092***	2.339***	2.641***	3.000***	3.419***	3.907***	4.474***	5.130***
	(0.112)	(0.119)	(0.140)	(0.173)	(0.219)	(0.279)	(0.355)	(0.449)	(0.566)	(0.710)	(0.885)
Children under 6	0.584***	0.568***	0.528***	0.477***	0.424***	0.372***	0.325***	0.282***	0.245***	0.211***	0.182***
	(0.035)	(0.036)	(0.037)	(0.037)	(0.038)	(0.037)	(0.037)	(0.035)	(0.034)	(0.032)	(0.030)
6-14 year old	0.717**	0.706***	0.675***	0.631***	0.583***	0.535***	0.490***	0.447***	0.407***	0.370***	0.336***
	(0.074)	(0.074)	(0.076)	(0.079)	(0.081)	(0.083)	(0.084)	(0.085)	(0.084)	(0.084)	(0.082)
Ln (Debt)	1.000	1.000	0.999	0.998	0.997	0.996	0.995	0.994	0.993	0.991	0.990
	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.010)	(0.011)	(0.012)	(0.013)
Only Daughters	1.137	1.143	1.161	1.187	1.219*	1.256*	1.297*	1.340*	1.385*	1.431*	1.478*
	(0.078)	(0.081)	(0.091)	(0.105)	(0.122)	(0.142)	(0.165)	(0.189)	(0.216)	(0.244)	(0.274)
Siblings	0.948***	0.946***	0.941***	0.934***	0.926***	0.917***	0.907***	0.898***	0.888***	0.879***	0.869***
	(0.012)	(0.012)	(0.013)	(0.015)	(0.017)	(0.018)	(0.020)	(0.023)	(0.025)	(0.027)	(0.029)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and it is uncorrelated with any of the observed explanatory variables.

Table B.6: Secondary to post-secondary transition for daughters: Sensitivity analysis for odds ratio estimates

Effect of unobserved heterogeneity βu (LnOdds)	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Father Education	1.059***	1.064***	1.075***	1.090***	1.107***	1.125***	1.145***	1.165***	1.187***	1.209***	1.231***
	(0.009)	(0.010)	(0.011)	(0.013)	(0.015)	(0.017)	(0.019)	(0.022)	(0.025)	(0.028)	(0.031)
Migrant	0.769	0.746	0.698	0.647	0.595	0.543	0.494	0.448	0.406	0.366	0.330
	(0.291)	(0.294)	(0.301)	(0.311)	(0.323)	(0.333)	(0.341)	(0.345)	(0.345)	(0.343)	(0.337)
Father Education × Migrant	1.006	1.009	1.016	1.022	1.028	1.034	1.039	1.045	1.051	1.057	1.063
	(0.051)	(0.053)	(0.059)	(0.067)	(0.076)	(0.087)	(0.098)	(0.110)	(0.123)	(0.136)	(0.149)
Mother Education	1.088***	1.092***	1.104***	1.120***	1.140***	1.161***	1.185***	1.210***	1.236***	1.263***	1.292***
	(0.010)	(0.011)	(0.012)	(0.014)	(0.016)	(0.018)	(0.021)	(0.024)	(0.027)	(0.031)	(0.034)
Remittances	0.982	0.980	0.975	0.969	0.963	0.956*	0.949*	0.942*	0.935*	0.928*	0.921*
	(0.013)	(0.014)	(0.015)	(0.017)	(0.019)	(0.021)	(0.024)	(0.026)	(0.029)	(0.031)	(0.034)
Child Age	2.147***	2.224***	2.440***	2.779***	3.247***	3.862***	4.648***	5.636***	6.864***	8.381***	10.25***
	(0.086)	(0.093)	(0.112)	(0.143)	(0.188)	(0.251)	(0.337)	(0.453)	(0.607)	(0.810)	(1.075)
Child Age^2	0.998***	0.998***	0.998***	0.998***	0.997***	0.997***	0.997***	0.996***	0.996***	0.995***	0.995***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.990	0.990	0.988	0.986	0.984	0.981	0.979	0.976	0.973	0.971	0.968
	(0.018)	(0.018)	(0.020)	(0.023)	(0.026)	(0.029)	(0.033)	(0.037)	(0.041)	(0.045)	(0.049)
Household Assets	1.105***	1.112***	1.131***	1.156***	1.184***	1.215***	1.249***	1.285***	1.323***	1.363***	1.405***
	(0.008)	(0.009)	(0.010)	(0.011)	(0.013)	(0.015)	(0.018)	(0.020)	(0.023)	(0.026)	(0.030)
Non-Hindu	0.704***	0.684***	0.638***	0.582***	0.526***	0.473***	0.425***	0.381***	0.342***	0.306***	0.274***
	(0.058)	(0.059)	(0.061)	(0.063)	(0.064)	(0.066)	(0.066)	(0.066)	(0.065)	(0.064)	(0.063)
OBC	1.000	0.998	0.991	0.982	0.972	0.962	0.954	0.948	0.943	0.939	0.936
	(0.080)	(0.084)	(0.092)	(0.104)	(0.117)	(0.131)	(0.146)	(0.161)	(0.178)	(0.194)	(0.211)
SC & ST	0.858	0.851	0.835	0.812	0.788	0.762	0.736	0.711	0.686	0.662	0.639
	(0.076)	(0.079)	(0.086)	(0.095)	(0.105)	(0.115)	(0.125)	(0.135)	(0.144)	(0.153)	(0.161)

Table B.6 (continued).

Urban	1.167*	1.190*	1.246*	1.321**	1.410**	1.511**	1.625***	1.750***	1.889***	2.041***	2.206***
	(0.088)	(0.094)	(0.109)	(0.132)	(0.160)	(0.194)	(0.234)	(0.281)	(0.336)	(0.399)	(0.471)
Children under 6	0.751**	0.727**	0.672***	0.611***	0.551***	0.493***	0.440***	0.392***	0.347***	0.308***	0.272***
	(0.075)	(0.076)	(0.077)	(0.079)	(0.081)	(0.082)	(0.082)	(0.081)	(0.080)	(0.077)	(0.075)
6-14 year old	0.806**	0.790**	0.751***	0.707***	0.662***	0.620***	0.579***	0.540***	0.503***	0.468***	0.436***
	(0.059)	(0.060)	(0.064)	(0.068)	(0.073)	(0.078)	(0.082)	(0.085)	(0.088)	(0.090)	(0.092)
Ln (Debt)	1.016**	1.017**	1.018*	1.020*	1.023*	1.025*	1.029*	1.032*	1.036*	1.039*	1.043*
	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.011)	(0.012)	(0.013)	(0.015)	(0.016)	(0.018)
Only Daughters	1.177	1.184	1.205	1.237	1.279	1.329	1.388*	1.455*	1.528*	1.608*	1.694*
	(0.102)	(0.107)	(0.121)	(0.140)	(0.165)	(0.194)	(0.228)	(0.266)	(0.309)	(0.357)	(0.411)
Siblings	1.009	1.009	1.008	1.007	1.005	1.003	1.000	0.998	0.995	0.993	0.992
	(0.018)	(0.019)	(0.021)	(0.024)	(0.027)	(0.030)	(0.034)	(0.037)	(0.041)	(0.045)	(0.049)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and it is uncorrelated with any of the observed explanatory variables.

