Intergenerational Educational Persistence among Daughters: Evidence from India

Mehtabul Azam

DISCUSSION PAPER SERIES

I Z A

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Intergenerational Educational Persistence among Daughters: Evidence from India

Mehtabul Azam
Oklahoma State University
and IZA

ABSTRACT

Intergenerational Educational Persistence among Daughters: Evidence from India

We examine educational transmission between fathers (mothers) and daughters in India for daughters born during 1962-1991. We find that educational persistence, as measured by the regression coefficient of father's (mother's) education as a predictor of daughter's education, has declined over time. However, the correlation between educational attainment of daughters and fathers (mothers), another commonly used measure of persistence, suggests only a marginal decline. Further decomposing the intergenerational correlation, we find that although persistence has declined at the lower end of the fathers' (mothers') educational distribution, it has been compensated by an increase in persistence at other parts of fathers' (mothers') educational distribution. We also find that "Equality of Opportunity" remains an elusive goal for India. Not only the probability of a daughter attaining senior secondary or above education (top end of educational distribution) is positively associated with father's (mother's) education levels, the gaps in those probabilities do not show any sign of convergence. Similarly, there is no sign of any convergence in the probability of a daughter attaining senior secondary or above education even with the same level of father's (mother's) education between Higher Hindu Castes' daughters and daughters belonging to disadvantaged groups such as Other Backward Castes or Scheduled Castes/Tribes.

Corresponding author:

Mehtabul Azam
326 Business Building
Spears School of Business
Oklahoma State University
Stillwater, OK 74078
USA
E-mail: mazam@okstate.edu

1   Introduction

The empirical literature on intergenerational mobility in developed countries has predominantly focused on sons. Only a few studies do examine intergenerational transmission between fathers and daughters (see, for example, DiPrete and Grusky 1990, Chadwick and Solon, 2002). Compared to developed countries, the intergenerational mobility in developing countries remains largely an under-researched area although there has been an increasing focus on the topic with the availability of new datasets that contain parents' information. However, the literature in developing countries too has primarily focused on father-son transmission. One of reason for this is non-availability of datasets that contain information on fathers' for women.

The concerns regarding the equality of opportunities are growing in developing countries.[1] Equality of opportunity is considered a key condition for a society to ensure distributional justice (Rama et al., 2015), and education is perhaps the most important policy instrument in the hands of policymakers in this regard.[2] For example, Stiglitz (2012, P. 275) notes "Opportunity is shaped, more than anything else, by access to education." One of the key roles of publicly provided education in many countries including India has been to increase access to education.[3] Intergenerational persistence in education can undermine the notion of equality of opportunity.

The intergenerational persistence in father (mother)-daughter is quite an important issue per se, however it is comparatively more important in the Indian context because of following

---

[1]As such this issues is also very important in the US. For example, President Obama's second Inaugural Address reaffirmed America's commitment to the dream of equality of opportunity: "We are true to our creed when a little girl born into the bleakest poverty knows that she has the same chance to succeed as anybody else, because she is an American; she is free, and she is equal, not just in the eyes of God but also in our own."

[2]Equality of opportunity can be broadly defined as: those who have the same talent and ability and have the same willingness to use them should have the same prospects of success regardless of their initial circumstances (Rawls 1971).

[3]Achievement of universal primary education by 2015 was one of the eight Millennium Development Goals (MDGs) established by United Nations in 2000. The Indian Parliament passed Right to Education Act in 2009 that makes education a fundamental right of every child between the ages of 6 and 14 and specifies minimum norms in elementary schools.

reasons. First, the notion of family background (economic and caste) determining destiny is quite pervasive in India. Second, there exists a strong son preference in Indian society, and the existing evidence suggests pro-male bias in educational investment (Kingdon, 2005). Third, there exists considerable inequality in India. For example, the Gini index for consumption calculated from 2004-05 IHDS was at 0.34, while the Gini index for income was estimated to be 20 points higher at 0.54. This is less than the income inequality observed in South Africa, but it is still more than in Mexico (Rama et al., 2015). Existing evidence suggests that countries with greater inequality of incomes also tend to be countries in which a greater fraction of economic advantage and disadvantage is passed on between parents and their children (Corak, 2013).

In this paper, we examine the father (mother)-daughter educational persistence over time in India utilizing a recently available nationally representative India Human Development Survey (IHDS)-2, 2011-12. The IHDS-2 has a separate women module that asks detailed questions from two women in age 15-49 per household. This helps us to identify fathers' (mothers') information for about 86 (88) percent of women in age 20-49. Dividing those into birth cohorts, we first examine the evolution of two commonly used measure of persistence: intergenerational regression coefficient and correlation coefficient. To reconcile the observed differences in trends using these two measures, we further decompose the correlation coefficient.

The findings of the paper are as follows. First, when compared to the birth cohort 1962-66, there is a fall in the intergenerational educational persistence, as measured by the regression coefficient of fathers' (mothers') schooling as a predictor of daughters' schooling in the recent birth cohort 1987-91, implying less persistence for more recent cohort in India. Importantly, another commonly used measure of persistence, correlation between fathers' (mothers') and daughters' years of schooling, presents a much sober picture of only a marginal decline in persistence. By decomposing the correlation, we find that persistence at the lower end of the fathers'(mothers') educational distribution has declined (father/mother being

2

below primary); however, the persistence has increased in other parts of fathers' (mothers') educational distribution resulting in an overall steady trend in the correlation coefficient. Second, we find that "Equality of Opportunity" remains an elusive goal for India. Not only the probability of a daughter attaining senior secondary or above education (top end of educational distribution) is positively associated with father's education levels, the gaps in those probabilities do not show any sign of convergence. Similarly, the probability of a daughter attaining senior secondary or above education is higher for Higher Hindu Castes' daughters irrespective of parental education. Importantly, there is no sign of any convergence in the probability of a daughter attaining senior secondary or above education even with the same level of father's (mother's) education between Higher Hindu Castes' daughters and daughters belonging to disadvantaged groups such as Other Backward Castes or Scheduled Castes/Tribes.

The remainder of the paper is organized as follows. Section 1.1 presents a brief review of the literature on the intergenerational mobility in educational attainment in India, and places our paper in the existing literature. Section 2 discusses the data, Section 3 outlines the analytical framework underlying the empirical analysis. Section 4 presents the results and Section 5 concludes.