Table B.7: Non-literate to primary transition for sons: Odds ratio estimates assuming correlation between omitted confounder and father's education

Correlation between unobserved heterogeneity (u) and father's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.502***	1.363***	1.233***	1.113***	1.001	0.899***	0.804***
	(0.017)	(0.017)	(0.016)	(0.014)	(0.013)	(0.011)	(0.009)
Migrant	0.741	0.717	0.704	0.700	0.704	0.717	0.741
	(0.154)	(0.161)	(0.164)	(0.166)	(0.164)	(0.161)	(0.154)
Father Education × Migrant	1.017	1.019	1.021	1.021	1.021	1.019	1.017
	(0.036)	(0.039)	(0.041)	(0.041)	(0.041)	(0.039)	(0.036)
Mother Education	1.038**	1.041**	1.042**	1.042**	1.042**	1.041**	1.038**
	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
Remittances	0.966*	0.963*	0.962*	0.961*	0.962*	0.963*	0.966*
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Child Age	1.207***	1.225***	1.235***	1.238***	1.235***	1.225***	1.207***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.983	0.982	0.981	0.981	0.981	0.982	0.983
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)
Household Assets	1.091***	1.097***	1.101***	1.102***	1.101***	1.097***	1.091***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Non-Hindu	0.508***	0.484***	0.471***	0.467***	0.471***	0.484***	0.508***
	(0.048)	(0.050)	(0.050)	(0.051)	(0.050)	(0.050)	(0.048)
OBC	0.813	0.804	0.799	0.798	0.799	0.804	0.813
	(0.087)	(0.093)	(0.096)	(0.097)	(0.096)	(0.093)	(0.087)
SC & ST	0.941	0.940	0.940	0.940	0.940	0.940	0.941

Table B.7 (continued).

	(0.111)	(0.119)	(0.124)	(0.125)	(0.124)	(0.119)	(0.111)
Urban	1.687***	1.753***	1.791***	1.804***	1.791***	1.753***	1.687***
	(0.166)	(0.186)	(0.198)	(0.202)	(0.198)	(0.186)	(0.166)
Children under 6	0.653***	0.631***	0.619***	0.615***	0.619***	0.631***	0.653***
	(0.056)	(0.059)	(0.060)	(0.060)	(0.060)	(0.059)	(0.056)
6-14 year old	0.625**	0.614**	0.607**	0.605**	0.607**	0.614**	0.625**
	(0.098)	(0.103)	(0.105)	(0.106)	(0.105)	(0.103)	(0.098)
Ln (Debt)	1.008	1.009	1.009	1.009	1.009	1.009	1.008
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Only Sons	0.774**	0.759**	0.751**	0.748**	0.751**	0.759**	0.774**
	(0.073)	(0.077)	(0.080)	(0.080)	(0.080)	(0.077)	(0.073)
Siblings	0.750***	0.733***	0.723***	0.720***	0.723***	0.733***	0.750***
	(0.024)	(0.025)	(0.025)	(0.026)	(0.025)	(0.025)	(0.024)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.8: Primary to secondary transition for Sons: Odds ratio estimates under different assumptions of correlation between unobserved heterogeneity and father's education

Correlation between unobserved heterogeneity (u) and father's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.471***	1.335***	1.208***	1.090***	1.076***	1.048***	1.033**
	(0.014)	(0.014)	(0.013)	(0.012)	(0.012)	(0.011)	(0.010)
Migrant	0.596*	0.571*	0.557*	0.552*	0.557*	0.571*	0.596*
	(0.135)	(0.139)	(0.141)	(0.141)	(0.141)	(0.139)	(0.135)
Father Education × Migrant	1.088*	1.095*	1.099*	1.100*	1.099*	1.095*	1.088*
	(0.040)	(0.044)	(0.045)	(0.046)	(0.045)	(0.044)	(0.040)
Mother Education	1.068***	1.073***	1.076***	1.077***	1.076***	1.073***	1.068***
	(0.012)	(0.013)	(0.014)	(0.014)	(0.014)	(0.013)	(0.012)
Remittances	0.958**	0.955**	0.953**	0.953**	0.953**	0.955**	0.958**
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Child Age	1.638***	1.704***	1.742***	1.755***	1.742***	1.704***	1.638***
	(0.016)	(0.018)	(0.019)	(0.019)	(0.019)	(0.018)	(0.016)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.988	0.986	0.986	0.985	0.986	0.986	0.988
	(0.014)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.014)
Household Assets	1.139***	1.151***	1.158***	1.160***	1.158***	1.151***	1.139***
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
Non-Hindu	0.422***	0.391***	0.375***	0.370***	0.375***	0.391***	0.422***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
OBC	1.207*	1.219*	1.225*	1.227*	1.225*	1.219*	1.207*
	(0.109)	(0.118)	(0.124)	(0.125)	(0.124)	(0.118)	(0.109)
SC & ST	1.008	1.006	1.004	1.004	1.004	1.006	1.008

Table B.8 (continued).

	(0.099)	(0.106)	(0.110)	(0.111)	(0.110)	(0.106)	(0.099)
Urban	2.143***	2.276***	2.353***	2.379***	2.353***	2.276***	2.143***
	(0.182)	(0.208)	(0.223)	(0.228)	(0.223)	(0.208)	(0.182)
Children under 6	0.588***	0.564***	0.550***	0.546***	0.550***	0.564***	0.588***
	(0.053)	(0.055)	(0.056)	(0.056)	(0.056)	(0.055)	(0.053)
6-14 year old	0.703**	0.680**	0.667**	0.663**	0.667**	0.680**	0.703**
	(0.082)	(0.085)	(0.086)	(0.087)	(0.086)	(0.085)	(0.082)
Ln (Debt)	1.008	1.008	1.009	1.009	1.009	1.008	1.008
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Only Sons	0.666***	0.643***	0.631***	0.627***	0.631***	0.643***	0.666***
	(0.054)	(0.056)	(0.057)	(0.057)	(0.057)	(0.056)	(0.054)
Siblings	0.724***	0.704***	0.694***	0.690***	0.694***	0.704***	0.724***
	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.9: Secondary to post-secondary transition for Sons: Sensitivity analysis for odds ratio estimates

Correlation between unobserved heterogeneity (u) and father's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.561***	1.423***	1.290***	1.165***	1.166***	1.098***	1.064***
	(0.018)	(0.018)	(0.017)	(0.015)	(0.014)	(0.014)	(0.013)
Migrant	1.299	1.309	1.316	1.318	1.316	1.309	1.299
	(0.518)	(0.560)	(0.585)	(0.593)	(0.585)	(0.560)	(0.518)
Father Education × Migrant	0.876*	0.869*	0.865*	0.864*	0.865*	0.869*	0.876*
	(0.052)	(0.055)	(0.057)	(0.058)	(0.057)	(0.055)	(0.052)
Mother Education	1.096***	1.105***	1.109***	1.111***	1.109***	1.105***	1.096***
	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
Remittances	0.998	0.997	0.996	0.996	0.996	0.997	0.998
	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.018)
Child Age	3.146***	3.440***	3.617***	3.676***	3.617***	3.440***	3.146***
	(0.174)	(0.204)	(0.222)	(0.228)	(0.222)	(0.204)	(0.174)
Child Age^2	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	1.006	1.006	1.006	1.006	1.006	1.006	1.006
	(0.021)	(0.022)	(0.023)	(0.024)	(0.023)	(0.022)	(0.021)
Household Assets	1.164***	1.180***	1.189***	1.192***	1.189***	1.180***	1.164***
	(0.012)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.012)
Non-Hindu	0.664***	0.635***	0.619***	0.614***	0.619***	0.635***	0.664***
	(0.075)	(0.077)	(0.078)	(0.078)	(0.078)	(0.077)	(0.075)
OBC	1.111	1.121	1.127	1.129	1.127	1.121	1.111
	(0.120)	(0.131)	(0.137)	(0.139)	(0.137)	(0.131)	(0.120)
SC & ST	0.830	0.818	0.812	0.810	0.812	0.818	0.830

Table B.9 (continued).