1.1   Related Literature

Literature on intergenerational economic mobility in developed countries mostly focuses on intergenerational correlation between fathers' and sons' incomes. Solon (1999), Black and Devereux (2011), and Blanden (2013) provide excellent survey of literature in developed countries. Hertz et al. (2007) study trends in intergenerational transmission of education for a sample of 42 countries. Daouli et al. (2010) examine educational transmission for Greek women.

The literature on intergenerational persistence in India has primarily focused on educa-

tional persistence and construct their parents and children sample through co-resident from cross-sectional data. Jalan and Murgai (2008) investigate educational mobility for both men and women in the age group 15-19 using 1992-93 and 1998-99 National Family Health Survey (NFHS) data. They rely on co-resident to identify parents' education. Similarly, Maitra and Sharma (2009) use the IHDS-2005, and explore the effect of parental education (both father and mother) on years of schooling of children, identifying children-parent pairs if they both reside in the same household. Hnatkovskay et al. (2013) use five rounds of National Sample Survey (NSS) that covers the period 1983-2005, to analyze intergenerational persistence in occupational choices, educational attainment and wages between fathers and sons. They also rely on co-resident to identify fathers' information. Emran and Shilpi (2015) examine correlation and sibling correlation in 16-27-year olds in 1992-93 and 2005-06 NFHS data. They use co-resident to identify parents' education. They find stagnant educational persistence over the two cross-sections using the correlation coefficient.

In contrast to the abovementioned studies on India, Azam and Bhatt (2015) do not rely on co-resident. Their sample include fathers' information for all the adult men. They explore transmission of education between fathers and sons starting from 1940 to 1985. They find that intergenerational educational persistence in India, as measured by the regression coefficient of fathers' education as a predictor of schooling in the next generation, has decreased significantly across birth cohorts in last 45 years. However, they do no find such a trend in the estimated correlation between father-son educational attainments. They further decompose the correlation and find that the decline in correlation at the lower end of fathers' education distribution is offset by the increase at the top end of fathers' education distribution. They also find a significant difference in the probability of achieving senior secondary or above education based on fathers' education levels. Moreover, they find no evidence for convergence in the probability of a son achieving senior secondary or above education conditional on father's education between Higher Hindu Castes versus others social groups.

Our paper extends Azam and Bhatt (2015) for the case of daughters. Azam and Bhatt (2015) use IHDS 2004-05 data, which although facilitated identification of fathers for adult men, however, it does not contain same information for adult women. The recently released IHDS 2011-12 data make the study of persistence in daughters' feasible (see Data Section for details).

2   Data

We use recently available India Human Development Survey-2 (IHDS-2), 2011-2012. IHDS-2 (Desai and Vanneman, 2015) is jointly conducted by National Council of Applied Economic Research and University of Maryland. One of the major problems faced by researchers interested in the study of intergenerational persistence in developing countries is non-availability of long panel data that help to identify parents' information. Researchers have used co-resident (parents and child residing in the same household at the time of survey) in cross-sectional data to identify parents' information. Azam and Bhatt (2015) show that this condition helps to identify fathers' information for less than a third of adult male population (20-65), and most of those adult males belong to 20-30 age group. Moreover, co-resident condition does not help much in the case of women in developing country context, particularly India. Majority of married women in India reside in different households than their parents, and household surveys typically collect information on members residing in the same household (through household roster) at the time of survey. For example, we are able to identify co-resident father only for 10.5 (20.5) percent of women in age group 20-49 (20-30) in IHDS-2.

This data constraint is relaxed to a large extent in the IHDS-2. IHDS-2 has a separate women module that collects information of parents for two women in each household in age 15-49.[4] This helps us to identify fathers' (mothers') education for 86 (88) percent of

---

[4]The IHDS-1 conducted in 2004-05 also contains a separate women module that asks detailed questions

women in age group 20-49 (see Table 1 for details). We chose the lower age limit at 20 as majority of individuals in India finish their college (about 15 years of education) around this age, and the upper age limit is driven by availability of parental information. Since our survey is from 2011-12, this implies we have data on daughters born between 1962 and 1991. We divide our sample into six five year birth cohorts: 1962-66, 1967-71, 1972-76, 1977-81, 1982-86, and 1987-91.[5] To examine the educational persistence among social groups, we further divide our sample in four social groups: Higher Hindu Caste (HHC), Other Backward Caste (OBC), Scheduled Caste/Tribe (SC/ST), and Muslims. SC/STs are historically disadvantaged groups in India, and have enjoyed affirmative policies in education and employment since the independence. OBCs were given reservation in employment in 1993.[6] Muslims are the largest minority religious group in India, and according to the Government of India (2006), their performance on many economic and education indicators are comparable to SC/STs. There exist certain differences among STs and SCs, however, because of small sample sizes of STs after dividing the data in cohorts, we group SCs and STs together.

We measure the economic/social status through years of schooling. Although, income (occupation) remains more popular measures of economic/social status in the economics (sociology) literature in developed countries, education is probably more suited in developing countries context for daughters, especially for India. The female labor force participation (LFPR) has been abysmally low in India. For example, in 2011-12, LFPR among women in age 15-59 was only 24.7 percent compared to 82 percent among men (Government of India,

---

from one ever married women in age 15-49 per household. However, IHDS-1 women's module does not contain parental information questions.

[5]Our most recent birth cohort 1987-1991 represents age 20-24 in the data. Including this cohort raises the concern that some of the daughters might still be in school and has not completed the maximum schooling. We find that about 16 percent of the daughters in 20-24 age group were still in school and has not completed the maximum 16 years of education. However, most of them has completed 12 years or more. Only 0.2 percent of 20-24 were still in school and have not completed the 12 years of education. Although this potentially introduce a marginal downward bias in the IGRC, however, in the rest of our analysis we have combined 12-16 years of education as senior secondary or above.

[6]Beteille (2002) provides a useful discussion on the caste-system and affirmative action in India.

6

2013). In addition majority of those working women are self-employed for whom no wages are reported in household survey datasets. Given the scarceness of information on income and occupation for daughters, education remains a popular choice as a measure of economic status in developing countries. Moreover, there are several advantages of using education as a measure of economic status in developing countries. First, on the measurement side, education is less prone to serious errors than earnings. Second, since most individuals complete their education by early or mid-twenties, life cycle biases are unlikely to bias estimation when compared with earnings. Finally, there is a vast literature that shows that higher education is associated with higher earnings, better health, and other economic outcomes (see Black and Devereux, 2011), rendering a measure of intergenerational transmission based on education a reasonable proxy for mobility in overall economic status.