	(0.100)	(0.106)	(0.110)	(0.111)	(0.110)	(0.106)	(0.100)
Urban	1.529***	1.595***	1.632***	1.644***	1.632***	1.595***	1.529***
	(0.155)	(0.174)	(0.186)	(0.189)	(0.186)	(0.174)	(0.155)
Children under 6	0.475***	0.446***	0.430***	0.425***	0.430***	0.446***	0.475***
	(0.066)	(0.066)	(0.067)	(0.067)	(0.067)	(0.066)	(0.066)
6-14 year old	0.796*	0.780*	0.771*	0.768*	0.771*	0.780*	0.796*
	(0.079)	(0.084)	(0.086)	(0.087)	(0.086)	(0.084)	(0.079)
Ln (Debt)	1.032***	1.035***	1.036***	1.037***	1.036***	1.035***	1.032***
	(0.008)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.008)
Only Sons	0.786*	0.767*	0.757*	0.754*	0.757*	0.767*	0.786*
	(0.080)	(0.084)	(0.086)	(0.087)	(0.086)	(0.084)	(0.080)
Siblings	0.822***	0.806***	0.797***	0.794***	0.797***	0.806***	0.822***
	(0.038)	(0.040)	(0.041)	(0.041)	(0.041)	(0.040)	(0.038)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.10: Non-literate to primary transition for daughters: Odds ratio estimates assuming correlation between omitted confounder and father's education

Correlation between unobserved heterogeneity (u) and father's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.520***	1.380***	1.248***	1.126***	1.066***	1.048***	1.029***
	(0.018)	(0.018)	(0.017)	(0.015)	(0.013)	(0.011)	(0.010)
Migrant	0.791	0.779	0.773	0.771	0.773	0.779	0.791
	(0.171)	(0.182)	(0.188)	(0.189)	(0.188)	(0.182)	(0.171)
Father Education × Migrant	1.066	1.073	1.077	1.078	1.077	1.073	1.066
	(0.044)	(0.048)	(0.050)	(0.051)	(0.050)	(0.048)	(0.044)
Mother Education	1.062***	1.066***	1.068***	1.068***	1.068***	1.066***	1.062***
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Remittances	0.956**	0.952**	0.950**	0.949**	0.950**	0.952**	0.956**
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Child Age	1.193***	1.210***	1.220***	1.223***	1.220***	1.210***	1.193***
	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.006)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.992	0.991	0.990	0.990	0.990	0.991	0.992
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Household Assets	1.109***	1.118***	1.122***	1.124***	1.122***	1.118***	1.109***
	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)
Non-Hindu	0.540***	0.517***	0.505***	0.501***	0.505***	0.517***	0.540***
	(0.054)	(0.056)	(0.057)	(0.057)	(0.057)	(0.056)	(0.054)
OBC	0.617***	0.597***	0.585***	0.582***	0.585***	0.597***	0.617***
	(0.072)	(0.074)	(0.076)	(0.076)	(0.076)	(0.074)	(0.072)
SC & ST	0.797	0.785	0.779	0.777	0.779	0.785	0.797

Table B.10 (continued).

	(0.102)	(0.108)	(0.112)	(0.113)	(0.112)	(0.108)	(0.102)
Urban	1.771***	1.857***	1.907***	1.924***	1.907***	1.857***	1.771***
	(0.187)	(0.211)	(0.225)	(0.230)	(0.225)	(0.211)	(0.187)
Children under 6	0.564***	0.539***	0.526***	0.521***	0.526***	0.539***	0.564***
	(0.050)	(0.052)	(0.053)	(0.053)	(0.053)	(0.052)	(0.050)
6-14 year old	0.477***	0.457***	0.445***	0.441***	0.445***	0.457***	0.477***
	(0.081)	(0.083)	(0.083)	(0.084)	(0.083)	(0.083)	(0.081)
Ln (Debt)	0.993	0.992	0.992	0.992	0.992	0.992	0.993
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Only Daughters	1.055	1.065	1.071	1.073	1.071	1.065	1.055
	(0.118)	(0.128)	(0.134)	(0.136)	(0.134)	(0.128)	(0.118)
Siblings	0.977	0.975	0.974	0.974	0.974	0.975	0.977
	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.018)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.11: Primary to secondary transition for daughters: Odds ratio estimates under different assumptions of correlation between unobserved heterogeneity and father's education

Correlation between unobserved heterogeneity (u) and father's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.496***	1.359***	1.230***	1.109***	1.076***	1.045***	1.031***
	(0.015)	(0.015)	(0.014)	(0.013)	(0.012)	(0.012)	(0.010)
Migrant	0.591*	0.566*	0.553*	0.548*	0.553*	0.566*	0.591*
	(0.135)	(0.140)	(0.141)	(0.142)	(0.141)	(0.140)	(0.135)
Father Education × Migrant	1.088*	1.096*	1.101*	1.102*	1.101*	1.096*	1.088*
	(0.043)	(0.047)	(0.049)	(0.049)	(0.049)	(0.047)	(0.043)
Mother Education	1.081***	1.088***	1.092***	1.094***	1.092***	1.088***	1.081***
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Remittances	0.949***	0.945***	0.942***	0.941***	0.942***	0.945***	0.949***
	(0.014)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.014)
Child Age	1.625***	1.691***	1.729***	1.741***	1.729***	1.691***	1.625***
	(0.017)	(0.019)	(0.020)	(0.021)	(0.020)	(0.019)	(0.017)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.988	0.987	0.986	0.986	0.986	0.987	0.988
	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
Household Assets	1.152***	1.166***	1.174***	1.176***	1.174***	1.166***	1.152***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Non-Hindu	0.491***	0.460***	0.444***	0.439***	0.444***	0.460***	0.491***
	(0.045)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.045)
OBC	0.884	0.869	0.861	0.858	0.861	0.869	0.884
	(0.087)	(0.092)	(0.094)	(0.095)	(0.094)	(0.092)	(0.087)
SC & ST	0.996	0.990	0.987	0.986	0.987	0.990	0.996

Table B.11 (continued).

	(0.107)	(0.114)	(0.118)	(0.120)	(0.118)	(0.114)	(0.107)
Urban	2.339***	2.509***	2.608***	2.641***	2.608***	2.509***	2.339***
	(0.219)	(0.252)	(0.272)	(0.279)	(0.272)	(0.252)	(0.219)
Children under 6	0.424***	0.393***	0.377***	0.372***	0.377***	0.393***	0.424***
	(0.038)	(0.038)	(0.037)	(0.037)	(0.037)	(0.038)	(0.038)
6-14 year old	0.583***	0.555***	0.540***	0.535***	0.540***	0.555***	0.583***
	(0.081)	(0.082)	(0.083)	(0.083)	(0.083)	(0.082)	(0.081)
Ln (Debt)	0.997	0.997	0.996	0.996	0.996	0.997	0.997
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
Only Daughters	1.219*	1.240*	1.252*	1.256*	1.252*	1.240*	1.219*
	(0.122)	(0.134)	(0.140)	(0.142)	(0.140)	(0.134)	(0.122)
Siblings	0.926***	0.921***	0.918***	0.917***	0.918***	0.921***	0.926***
	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.016)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.12 : Secondary to post-secondary transition for daughters: Sensitivity analysis for odds ratio estimates

Correlation between unobserved heterogeneity (u) and father's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.515***	1.377***	1.247***	1.125***	1.092***	1.073***	1.044***
	(0.020)	(0.020)	(0.018)	(0.017)	(0.015)	(0.013)	(0.011)
Migrant	0.595	0.565	0.548	0.543	0.548	0.565	0.595
	(0.323)	(0.329)	(0.332)	(0.333)	(0.332)	(0.329)	(0.323)
Father Education × Migrant	1.028	1.031	1.033	1.034	1.033	1.031	1.028
	(0.076)	(0.082)	(0.085)	(0.086)	(0.085)	(0.082)	(0.076)
Mother Education	1.140***	1.152***	1.159***	1.161***	1.159***	1.152***	1.140***
	(0.016)	(0.017)	(0.018)	(0.018)	(0.018)	(0.017)	(0.016)
Remittances	0.963	0.959*	0.957*	0.956*	0.957*	0.959*	0.963
	(0.019)	(0.020)	(0.021)	(0.021)	(0.021)	(0.020)	(0.019)
Child Age	3.247***	3.586***	3.793***	3.862***	3.793***	3.586***	3.247***
	(0.188)	(0.222)	(0.243)	(0.251)	(0.243)	(0.222)	(0.188)
Child Age^2	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.984	0.982	0.981	0.981	0.981	0.982	0.984
	(0.026)	(0.028)	(0.029)	(0.029)	(0.029)	(0.028)	(0.026)
Household Assets	1.184***	1.202***	1.212***	1.215***	1.212***	1.202***	1.184***
	(0.013)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.013)
Non-Hindu	0.526***	0.495***	0.479***	0.473***	0.479***	0.495***	0.526***
	(0.064)	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)	(0.064)
OBC	0.972	0.966	0.963	0.962	0.963	0.966	0.972
	(0.117)	(0.125)	(0.129)	(0.131)	(0.129)	(0.125)	(0.117)
SC & ST	0.788	0.773	0.764	0.762	0.764	0.773	0.788

Table B.12 (continued).