The years of schooling is reported as a continuous variable in our data, and varies from 0 to 16, with 0 representing illiterate and 16 representing above bachelor degree. In the literature, parental education is proxy by either father's education, or the maximum of father's or mother's education, or the average of both parents education. In our analysis, we use father's years of schooling to proxy for parents' education.[7] In our sample, fathers have either the same or more education for about 94.4 percent of daughters. For 40.2 percent of daughters in our sample, father has more education than mothers. Father has the same education as mothers for 54.2 percent of daughters'. Interestingly, among the daughters who have similarly educated fathers and mothers, more than 90 percent of those fathers and mothers are illiterate. Nevertheless, we also present the results using mothers' years of schooling as proxy for parents education in an online appendix, and overall conclusions remain same.

---

[7]It is not a priori clear whether one should include spousal education as an additional explanatory variable. Without the inclusion of the partner's schooling, the effect of parental schooling as it is estimated represents both the direct transfer from the given parent and the indirect transfer from the other parent, which is due to assortative mating effects. If we are interested in the schooling of the children, we should not care whether parental schooling effects run through assortative mating or something else, and we can estimate separate regressions for mothers and fathers, without controlling for the spouses' schooling (Holmlund et al., 2011).

Table 2 presents the descriptive statistics of our sample. The average years of schooling for daughters has been increasing over time. For example, the daughters born during 1962-66 attended on average of 3.53 years of school, while daughters born during 1987-91 attended about 8.51 years of school on average. This steady increase has been observed among all social groups. Similarly, the average education of fathers and mothers also has improved over time. There are few interesting facts observed in the data. First, fathers tend to have much higher educational attainment compared to mothers among each birth cohort (Table 1). This is true among all social groups. This is not surprising given the patriarchal nature of Indian society. Second, there is a significant advantage witnessed by HHC daughters in terms of parents' education compared to other social groups. For example, the average education of fathers for HHC daughters born during 1962-66 is more than four times higher than SC/ST daughters. This disadvantage of SC/ST daughters has declined over time, however a significant gap remains: the average education of fathers for HHC daughters born during 1986-91 is about 2.3 times higher than SC/ST daughters. Similar is the case for mothers also. Not surprisingly, a considerable advantage of HHC is also witnessed in daughters' education.

3   Analytical Framework

To capture the intergenerational transmission of education, we estimate the following regression:

$$S_i^d = \alpha + \beta S_i^f + \epsilon_i \tag{1}$$

where S_i^d and S_i^f represent the education of daughter i and education of her father, respectively. ϵ_i is an error term and β is the parameter of interest. The OLS estimate of β is reported as one of the measure of intergenerational persistence of educational attainment. The $\hat{\beta}$ is given by:

$$\hat{\beta} = \frac{\sigma_{df}}{\sigma_f^2} = \rho_{df} \frac{\sigma_d}{\sigma_f} \tag{2}$$

where σ_d and σ_f are the standard deviations of daughters' and fathers' schooling, σ_{df} is the covariance between daughters' and fathers' schooling, and ρ_{df} is the correlation between daughters' and fathers' schooling. To ensure that the evolution of $\hat{\beta}$ is not entirely driven by the evolution of $\frac{\sigma_d}{\sigma_f}$, we also normalized the years of schooling of daughters and fathers by the corresponding standard deviations and estimate the following equation:

$$\frac{S_i^d}{\sigma_d} = \delta + \rho \frac{S_i^f}{\sigma_f} + \epsilon_i \tag{3}$$

As argued by Checchi et al. (2008) the main difference between the β coefficient in equation (1) and ρ coefficient in equation (3) is that the former by considering the ratio of variances, takes into account a change of inequality of educational outcomes in daughters and fathers generations, providing a relative measure of intergenerational mobility. The latter provides an absolute measure of intergenerational transmission, i.e. cleansed from possible evolution of the distribution of educational attainments, for instance, due to school reforms that increased the average schooling of the population, reducing its variance. The changes in the relative standard deviations will cause both measures to evolve differently over time, and evidence (Hertz et al., 2007; Azam and Bhatt, 2015) shows that in several countries β and ρ behave differently. In our empirical results we report both the intergenerational regression coefficient (IGRC) ($\hat{\beta}$) and the intergenerational correlation coefficient (IGC) (ρ) across different cohorts. It is common among economists to refer to both intergenerational regression coefficients and correlation coefficients as inverse measures of intergenerational mobility (Solon 1999).

We estimate equation (1) and equation (3) separately for six five-year cohorts starting with 1962. Note that the interpretation of β and ρ is descriptive and not causal. However, assuming that the factors potentially biasing the persistence estimates are time invariant, the evolution of these estimates over time can be reliably inferred from the above approach (Checchi et al. 2008).

Following Checchi et al. (2013) and denoting the normalized schooling (by their corresponding standard deviations) daughters and fathers with d and f, we rewrite the correlation coefficient as:

$$\rho = \int \underbrace{(d - E(d))(f - E(f))}_{A} \underbrace{P(d/f)}_{B} \underbrace{P(f)}_{C} \tag{4}$$

Thus, ρ can change over time because of changes in the dispersion of daughters' and fathers' (standardized) education around their respective means (term A), because of changes in daughters' educational attainment conditional on fathers' education (term B), or because of changes in the unconditional distribution of fathers' education (term C). Checchi et al. (2013) suggest that term B should be the policy-relevant indicator of intergenerational persistence as changes in term A can be due to uniform convergence towards higher levels of education. In addition, as countries develop, one would expect an increase in the level of education of fathers across generations.

To explore the stability of correlation coefficients further, we decompose the correlation coefficient using the empirical analogue of equation (4) (Checchi et al 2013):

$$\rho = \sum_{d,f} \underbrace{(d - E(d))(f - E(f))}_{A} \underbrace{P(d/f)}_{B} \underbrace{P(f)}_{C} \tag{5}$$

where $d, f = 0, 1, 2, ..., 15, 16$ and thus $\hat{\rho}$ for each cohort is the sum of 289 elements.

4   Results

Table 3 presents estimates for both measures of persistence for six five-year birth cohorts. As is evident from Table 3, a one year difference in fathers' education has been associated with a 0.627 (0.535) year difference in daughters' education for daughters born during 1962-1966 (1987-1991). Thus there is 9 point decline in estimated regression coefficient over three decades. A Chi-square test of equality of $\hat{\beta}$ for cohorts 1962-1966 and 1987-1991 rejects the

null (p-value=0.000). A Chi-square test of equality of $\hat{\beta}$ for successive cohorts rejects the null for 1962-66 vs. 1967-71, 1972-76 vs. 1977-81, 1977-81 vs. 1982-86, and 1982-86 vs. 1987-91 at 5% significance level. However, we are unable to reject the null of equality of $\hat{\beta}$ between birth cohort 1967-71 and birth cohort 1972-76. Although there is no discernible trend in IGRC over the entire period, a steady decline in IGRC is observed after mid-1970s. The IGC shows a marginal decline of 1.3 points between 1962-66 and 1987-91. A Chi-square test of equality of $\hat{\rho}$ for cohorts 1962-1966 and 1987-1991 rejects the null (p-value=0.022). A Chi-square test of equality of $\hat{\rho}$ for successive cohorts rejects the null for 1962-66 vs. 1967-71, 1967-71 vs. 1972-76, 1972-76 vs. 1977-81, and 1977-81 vs. 1982-86. However, we are unable to reject the null of equality of $\hat{\rho}$ between birth cohort 1982-86 and birth cohort 1987-91. In contrast to IGRC, the persistence based on IRC does not show a decline in the 1980s. This results are very much in line with to Azam and Bhatt (2015) finding for men. They find a decline in IGRC, however no definite trend in IGC.

Online appendix Table A1 presents similar results for mother-daughter transmission. We find a definite negative trend in the IGRC over the entire period, however no definite trend in the IGC over the entire period. The IGRC fall from implausibly high 1.030 for the 1962-66 birth cohort to 0.640 for the 1987-91 birth cohort. The very high IGRC estimates in the earlier cohorts are driven by a large number of zeros in mothers' education.

Table 3 also presents the standard deviation (SD) in daughters and fathers years of schooling. The SD in daughters' years of schooling has been increasing throughout except for the recent 1987-91 birth cohort. Similarly, SD in fathers' schooling has been increasing over time. Except for the most recent cohort, the variance of daughters' schooling is greater than the variance of fathers' schooling. This implies the ratio of the SD of fathers' years of schooling to that of daughters' years of schooling will be less than one because of which $\hat{\rho}$ is less than $\hat{\beta}$ for all cohorts except the 1987-91 birth cohort.

Table 4 presents decomposition of $\hat{\rho}$ grouped by stages of schooling attended by fathers

and daughters.[8] Line 31 of Table 4 reports the correlation coefficient $\hat{\rho}$, which is the sum of each combination of daughter's and father's education. Line 6 shows the total contribution of daughters with uneducated fathers to the intergenerational correlation coefficient. This group accounts for a large part of the correlation in each cohort but its weight declined from about 66 percent to 38 percent over 1962-67 and 1987-91 birth cohort. This is a natural consequence of increase in average education over time starting with a largely uneducated society.

However, this decline in correlation at the lower end of fathers' education distribution is compensated by an increase at the other parts of the fathers' educational distribution. As evident from lines (12), (18), (24), and (30), the contribution of sons of whose fathers have attended primary, middle school, or secondary schools has increased steadily across cohorts. This leads to a steady trend in the overall correlation coefficient. The total contribution of daughters with secondary attended father to the intergenerational correlation coefficient increased from about 10 percent to 24 percent over 1962-67 and 1987-91 birth cohort.

Online Appendix Table A2 presents similar results for mother-daughter transmission. The overall findings are similar to the findings for father-daughter educational transmission. The total contribution of daughters with uneducated fathers to the intergenerational correlation coefficient falls from 83 percent to 61 percent over 1962-67 and 1987-91 birth cohort. This decline in correlation at the lower end of mothers' education distribution is compensated by an increase at other parts of the mothers' educational distribution.

Checchi et al. (2008, 2013) argue that term B of equation (4) is the correct measure for analyzing the transmission of education: a system would achieve equality of opportunity if the probability of obtaining a particular degree for the daughter was independent of the father's educational achievement. To investigate the persistence in education, or term B, we collapse our years of schooling into stages of schooling achieved by daughters and fathers.

---

[8]Note that here stage of schooling implies attended those stage. For example, a person will be classified as attended primary school if he/she has completed 1-5 years of education.

We group the years of schooling into five achievement levels: years of schooling 0-4: below primary, 5-7: primary, 8-9: middle, 10-11: secondary, and 12-16: senior secondary or above. Figure 1 presents the probability of a daughter achieving either below primary or senior secondary or above education conditional on her father's education.[9] Left panel of Figure 1 plots the probability of a daughter being below primary conditional on different levels of her father's education. As expected, with the expansion of primary education, the probability of the daughter being below primary declines over time with the highest decline witnessed by daughters of fathers with either below primary or primary education. Our most recent birth-cohort is 1987-91, which implies that the daughters born during 1987-91 attended primary schools in late 1990s and early 2000s, and not in the last decade. With the near universalization of primary education in recent years, one should expect the probability of below primary education approaches to zero irrespective of the father's education level for daughters born in 1990s and 2000s.

Right panel of Figure 1 presents the probability of a daughter achieving senior secondary or above education. Here, the differences are quite striking, and most importantly there is no evidence of convergence among daughters of fathers with different levels of education. The probability of a daughter attaining senior secondary or above education increases with the level of father's education. More importantly, there remains a considerable gap in probability between top end and bottom end of fathers' education distribution. For example, the gap in probability of a daughter attaining senior secondary or above education between a daughter of senior secondary or above educated father and a daughter of below primary educated father is about 0.5 points for the 1962-67 birth cohort, and that increases to about 0.6 points for the 1987-91 birth cohort. This results are in line with the results reported in Azam and Bhatt (2015) for men. For example, Azam and Bhatt (2015) finds that the probability of

---

[9]We only present results for the top and bottom education levels as the convergence in probability of achieving middle education levels may be misleading as this convergence may be achieved by increase in probability of achieving middle levels of education by daughters with low educated fathers, while decline in probability of achieving middle levels of education by daughters with high educated fathers as they achieve more higher education.

achieving a senior secondary or above education for someone born to a senior secondary or above educated father in 1940-1945 is about 0.75 points higher than for someone born to an illiterate father in the same period, and there has been no decline in this gap over time. Overall, one may conclude that the probability of achieving higher education is definitely associated with the family background, and there is not much improvement over time.