	(0.105)	(0.111)	(0.114)	(0.115)	(0.114)	(0.111)	(0.105)
Urban	1.410**	1.468**	1.501**	1.511**	1.501**	1.468**	1.410**
	(0.160)	(0.179)	(0.190)	(0.194)	(0.190)	(0.179)	(0.160)
Children under 6	0.551***	0.517***	0.499***	0.493***	0.499***	0.517***	0.551***
	(0.081)	(0.082)	(0.082)	(0.082)	(0.082)	(0.082)	(0.081)
6-14 year old	0.662***	0.637***	0.624***	0.620***	0.624***	0.637***	0.662***
	(0.073)	(0.076)	(0.077)	(0.078)	(0.077)	(0.076)	(0.073)
Ln (Debt)	1.023*	1.024*	1.025*	1.025*	1.025*	1.024*	1.023*
	(0.009)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.009)
Only Daughters	1.279	1.307	1.324	1.329	1.324	1.307	1.279
	(0.165)	(0.181)	(0.191)	(0.194)	(0.191)	(0.181)	(0.165)
Siblings	1.005	1.004	1.003	1.003	1.003	1.004	1.005
	(0.027)	(0.029)	(0.030)	(0.030)	(0.030)	(0.029)	(0.027)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.13: Non-literate to primary transition for sons: Odds ratio estimates assuming correlation between omitted confounder and mother's education

Correlation between unobserved heterogeneity (u) and mother's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.099***	1.107***	1.111***	1.113***	1.111***	1.107***	1.099***
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)
Migrant	0.741	0.717	0.704	0.700	0.704	0.717	0.741
	(0.154)	(0.161)	(0.164)	(0.166)	(0.164)	(0.161)	(0.154)
Father Education × Migrant	1.017	1.019	1.021	1.021	1.021	1.019	1.017
	(0.036)	(0.039)	(0.041)	(0.041)	(0.041)	(0.039)	(0.036)
Mother Education	1.453***	1.302***	1.165***	1.082***	1.067***	1.044***	1.023*
	(0.020)	(0.019)	(0.018)	(0.015)	(0.014)	(0.012)	(0.009)
Remittances	0.966*	0.963*	0.962*	0.961*	0.962*	0.963*	0.966*
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Child Age	1.207***	1.225***	1.235***	1.238***	1.235***	1.225***	1.207***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.983	0.982	0.981	0.981	0.981	0.982	0.983
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)
Household Assets	1.091***	1.097***	1.101***	1.102***	1.101***	1.097***	1.091***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Non-Hindu	0.508***	0.484***	0.471***	0.467***	0.471***	0.484***	0.508***
	(0.048)	(0.050)	(0.050)	(0.051)	(0.050)	(0.050)	(0.048)
OBC	0.813	0.804	0.799	0.798	0.799	0.804	0.813
	(0.087)	(0.093)	(0.096)	(0.097)	(0.096)	(0.093)	(0.087)
SC & ST	0.941	0.940	0.940	0.940	0.940	0.940	0.941

Table B.13 (continued).

	(0.111)	(0.119)	(0.124)	(0.125)	(0.124)	(0.119)	(0.111)
Urban	1.687***	1.753***	1.791***	1.804***	1.791***	1.753***	1.687***
	(0.166)	(0.186)	(0.198)	(0.202)	(0.198)	(0.186)	(0.166)
Children under 6	0.653***	0.631***	0.619***	0.615***	0.619***	0.631***	0.653***
	(0.056)	(0.059)	(0.060)	(0.060)	(0.060)	(0.059)	(0.056)
6-14 year old	0.625**	0.614**	0.607**	0.605**	0.607**	0.614**	0.625**
	(0.098)	(0.103)	(0.105)	(0.106)	(0.105)	(0.103)	(0.098)
Ln (Debt)	1.008	1.009	1.009	1.009	1.009	1.009	1.008
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Only Sons	0.774**	0.759**	0.751**	0.748**	0.751**	0.759**	0.774**
	(0.073)	(0.077)	(0.080)	(0.080)	(0.080)	(0.077)	(0.073)
Siblings	0.750***	0.733***	0.723***	0.720***	0.723***	0.733***	0.750***
	(0.024)	(0.025)	(0.025)	(0.026)	(0.025)	(0.025)	(0.024)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.14: Primary to secondary transition for Sons: Odds ratio estimates under different assumptions of correlation between unobserved heterogeneity and mother's education

Correlation between unobserved heterogeneity (u) and mother's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.076***	1.084***	1.088***	1.090***	1.088***	1.084***	1.076***
	(0.010)	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.010)
Migrant	0.596*	0.571*	0.557*	0.552*	0.557*	0.571*	0.596*
	(0.135)	(0.139)	(0.141)	(0.141)	(0.141)	(0.139)	(0.135)
Father Education × Migrant	1.088*	1.095*	1.099*	1.100*	1.099*	1.095*	1.088*
	(0.040)	(0.044)	(0.045)	(0.046)	(0.045)	(0.044)	(0.040)
Mother Education	1.495***	1.343***	1.204***	1.077***	1.059***	1.038***	1.022*
	(0.017)	(0.016)	(0.015)	(0.014)	(0.012)	(0.010)	(0.008)
Remittances	0.958**	0.955**	0.953**	0.953**	0.953**	0.955**	0.958**
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Child Age	1.638***	1.704***	1.742***	1.755***	1.742***	1.704***	1.638***
	(0.016)	(0.018)	(0.019)	(0.019)	(0.019)	(0.018)	(0.016)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.988	0.986	0.986	0.985	0.986	0.986	0.988
	(0.014)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.014)
Household Assets	1.139***	1.151***	1.158***	1.160***	1.158***	1.151***	1.139***
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
Non-Hindu	0.422***	0.391***	0.375***	0.370***	0.375***	0.391***	0.422***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
OBC	1.207*	1.219*	1.225*	1.227*	1.225*	1.219*	1.207*
	(0.109)	(0.118)	(0.124)	(0.125)	(0.124)	(0.118)	(0.109)
SC & ST	1.008	1.006	1.004	1.004	1.004	1.006	1.008

Table B.14 (continued).

	(0.099)	(0.106)	(0.110)	(0.111)	(0.110)	(0.106)	(0.099)
Urban	2.143***	2.276***	2.353***	2.379***	2.353***	2.276***	2.143***
	(0.182)	(0.208)	(0.223)	(0.228)	(0.223)	(0.208)	(0.182)
Children under 6	0.588***	0.564***	0.550***	0.546***	0.550***	0.564***	0.588***
	(0.053)	(0.055)	(0.056)	(0.056)	(0.056)	(0.055)	(0.053)
6-14 year old	0.703**	0.680**	0.667**	0.663**	0.667**	0.680**	0.703**
	(0.082)	(0.085)	(0.086)	(0.087)	(0.086)	(0.085)	(0.082)
Ln (Debt)	1.008	1.008	1.009	1.009	1.009	1.008	1.008
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Only Sons	0.666***	0.643***	0.631***	0.627***	0.631***	0.643***	0.666***
	(0.054)	(0.056)	(0.057)	(0.057)	(0.057)	(0.056)	(0.054)
Siblings	0.724***	0.704***	0.694***	0.690***	0.694***	0.704***	0.724***
	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.15: Secondary to post-secondary transition for Sons: Odds ratio estimates assuming correlation between omitted confounder and mother's education

Correlation between unobserved heterogeneity (u) and mother's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.142***	1.155***	1.163***	1.165***	1.163***	1.155***	1.142***
	(0.013)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.013)
Migrant	1.299	1.309	1.316	1.318	1.316	1.309	1.299
	(0.518)	(0.560)	(0.585)	(0.593)	(0.585)	(0.560)	(0.518)
Father Education × Migrant	0.876*	0.869*	0.865*	0.864*	0.865*	0.869*	0.876*
	(0.052)	(0.055)	(0.057)	(0.058)	(0.057)	(0.055)	(0.052)
Mother Education	1.534***	1.382***	1.241***	1.111***	0.992	0.883***	0.783***
	(0.020)	(0.019)	(0.018)	(0.016)	(0.014)	(0.012)	(0.010)
Remittances	0.998	0.997	0.996	0.996	0.996	0.997	0.998
	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.018)
Child Age	3.146***	3.440***	3.617***	3.676***	3.617***	3.440***	3.146***
	(0.174)	(0.204)	(0.222)	(0.228)	(0.222)	(0.204)	(0.174)
Child Age^2	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	1.006	1.006	1.006	1.006	1.006	1.006	1.006
	(0.021)	(0.022)	(0.023)	(0.024)	(0.023)	(0.022)	(0.021)
Household Assets	1.164***	1.180***	1.189***	1.192***	1.189***	1.180***	1.164***
	(0.012)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.012)
Non-Hindu	0.664***	0.635***	0.619***	0.614***	0.619***	0.635***	0.664***
	(0.075)	(0.077)	(0.078)	(0.078)	(0.078)	(0.077)	(0.075)
OBC	1.111	1.121	1.127	1.129	1.127	1.121	1.111
	(0.120)	(0.131)	(0.137)	(0.139)	(0.137)	(0.131)	(0.120)
SC & ST	0.830	0.818	0.812	0.810	0.812	0.818	0.830

Table B.15 (continued).

	(0.100)	(0.106)	(0.110)	(0.111)	(0.110)	(0.106)	(0.100)
Urban	1.529***	1.595***	1.632***	1.644***	1.632***	1.595***	1.529***
	(0.155)	(0.174)	(0.186)	(0.189)	(0.186)	(0.174)	(0.155)
Children under 6	0.475***	0.446***	0.430***	0.425***	0.430***	0.446***	0.475***
	(0.066)	(0.066)	(0.067)	(0.067)	(0.067)	(0.066)	(0.066)
6-14 year old	0.796*	0.780*	0.771*	0.768*	0.771*	0.780*	0.796*
	(0.079)	(0.084)	(0.086)	(0.087)	(0.086)	(0.084)	(0.079)
Ln (Debt)	1.032***	1.035***	1.036***	1.037***	1.036***	1.035***	1.032***
	(0.008)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.008)
Only Sons	0.786*	0.767*	0.757*	0.754*	0.757*	0.767*	0.786*
	(0.080)	(0.084)	(0.086)	(0.087)	(0.086)	(0.084)	(0.080)
Siblings	0.822***	0.806***	0.797***	0.794***	0.797***	0.806***	0.822***
	(0.038)	(0.040)	(0.041)	(0.041)	(0.041)	(0.040)	(0.038)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.16: Non-literate to primary transition for daughters: Odds ratio estimates assuming correlation between omitted confounder and mother's education

Correlation between unobserved heterogeneity (u) and mother's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.111***	1.119***	1.124***	1.126***	1.124***	1.119***	1.111***
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Migrant	0.791	0.779	0.773	0.771	0.773	0.779	0.791
	(0.171)	(0.182)	(0.188)	(0.189)	(0.188)	(0.182)	(0.171)
Father Education × Migrant	1.066	1.073	1.077	1.078	1.077	1.073	1.066
	(0.044)	(0.048)	(0.050)	(0.051)	(0.050)	(0.048)	(0.044)
Mother Education	1.482***	1.331***	1.193***	1.089***	1.067***	1.054***	1.031**
	(0.021)	(0.020)	(0.019)	(0.017)	(0.015)	(0.013)	(0.010)
Remittances	0.956**	0.952**	0.950**	0.949**	0.950**	0.952**	0.956**
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Child Age	1.193***	1.210***	1.220***	1.223***	1.220***	1.210***	1.193***
	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.006)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.992	0.991	0.990	0.990	0.990	0.991	0.992
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Household Assets	1.109***	1.118***	1.122***	1.124***	1.122***	1.118***	1.109***
	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)
Non-Hindu	0.540***	0.517***	0.505***	0.501***	0.505***	0.517***	0.540***
	(0.054)	(0.056)	(0.057)	(0.057)	(0.057)	(0.056)	(0.054)
OBC	0.617***	0.597***	0.585***	0.582***	0.585***	0.597***	0.617***
	(0.072)	(0.074)	(0.076)	(0.076)	(0.076)	(0.074)	(0.072)
SC & ST	0.797	0.785	0.779	0.777	0.779	0.785	0.797

Table B.16 (continued).

	(0.102)	(0.108)	(0.112)	(0.113)	(0.112)	(0.108)	(0.102)
Urban	1.771***	1.857***	1.907***	1.924***	1.907***	1.857***	1.771***
	(0.187)	(0.211)	(0.225)	(0.230)	(0.225)	(0.211)	(0.187)
Children under 6	0.564***	0.539***	0.526***	0.521***	0.526***	0.539***	0.564***
	(0.050)	(0.052)	(0.053)	(0.053)	(0.053)	(0.052)	(0.050)
6-14 year old	0.477***	0.457***	0.445***	0.441***	0.445***	0.457***	0.477***
	(0.081)	(0.083)	(0.083)	(0.084)	(0.083)	(0.083)	(0.081)
Ln (Debt)	0.993	0.992	0.992	0.992	0.992	0.992	0.993
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Only Daughters	1.055	1.065	1.071	1.073	1.071	1.065	1.055
	(0.118)	(0.128)	(0.134)	(0.136)	(0.134)	(0.128)	(0.118)
Siblings	0.977	0.975	0.974	0.974	0.974	0.975	0.977
	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.018)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.17 :Primary to secondary transition for daughters: Odds ratio estimates assuming correlation between omitted confounder and mother's education

Correlation between unobserved heterogeneity (u) and mother's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.093***	1.102***	1.108***	1.109***	1.108***	1.102***	1.093***
	(0.011)	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)	(0.011)
Migrant	0.591*	0.566*	0.553*	0.548*	0.553*	0.566*	0.591*
	(0.135)	(0.140)	(0.141)	(0.142)	(0.141)	(0.140)	(0.135)
Father Education × Migrant	1.088*	1.096*	1.101*	1.102*	1.101*	1.096*	1.088*
	(0.043)	(0.047)	(0.049)	(0.049)	(0.049)	(0.047)	(0.043)
Mother Education	1.508***	1.358***	1.220***	1.094***	1.078***	1.052***	1.035***
	(0.018)	(0.017)	(0.016)	(0.015)	(0.013)	(0.011)	(0.009)
Remittances	0.949***	0.945***	0.942***	0.941***	0.942***	0.945***	0.949***
	(0.014)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.014)
Child Age	1.625***	1.691***	1.729***	1.741***	1.729***	1.691***	1.625***
	(0.017)	(0.019)	(0.020)	(0.021)	(0.020)	(0.019)	(0.017)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.988	0.987	0.986	0.986	0.986	0.987	0.988
	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
Household Assets	1.152***	1.166***	1.174***	1.176***	1.174***	1.166***	1.152***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Non-Hindu	0.491***	0.460***	0.444***	0.439***	0.444***	0.460***	0.491***
	(0.045)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.045)
OBC	0.884	0.869	0.861	0.858	0.861	0.869	0.884
	(0.087)	(0.092)	(0.094)	(0.095)	(0.094)	(0.092)	(0.087)
SC & ST	0.996	0.990	0.987	0.986	0.987	0.990	0.996

Table B.17 (continued).