4.1 Educational persistence by social groups

In Table 5, we present the IGRC and IGC by caste for each of the five-year birth cohorts. There is no distinct trend over the entire period across all four social groups in both measures of persistence. The IGRC is lower in the 1987-91 birth cohort compared to the 1962-66 birth cohort for HHC and OBC, while higher for SC/ST and Muslims. Interestingly, the IGC is also higher (lower) for SC/ST (HHC) in the 1987-91 birth cohort compared to the 1962-66 birth cohort. For Muslims and OBC, the IGC provide a conflicting evidence. The IGC is lower (the same) for Muslims (OBC) in the 1987-91 birth cohort compared to the 1962-66 birth cohort.

The SD in daughters schooling shows a declining trend for HHCs, however, an increasing trend in SD is witnessed for rest of the social groups. This is because of largely uneducated daughters to start with. Similarly, the SD in fathers' education has a positive trend for all social groups except HHC. For HHC, the SD in fathers' schooling shows a declining trend in the late 1970s and 1980s after increasing in the 1960s.

Based on both measures of persistence, no clear cut picture emerges. To explore the issue further, we turn our focus to term B of equation (4). Unlike the regression and correlation coefficients which are not suitable for inter-group comparisons based on stratification, term B of equation (4) can be used to compare groups (Checchi et al., 2013).[10]

Figure 2 presents the probability of a daughters achieving below primary education con-

---

[10]This is because the estimated persistence for any group only provides an estimate of the rate to regression to the mean for that particular group and not for the overall education distribution.

ditional on father's education (term B of equation (4)) for different caste groups.[11] It is evident that daughters belonging to HHC have the lowest probability of remaining illiterate/below primary irrespective of fathers' education. The gap between HHCs and other social groups is substantial. While the probability of a daughter being illiterate/below primary is more or less similar for SC/STs, OBCs, and Muslims, this probability is much lower for HHCs. Importantly, the probability of being illiterate/below primary declined over time for all social groups. Nevertheless, the probability of getting education is associated with family background, as the probability of a daughter being illiterate/below primary declines as fathers' education level increase irrespective of social group.

Figure 3 plots the probability of attaining senior secondary or above education conditional on fathers' education for different social groups. The daughters of HHCs have the highest probability of attaining senior secondary or above education for the same level of fathers' education. Although not plotted in the figure for clarity, the 95% confidence bands for HHCs do not overlap with the rest of the social groups, while the confidence bands for rest of the groups show substantial overlap. What is striking is that the probability of a daughter attaining senior secondary or above education for Muslims is either similar or marginally worse than that of daughters belonging to SC/ST group.[12] These results for women are very similar to results for men presented in Azam and Bhatt (2015) who find that Muslim men have a lower probability of achieving secondary and above education for each level of father's education, whereas HHC men have a significantly higher probability of achieving secondary and above education, compared to any other group.

The probability of a daughter achieving senior secondary or above education conditional on father's education shows convergence among SC/STs, OBCs, and Muslims. However, there is no convergence of probabilities between HHC and others. This suggests that not only inequality of opportunities based on caste membership (especially between HHC and others)

---

[11]Pr(Daughter=Below Primary/Father=Below primary) is excluded from the graph to preserve space.

[12]Pr(Daughter=Senior Secondary or above/Father=Below primary) is excluded from the graph to preserve space.

exists in India but such inequality has shown little improvement over time. This findings are similar to the findings of Azam and Bhatt (2015) for men, They find no convergence between HHCs and other social groups. Moreover, the probability of attaining senior secondary or above education for sons for all levels of fathers' education in Azam and Bhatt (2015) is higher than for daughters in our study. This obviously suggests pro-son bias in educational investment by parents.

Online Appendix Figure A2 and A3 present corresponding findings for mother-daughter association by caste. The overall findings are similar to the findings reported for father-daughter associations by caste.

5 Conclusion

We examine the intergenerational education transmission between fathers (mothers) and daughters in India for daughters born during 1962-1991. We find that persistence at the bottom of the fathers' or mothers' educational distribution has declined, however, this is compensated by an increase in persistence in other parts of fathers'/mothers' education distribution. Although there has been a significant decline in the probability of a daughter being illiterate/below primary irrespective of parents education, yet, those probabilities are associated with parents' education and caste. Probability of a daughter being illiterate/below primary is lower for more educated fathers (mothers). Similarly, the probability of a daughter being illiterate is lowest for Higher Hindu Castes irrespective of fathers' (mothers') education.

The inequality of opportunities is starker once we consider probability of a daughter attaining senior secondary or above education (top end of the education distribution). Not only the probability of a daughter attaining senior secondary or above education is positively associated with father education levels, the gaps in those probabilities do not show any signs of convergence. For example, the gap in probability of a daughter attaining senior secondary or above education between a daughter of senior secondary or above educated father and a

daughter of below primary educated father is about 0.5 points for the 1962-67 birth cohort, and that increases to about 0.6 points for the 1987-91 birth cohort. Similarly, although the probability of a daughter achieving senior secondary or above education conditional on father's education shows convergence among SC/STs, OBCs, and Muslims, there is no convergence of probabilities between Higher Hindu Castes and the rest. Probability of a daughter attaining senior secondary or above education is higher for Higher Hindu Caste daughters irrespective of parental education.

Our findings are in line with Azam and Bhatt (2015)' findings for father-son educational persistence. Therefore, one may conclude that "Equality of Opportunity" remains an elusive goal for India. More importantly, the gap between the Higher Hindu Castes and the disadvantaged groups such as Other Backward Castes, Scheduled Castes/Tribes remains, and does not show any sign of decline over time.

References

[1] Azam, M. and Bhatt, V. (2015). Like Father, Like Son? Intergenerational Education Mobility in India. Demography, 52(6), 1929-1959.

[2] Basant, R. and Sen, G. (2014). Parental Education as a Criterion for Affirmative Action in Higher Education. World Development, 64, 803-816.

[3] Beteille, A. (2002). Caste, Inequality, and Affirmative Action. Institute of Labor Studies, Geneva.