	(0.107)	(0.114)	(0.118)	(0.120)	(0.118)	(0.114)	(0.107)
Urban	2.339***	2.509***	2.608***	2.641***	2.608***	2.509***	2.339***
	(0.219)	(0.252)	(0.272)	(0.279)	(0.272)	(0.252)	(0.219)
Children under 6	0.424***	0.393***	0.377***	0.372***	0.377***	0.393***	0.424***
	(0.038)	(0.038)	(0.037)	(0.037)	(0.037)	(0.038)	(0.038)
6-14 year old	0.583***	0.555***	0.540***	0.535***	0.540***	0.555***	0.583***
	(0.081)	(0.082)	(0.083)	(0.083)	(0.083)	(0.082)	(0.081)
Ln (Debt)	0.997	0.997	0.996	0.996	0.996	0.997	0.997
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
Only Daughters	1.219*	1.240*	1.252*	1.256*	1.252*	1.240*	1.219*
	(0.122)	(0.134)	(0.140)	(0.142)	(0.140)	(0.134)	(0.122)
Siblings	0.926***	0.921***	0.918***	0.917***	0.918***	0.921***	0.926***
	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.016)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.18: Secondary to post-secondary transition for daughters: Odds ratio estimates assuming correlation between omitted confounder and mother's education

Correlation between unobserved heterogeneity (u) and mother's education	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.107***	1.117***	1.123***	1.125***	1.123***	1.117***	1.107***
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Migrant	0.595	0.565	0.548	0.543	0.548	0.565	0.595
	(0.323)	(0.329)	(0.332)	(0.333)	(0.332)	(0.329)	(0.323)
Father Education × Migrant	1.028	1.031	1.033	1.034	1.033	1.031	1.028
	(0.076)	(0.082)	(0.085)	(0.086)	(0.085)	(0.082)	(0.076)
Mother Education	1.590***	1.438***	1.295***	1.161***	1.077***	1.059***	1.029***
	(0.022)	(0.022)	(0.020)	(0.018)	(0.016)	(0.014)	(0.01)
Remittances	0.963	0.959*	0.957*	0.956*	0.957*	0.959*	0.963
	(0.019)	(0.020)	(0.021)	(0.021)	(0.021)	(0.020)	(0.019)
Child Age	3.247***	3.586***	3.793***	3.862***	3.793***	3.586***	3.247***
	(0.188)	(0.222)	(0.243)	(0.251)	(0.243)	(0.222)	(0.188)
Child Age^2	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.984	0.982	0.981	0.981	0.981	0.982	0.984
	(0.026)	(0.028)	(0.029)	(0.029)	(0.029)	(0.028)	(0.026)
Household Assets	1.184***	1.202***	1.212***	1.215***	1.212***	1.202***	1.184***
	(0.013)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.013)
Non-Hindu	0.526***	0.495***	0.479***	0.473***	0.479***	0.495***	0.526***
	(0.064)	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)	(0.064)
OBC	0.972	0.966	0.963	0.962	0.963	0.966	0.972
	(0.117)	(0.125)	(0.129)	(0.131)	(0.129)	(0.125)	(0.117)
SC & ST	0.788	0.773	0.764	0.762	0.764	0.773	0.788

Table B.18 (continued).

	(0.105)	(0.111)	(0.114)	(0.115)	(0.114)	(0.111)	(0.105)
Urban	1.410**	1.468**	1.501**	1.511**	1.501**	1.468**	1.410**
	(0.160)	(0.179)	(0.190)	(0.194)	(0.190)	(0.179)	(0.160)
Children under 6	0.551***	0.517***	0.499***	0.493***	0.499***	0.517***	0.551***
	(0.081)	(0.082)	(0.082)	(0.082)	(0.082)	(0.082)	(0.081)
6-14 year old	0.662***	0.637***	0.624***	0.620***	0.624***	0.637***	0.662***
	(0.073)	(0.076)	(0.077)	(0.078)	(0.077)	(0.076)	(0.073)
Ln (Debt)	1.023*	1.024*	1.025*	1.025*	1.025*	1.024*	1.023*
	(0.009)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.009)
Only Daughters	1.279	1.307	1.324	1.329	1.324	1.307	1.279
	(0.165)	(0.181)	(0.191)	(0.194)	(0.191)	(0.181)	(0.165)
Siblings	1.005	1.004	1.003	1.003	1.003	1.004	1.005
	(0.027)	(0.029)	(0.030)	(0.030)	(0.030)	(0.029)	(0.027)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.19: Non-literate to primary transition for sons: Odds ratio estimates under different assumptions of correlation between unobserved heterogeneity and father's migration status

Correlation between unobserved heterogeneity (u) and father's migration status	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.099***	1.107***	1.111***	1.113***	1.111***	1.107***	1.099***
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)
Migrant	482.7***	53.93***	6.107***	0.700	0.0812***	0.00954***	0.00114***
	(100.514)	(12.100)	(1.426)	(0.166)	(0.019)	(0.002)	(0.000)
Father Education × Migrant	1.017	1.019	1.021	1.021	1.021	1.019	1.017
	(0.036)	(0.039)	(0.041)	(0.041)	(0.041)	(0.039)	(0.036)
Mother Education	1.038**	1.041**	1.042**	1.042**	1.042**	1.041**	1.038**
	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
Remittances	0.966*	0.963*	0.962*	0.961*	0.962*	0.963*	0.966*
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Child Age	1.207***	1.225***	1.235***	1.238***	1.235***	1.225***	1.207***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.983	0.982	0.981	0.981	0.981	0.982	0.983
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)
Household Assets	1.091***	1.097***	1.101***	1.102***	1.101***	1.097***	1.091***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Non-Hindu	0.508***	0.484***	0.471***	0.467***	0.471***	0.484***	0.508***
	(0.048)	(0.050)	(0.050)	(0.051)	(0.050)	(0.050)	(0.048)
OBC	0.813	0.804	0.799	0.798	0.799	0.804	0.813
	(0.087)	(0.093)	(0.096)	(0.097)	(0.096)	(0.093)	(0.087)
SC & ST	0.941	0.940	0.940	0.940	0.940	0.940	0.941

Table B.19 (continued).

	(0.111)	(0.119)	(0.124)	(0.125)	(0.124)	(0.119)	(0.111)
Urban	1.687***	1.753***	1.791***	1.804***	1.791***	1.753***	1.687***
	(0.166)	(0.186)	(0.198)	(0.202)	(0.198)	(0.186)	(0.166)
Children under 6	0.653***	0.631***	0.619***	0.615***	0.619***	0.631***	0.653***
	(0.056)	(0.059)	(0.060)	(0.060)	(0.060)	(0.059)	(0.056)
6-14 year old	0.625**	0.614**	0.607**	0.605**	0.607**	0.614**	0.625**
	(0.098)	(0.103)	(0.105)	(0.106)	(0.105)	(0.103)	(0.098)
Ln (Debt)	1.008	1.009	1.009	1.009	1.009	1.009	1.008
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Only Sons	0.774**	0.759**	0.751**	0.748**	0.751**	0.759**	0.774**
	(0.073)	(0.077)	(0.080)	(0.080)	(0.080)	(0.077)	(0.073)
Siblings	0.750***	0.733***	0.723***	0.720***	0.723***	0.733***	0.750***
	(0.024)	(0.025)	(0.025)	(0.026)	(0.025)	(0.025)	(0.024)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.20: Primary to secondary transition for Sons: Odds ratio estimates under different assumptions of correlation between unobserved heterogeneity and father's migration status