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

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

[6] Chadwick, L. and Solon, G. (2002). Intergenerational Income Mobility Among Daughters. American Economic Review, 92(1), 335-344.

[7] Checchi, D., Fiorio, C.V., and Leonardi, M. (2008). Intergenerational Persistence in Educational Attainment in Italy. IZA Discussion Paper, 3622.

[8] Checchi, D, Fiorio, C.V, and Leonardi, M. (2013). Intergenerational Persistence in Educational Attainment in Italy. Economics Letters, 118(1), 229-232.

[9] Corak, M. (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. Journal of Economic Perspectives, 27 (3), 79-102.

[10] Daouli, J., Demoussis, M. and Giannakopoulos, N (2010). Mothers, fathers and daughters: Intergenerational transmission of education in Greece. Economics of Education Review, 29 (1), 83-93.

[11] Desai, S and Vanneman, R (2015). India Human Development Survey-II (IHDS-II), 2011-12. ICPSR36151-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2015-07-31. http://doi.org/10.3886/ICPSR36151.v2

[12] DiPrete, T. A. and Grusky, D. B. (1990). Recent Trends in the Process of Stratification. Demography, 27(4), 617-37.

[13] Emran, M. S. and Shilpi, F. (2015). Gender, geography and generations: intergenerational educational mobility in post-reform India. World Development, 72(C), 362-380.

[14] Government of India. (2006). Social, Economic and Education Status of the Muslim community of India. New Delhi.

[15] Government of India. (2013). Key Indicators of Employment and Unemployment in India, 68th Round, National Sample Survey Office.

[16] Hertz, T., Jayasunderay, T., Pirainoz, P., Selcuk, S., Smithyy, N., and Verashchagina, A. (2007). The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. The B.E. Journal of Economic Analysis and Policy (Advances), 7 (2), 1-46.

[17] Holmlund, H., Lindahl, M., and Plug, E. (2011). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. Journal of Economic Literature, 49 (3), 615-51.

[18] Hnatkovskay, V., Lahiri, A., and Paul, S.B. (2013). Breaking the Caste Barrier: Intergenerational Mobility in India. Journal of Human Resources, 48(2), 435-73.

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

[20] Kingdon, G. (2005). Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure. Economic Development and Cultural Change, 53(2), 409-51.

[21] Maitra, P. and Sharma, A. (2009). Parents and Children: Education Across Generations in India. Paper presented at 5th Annual Conference on Economic Growth and Development, Indian Statistical Institute, Delhi.

[22] Muralidharan, K. and Prakash, N. (2013). Cycling to School: Increasing Secondary School Enrollment for Girls in India. Working papers 2013-24, University of Connecticut.

[23] Rama, M., Béteille, T., Li, Y., Mitra, P.K., and Newman, J.L. (2015). Addressing Inequality in South Asia. South Asia Development Matters, Washington, DC: World Bank.

[24] Rawls, J. (1971). A Theory of Justice. Cambridge, MA: Harvard University Press.

[25] Solon, G. (1999). Intergenerational mobility in the labor market. In Ashenfelter, O, and Card D (eds) Handbook of Labor Economics, Vol. 3, Ch. 29.

[26] Stiglitz, J (2012). The Price of Inequality. W. W. Norton and Company.

Figure 1: Probability of daughters' education conditional on fathers' education
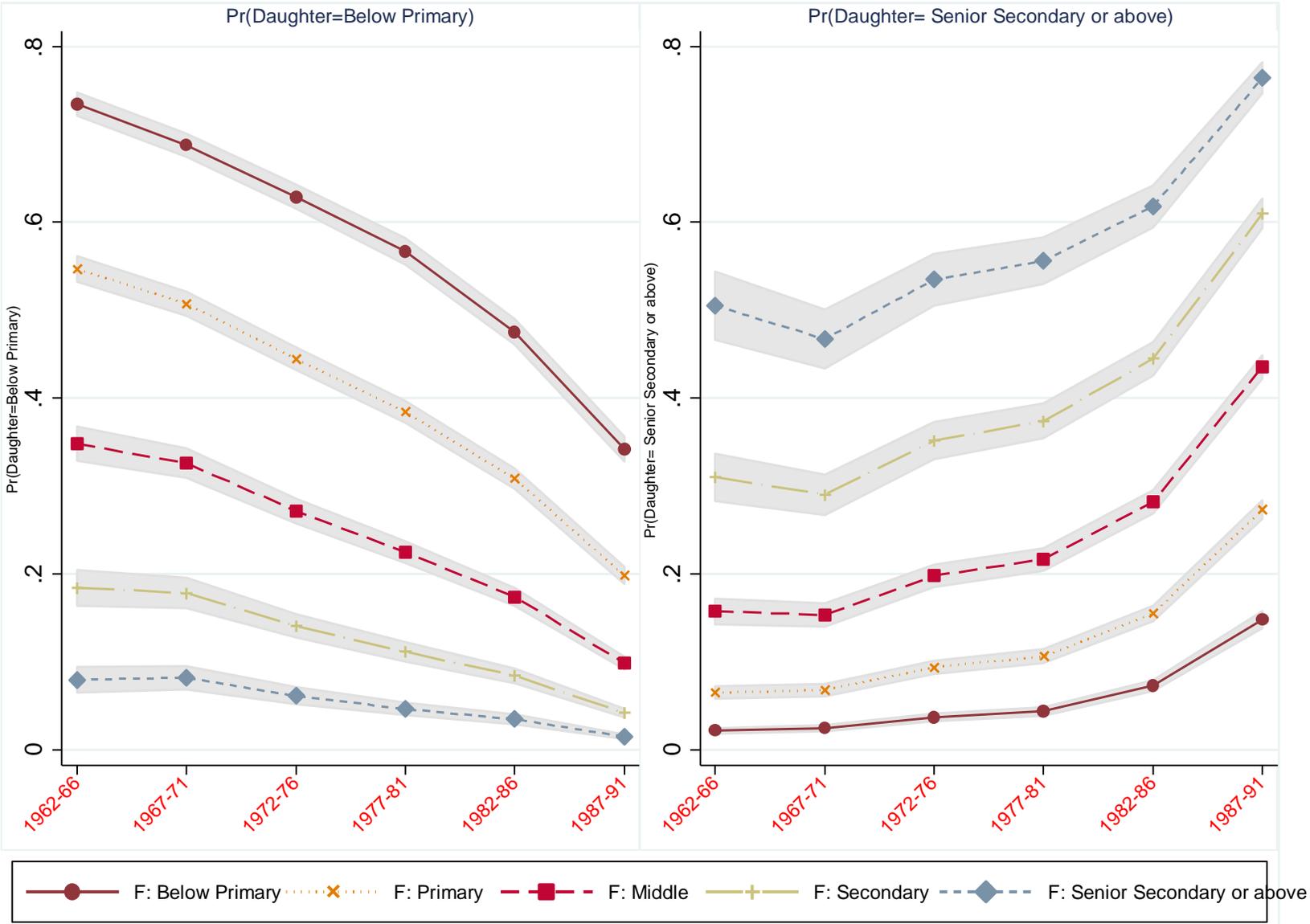
Note: The shaded area represents 95% confidence intervals.

Figure 2: Probability of daughters achieving Below Primary conditional on fathers' education by caste
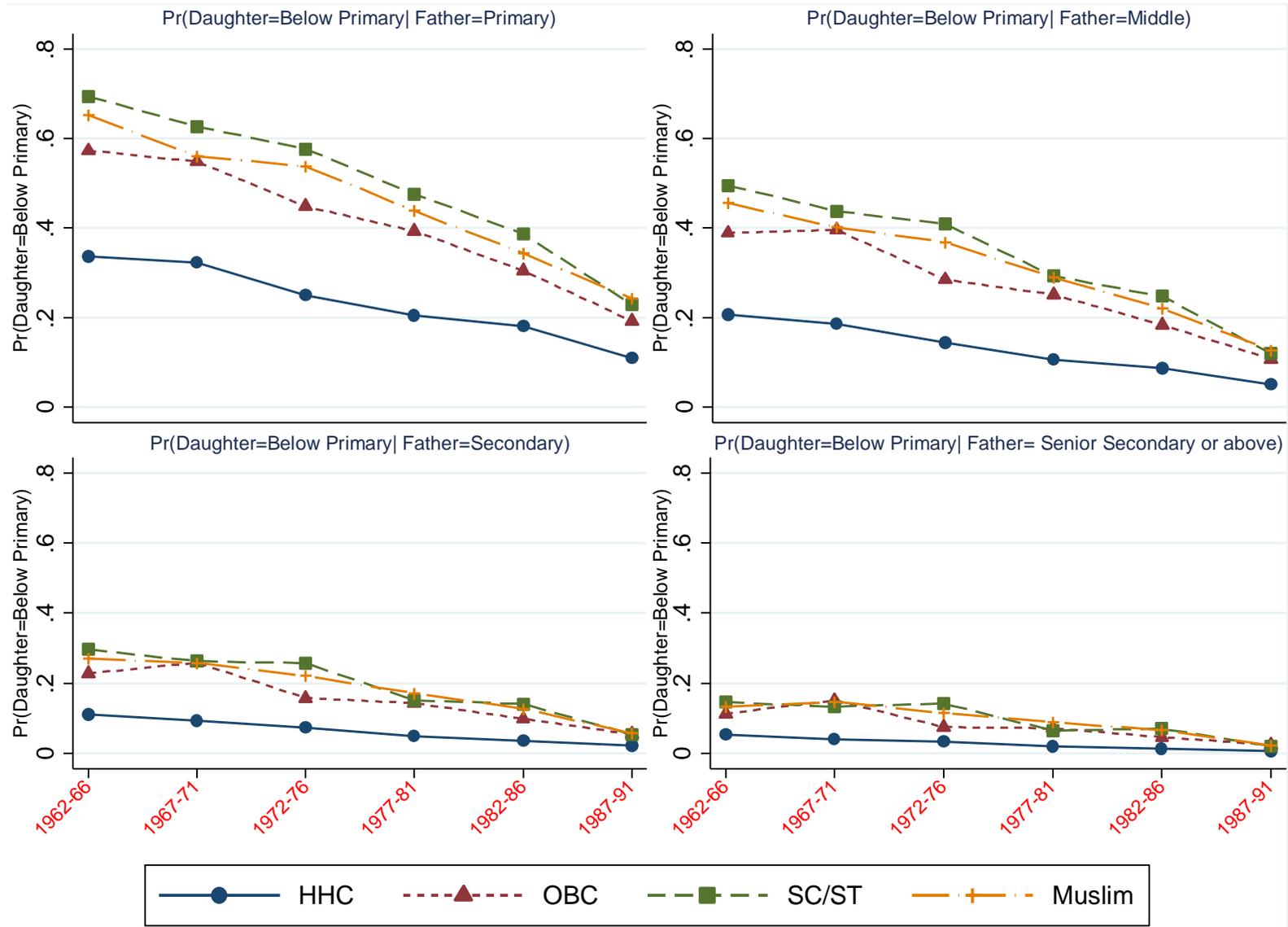
Figure 3: Probability of daughters achieving Senior Secondary or above conditional on fathers' education by caste

Pr(Daughter=Senior Secondary or above| Father=Primary)

Pr(Daughter=Senior Secondary or above| Father=Middle)

Pr(Daughter=Senior Secondary or above| Father=Secondary)

Pr(Daughter=Senior Secondary or above| Father= Senior Secondary or above)