Correlation between unobserved heterogeneity (u) and father's migration status	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.076***	1.084***	1.088***	1.090***	1.088***	1.084***	1.076***
	(0.010)	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.010)
Migrant	388.2***	42.89***	4.827***	0.552*	0.0642***	0.00759***	0.000914***
	(88.178)	(10.447)	(1.220)	(0.141)	(0.016)	(0.002)	(0.000)
Father Education × Migrant	1.088*	1.095*	1.099*	1.100*	1.099*	1.095*	1.088*
	(0.040)	(0.044)	(0.045)	(0.046)	(0.045)	(0.044)	(0.040)
Mother Education	1.068***	1.073***	1.076***	1.077***	1.076***	1.073***	1.068***
	(0.012)	(0.013)	(0.014)	(0.014)	(0.014)	(0.013)	(0.012)
Remittances	0.958**	0.955**	0.953**	0.953**	0.953**	0.955**	0.958**
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Child Age	1.638***	1.704***	1.742***	1.755***	1.742***	1.704***	1.638***
	(0.016)	(0.018)	(0.019)	(0.019)	(0.019)	(0.018)	(0.016)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.988	0.986	0.986	0.985	0.986	0.986	0.988
	(0.014)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.014)
Household Assets	1.139***	1.151***	1.158***	1.160***	1.158***	1.151***	1.139***
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
Non-Hindu	0.422***	0.391***	0.375***	0.370***	0.375***	0.391***	0.422***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
OBC	1.207*	1.219*	1.225*	1.227*	1.225*	1.219*	1.207*
	(0.109)	(0.118)	(0.124)	(0.125)	(0.124)	(0.118)	(0.109)
SC & ST	1.008	1.006	1.004	1.004	1.004	1.006	1.008

Table B.20 (continued).

	(0.099)	(0.106)	(0.110)	(0.111)	(0.110)	(0.106)	(0.099)
Urban	2.143***	2.276***	2.353***	2.379***	2.353***	2.276***	2.143***
	(0.182)	(0.208)	(0.223)	(0.228)	(0.223)	(0.208)	(0.182)
Children under 6	0.588***	0.564***	0.550***	0.546***	0.550***	0.564***	0.588***
	(0.053)	(0.055)	(0.056)	(0.056)	(0.056)	(0.055)	(0.053)
6-14 year old	0.703**	0.680**	0.667**	0.663**	0.667**	0.680**	0.703**
	(0.082)	(0.085)	(0.086)	(0.087)	(0.086)	(0.085)	(0.082)
Ln (Debt)	1.008	1.008	1.009	1.009	1.009	1.008	1.008
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Only Sons	0.666***	0.643***	0.631***	0.627***	0.631***	0.643***	0.666***
	(0.054)	(0.056)	(0.057)	(0.057)	(0.057)	(0.056)	(0.054)
Siblings	0.724***	0.704***	0.694***	0.690***	0.694***	0.704***	0.724***
	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.21: Secondary to post-secondary transition for Sons: Odds ratio estimates under different assumptions of correlation between unobserved heterogeneity and father's migration status

Correlation between unobserved heterogeneity (u) and father's migration status	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.142***	1.155***	1.163***	1.165***	1.163***	1.155***	1.142***
	(0.013)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.013)
Migrant	846.5***	98.41***	11.41***	1.318	0.152***	0.0174***	0.00199***
	(337.753)	(42.120)	(5.069)	(0.593)	(0.067)	(0.007)	(0.001)
Father Education × Migrant	0.876*	0.869*	0.865*	0.864*	0.865*	0.869*	0.876*
	(0.052)	(0.055)	(0.057)	(0.058)	(0.057)	(0.055)	(0.052)
Mother Education	1.096***	1.105***	1.109***	1.111***	1.109***	1.105***	1.096***
	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
Remittances	0.998	0.997	0.996	0.996	0.996	0.997	0.998
	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.018)
Child Age	3.146***	3.440***	3.617***	3.676***	3.617***	3.440***	3.146***
	(0.174)	(0.204)	(0.222)	(0.228)	(0.222)	(0.204)	(0.174)
Child Age^2	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	1.006	1.006	1.006	1.006	1.006	1.006	1.006
	(0.021)	(0.022)	(0.023)	(0.024)	(0.023)	(0.022)	(0.021)
Household Assets	1.164***	1.180***	1.189***	1.192***	1.189***	1.180***	1.164***
	(0.012)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.012)
Non-Hindu	0.664***	0.635***	0.619***	0.614***	0.619***	0.635***	0.664***
	(0.075)	(0.077)	(0.078)	(0.078)	(0.078)	(0.077)	(0.075)
OBC	1.111	1.121	1.127	1.129	1.127	1.121	1.111
	(0.120)	(0.131)	(0.137)	(0.139)	(0.137)	(0.131)	(0.120)
SC & ST	0.830	0.818	0.812	0.810	0.812	0.818	0.830

Table B.21 (continued).

	(0.100)	(0.106)	(0.110)	(0.111)	(0.110)	(0.106)	(0.100)
Urban	1.529***	1.595***	1.632***	1.644***	1.632***	1.595***	1.529***
	(0.155)	(0.174)	(0.186)	(0.189)	(0.186)	(0.174)	(0.155)
Children under 6	0.475***	0.446***	0.430***	0.425***	0.430***	0.446***	0.475***
	(0.066)	(0.066)	(0.067)	(0.067)	(0.067)	(0.066)	(0.066)
6-14 year old	0.796*	0.780*	0.771*	0.768*	0.771*	0.780*	0.796*
	(0.079)	(0.084)	(0.086)	(0.087)	(0.086)	(0.084)	(0.079)
Ln (Debt)	1.032***	1.035***	1.036***	1.037***	1.036***	1.035***	1.032***
	(0.008)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.008)
Only Sons	0.786*	0.767*	0.757*	0.754*	0.757*	0.767*	0.786*
	(0.080)	(0.084)	(0.086)	(0.087)	(0.086)	(0.084)	(0.080)
Siblings	0.822***	0.806***	0.797***	0.794***	0.797***	0.806***	0.822***
	(0.038)	(0.040)	(0.041)	(0.041)	(0.041)	(0.040)	(0.038)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.22: Non-literate to primary transition for daughters Odds ratio estimates assuming correlation between omitted confounder and father's migration status

Correlation between unobserved heterogeneity (u) and father's migration status	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.111***	1.119***	1.124***	1.126***	1.124***	1.119***	1.111***
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Migrant	458.3***	54.17***	6.442***	0.771	0.0927***	0.0112***	0.00136***
	(99.044)	(12.628)	(1.563)	(0.189)	(0.022)	(0.003)	(0.000)
Father Education × Migrant	1.066	1.073	1.077	1.078	1.077	1.073	1.066
	(0.044)	(0.048)	(0.050)	(0.051)	(0.050)	(0.048)	(0.044)
Mother Education	1.062***	1.066***	1.068***	1.068***	1.068***	1.066***	1.062***
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Remittances	0.956**	0.952**	0.950**	0.949**	0.950**	0.952**	0.956**
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Child Age	1.193***	1.210***	1.220***	1.223***	1.220***	1.210***	1.193***
	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.006)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.992	0.991	0.990	0.990	0.990	0.991	0.992
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Household Assets	1.109***	1.118***	1.122***	1.124***	1.122***	1.118***	1.109***
	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)
Non-Hindu	0.540***	0.517***	0.505***	0.501***	0.505***	0.517***	0.540***
	(0.054)	(0.056)	(0.057)	(0.057)	(0.057)	(0.056)	(0.054)
OBC	0.617***	0.597***	0.585***	0.582***	0.585***	0.597***	0.617***
	(0.072)	(0.074)	(0.076)	(0.076)	(0.076)	(0.074)	(0.072)
SC & ST	0.797	0.785	0.779	0.777	0.779	0.785	0.797

Table B.22 (continued).

	(0.102)	(0.108)	(0.112)	(0.113)	(0.112)	(0.108)	(0.102)
Urban	1.771***	1.857***	1.907***	1.924***	1.907***	1.857***	1.771***
	(0.187)	(0.211)	(0.225)	(0.230)	(0.225)	(0.211)	(0.187)
Children under 6	0.564***	0.539***	0.526***	0.521***	0.526***	0.539***	0.564***
	(0.050)	(0.052)	(0.053)	(0.053)	(0.053)	(0.052)	(0.050)
6-14 year old	0.477***	0.457***	0.445***	0.441***	0.445***	0.457***	0.477***
	(0.081)	(0.083)	(0.083)	(0.084)	(0.083)	(0.083)	(0.081)
Ln (Debt)	0.993	0.992	0.992	0.992	0.992	0.992	0.993
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Only Daughters	1.055	1.065	1.071	1.073	1.071	1.065	1.055
	(0.118)	(0.128)	(0.134)	(0.136)	(0.134)	(0.128)	(0.118)
Siblings	0.977	0.975	0.974	0.974	0.974	0.975	0.977
	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.018)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.23: Primary to secondary transition for daughters: Odds ratio estimates under different assumptions of correlation between unobserved heterogeneity and father's migration status

Correlation between unobserved heterogeneity (u) and father's migration status	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.093***	1.102***	1.108***	1.109***	1.108***	1.102***	1.093***
	(0.011)	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)	(0.011)
Migrant	342.6***	39.37***	4.609***	0.548*	0.0663***	0.00815***	0.00102***
	(78.516)	(9.697)	(1.179)	(0.142)	(0.017)	(0.002)	(0.000)
Father Education × Migrant	1.088*	1.096*	1.101*	1.102*	1.101*	1.096*	1.088*
	(0.043)	(0.047)	(0.049)	(0.049)	(0.049)	(0.047)	(0.043)
Mother Education	1.081***	1.088***	1.092***	1.094***	1.092***	1.088***	1.081***
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)
Remittances	0.949***	0.945***	0.942***	0.941***	0.942***	0.945***	0.949***
	(0.014)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.014)
Child Age	1.625***	1.691***	1.729***	1.741***	1.729***	1.691***	1.625***
	(0.017)	(0.019)	(0.020)	(0.021)	(0.020)	(0.019)	(0.017)
Child Age^2	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***	0.999***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.988	0.987	0.986	0.986	0.986	0.987	0.988
	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
Household Assets	1.152***	1.166***	1.174***	1.176***	1.174***	1.166***	1.152***
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Non-Hindu	0.491***	0.460***	0.444***	0.439***	0.444***	0.460***	0.491***
	(0.045)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.045)
OBC	0.884	0.869	0.861	0.858	0.861	0.869	0.884
	(0.087)	(0.092)	(0.094)	(0.095)	(0.094)	(0.092)	(0.087)
SC & ST	0.996	0.990	0.987	0.986	0.987	0.990	0.996

Table B.23 (continued).

	(0.107)	(0.114)	(0.118)	(0.120)	(0.118)	(0.114)	(0.107)
Urban	2.339***	2.509***	2.608***	2.641***	2.608***	2.509***	2.339***
	(0.219)	(0.252)	(0.272)	(0.279)	(0.272)	(0.252)	(0.219)
Children under 6	0.424***	0.393***	0.377***	0.372***	0.377***	0.393***	0.424***
	(0.038)	(0.038)	(0.037)	(0.037)	(0.037)	(0.038)	(0.038)
6-14 year old	0.583***	0.555***	0.540***	0.535***	0.540***	0.555***	0.583***
	(0.081)	(0.082)	(0.083)	(0.083)	(0.083)	(0.082)	(0.081)
Ln (Debt)	0.997	0.997	0.996	0.996	0.996	0.997	0.997
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
Only Daughters	1.219*	1.240*	1.252*	1.256*	1.252*	1.240*	1.219*
	(0.122)	(0.134)	(0.140)	(0.142)	(0.140)	(0.134)	(0.122)
Siblings	0.926***	0.921***	0.918***	0.917***	0.918***	0.921***	0.926***
	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.016)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5

Table B.24: Secondary to post-secondary transition for daughters: Odds ratio estimates under different assumptions of correlation between unobserved heterogeneity and father's migration status

Correlation between unobserved heterogeneity (u) and father's migration status	-0.6	-0.4	-0.2	0	0.2	0.4	0.6
Father Education	1.107***	1.117***	1.123***	1.125***	1.123***	1.117***	1.107***
	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)
Migrant	344.6***	39.24***	4.572*	0.543	0.0658***	0.00812***	0.00103***
	(186.984)	(22.862)	(2.768)	(0.333)	(0.040)	(0.005)	(0.001)
Father Education × Migrant	1.028	1.031	1.033	1.034	1.033	1.031	1.028
	(0.076)	(0.082)	(0.085)	(0.086)	(0.085)	(0.082)	(0.076)
Mother Education	1.140***	1.152***	1.159***	1.161***	1.159***	1.152***	1.140***
	(0.016)	(0.017)	(0.018)	(0.018)	(0.018)	(0.017)	(0.016)
Remittances	0.963	0.959*	0.957*	0.956*	0.957*	0.959*	0.963
	(0.019)	(0.020)	(0.021)	(0.021)	(0.021)	(0.020)	(0.019)
Child Age	3.247***	3.586***	3.793***	3.862***	3.793***	3.586***	3.247***
	(0.188)	(0.222)	(0.243)	(0.251)	(0.243)	(0.222)	(0.188)
Child Age^2	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***	0.997***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months of Migration	0.984	0.982	0.981	0.981	0.981	0.982	0.984
	(0.026)	(0.028)	(0.029)	(0.029)	(0.029)	(0.028)	(0.026)
Household Assets	1.184***	1.202***	1.212***	1.215***	1.212***	1.202***	1.184***
	(0.013)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.013)
Non-Hindu	0.526***	0.495***	0.479***	0.473***	0.479***	0.495***	0.526***
	(0.064)	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)	(0.064)
OBC	0.972	0.966	0.963	0.962	0.963	0.966	0.972
	(0.117)	(0.125)	(0.129)	(0.131)	(0.129)	(0.125)	(0.117)
SC & ST	0.788	0.773	0.764	0.762	0.764	0.773	0.788

Table B.24 (continued).

	(0.105)	(0.111)	(0.114)	(0.115)	(0.114)	(0.111)	(0.105)
Urban	1.410**	1.468**	1.501**	1.511**	1.501**	1.468**	1.410**
	(0.160)	(0.179)	(0.190)	(0.194)	(0.190)	(0.179)	(0.160)
Children under 6	0.551***	0.517***	0.499***	0.493***	0.499***	0.517***	0.551***
	(0.081)	(0.082)	(0.082)	(0.082)	(0.082)	(0.082)	(0.081)
6-14 year old	0.662***	0.637***	0.624***	0.620***	0.624***	0.637***	0.662***
	(0.073)	(0.076)	(0.077)	(0.078)	(0.077)	(0.076)	(0.073)
Ln (Debt)	1.023*	1.024*	1.025*	1.025*	1.025*	1.024*	1.023*
	(0.009)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.009)
Only Daughters	1.279	1.307	1.324	1.329	1.324	1.307	1.279
	(0.165)	(0.181)	(0.191)	(0.194)	(0.191)	(0.181)	(0.165)
Siblings	1.005	1.004	1.003	1.003	1.003	1.004	1.005
	(0.027)	(0.029)	(0.030)	(0.030)	(0.030)	(0.029)	(0.027)

Notes: This table reports exponentiated coefficients with robust standard errors clustered by PSU reported in parentheses. *,**,*** represent statistical significance at 5%, 1% and 0.1% confidence levels, respectively. The scenarios assume that the unobserved heterogeneity is normally distributed and its effect size (coefficient) βu is 2.5