Legend: HHC — OBC — SC/ST — Muslim

X-axis categories: 1962-66, 1967-71, 1972-76, 1977-81, 1982-86, 1987-91

Y-axis: Probability
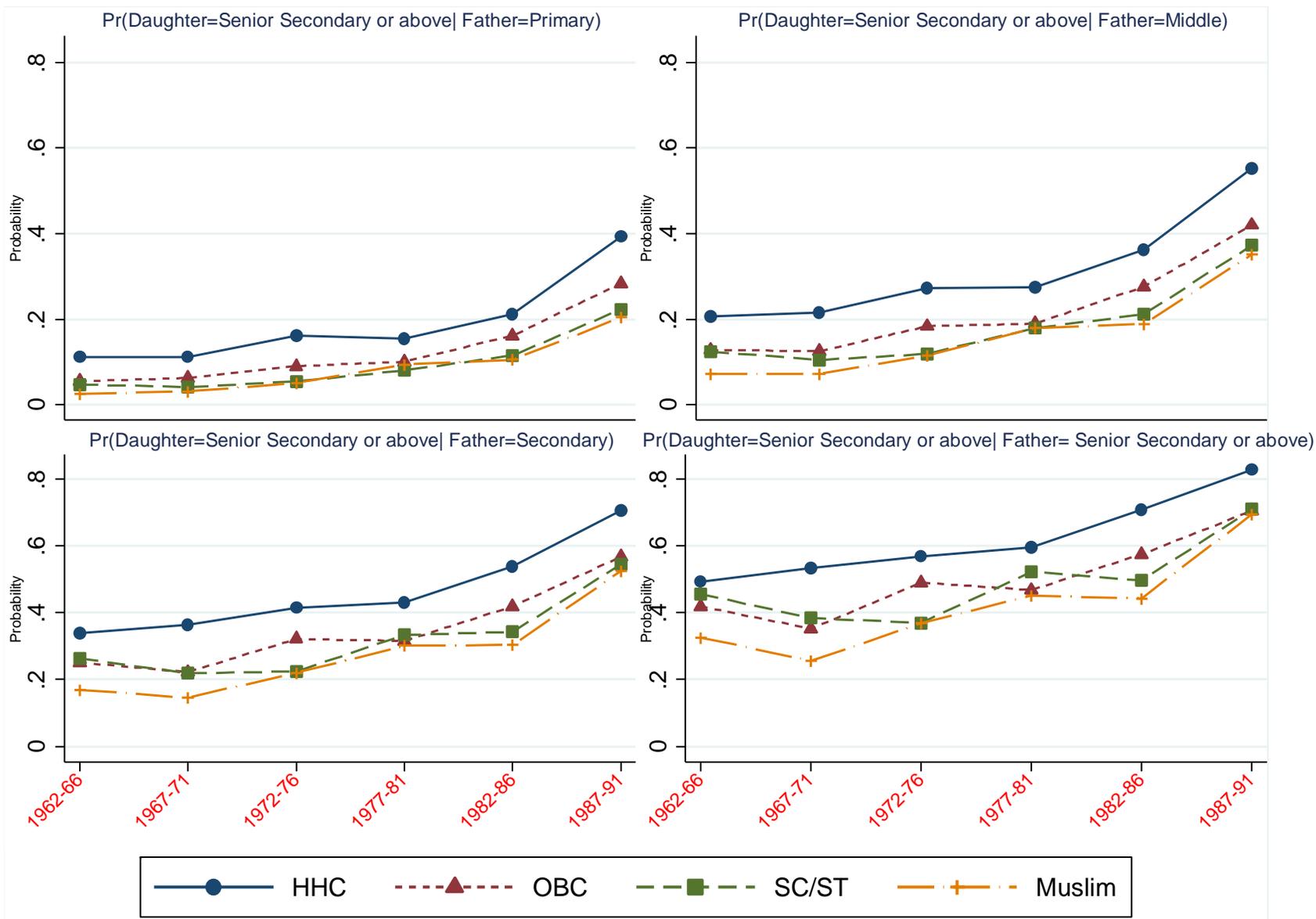
Table 1: Identification of parents' education for adult women in age 20-49

Panel A: Identification of fathers' years of schooling

Year of birth	Total surveyed women with non-missing education information*	Father years of education from women's module	Father years of education from household roster-co-resident father	*Number of women for whom father's years of education is available*	*% of surveyed women for whom father's years of education is available*
1962-66	6,129	5,458	25	5,483	*89.5*
1967-71	6,473	5,915	38	5,953	*92.0*
1972-76	7,157	6,456	97	6,553	*91.6*
1977-81	7,150	6,105	214	6,319	*88.4*
1982-86	8,512	6,147	773	6,920	*81.3*
1987-91	9,855	4,209	3,269	7,478	*75.9*
Total	**45,276**	**34,290**	**4,416**	**38,706**	*85.5*

Panel A: Identification of mothers' years of schooling

Year of birth	Total surveyed women with non-missing education information*	Mother years of education from women's module	Mother years of education from household roster-co-resident father	Number of women for whom mother's years of education available	*% of surveyed women for whom mother's education is available*
1962-66	6,129	5,481	50	5,531	*90.2*
1967-71	6,473	5,928	72	6,000	*92.7*
1972-76	7,157	6,480	157	6,637	*92.7*
1977-81	7,150	6,113	299	6,412	*89.7*
1982-86	8,512	6,162	961	7,123	*83.7*
1987-91	9,855	4,220	3,765	7,985	*81.0*
Total	**45,276**	**34,384**	**5,304**	**39,688**	*87.7*

Note: * IHDS surveyed 45319 women in age group 20-49. 43 observations are dropped because of missing education information.

Table 2: Descriptive statistics

Cohort	Sample size	Years of schooling-daughters		Years of schooling-fathers		Years of schooling-mothers	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
All sample							
1962-66	5483	3.53	4.55	2.51	3.99	0.97	2.41
1967-71	5953	4.01	4.66	2.93	4.27	1.15	2.70
1972-76	6553	4.86	4.90	3.44	4.50	1.45	3.03
1977-81	6319	5.76	5.08	4.11	4.80	1.87	3.44
1982-86	6920	6.66	5.12	4.31	4.84	2.01	3.54
1987-91	7478	8.51	4.97	5.29	4.99	2.84	4.09
All	*38706*	*5.70*	*5.19*	*3.84*	*4.70*	*1.77*	*3.35*
Social Group: Higher Hindu Castes (HHC)							
1962-66	1318	6.45	4.99	4.72	4.89	2.01	3.26
1967-71	1401	6.72	4.92	5.15	4.99	2.26	3.54
1972-76	1478	7.80	4.83	5.90	5.21	2.77	3.92
1977-81	1387	8.62	4.64	6.82	5.14	3.70	4.37
1982-86	1426	9.29	4.77	6.87	5.19	3.92	4.50
1987-91	1520	10.91	4.12	8.20	4.99	5.29	4.74
All HHC	*8530*	*8.35*	*4.95*	*6.32*	*5.20*	*3.36*	*4.25*
Social Group: Other Backward Castes (OBC)							
1962-66	1826	3.19	4.25	2.30	3.72	0.73	2.09
1967-71	1984	3.56	4.40	2.69	4.00	0.97	2.42
1972-76	2289	4.67	4.76	3.24	4.26	1.28	2.79
1977-81	2141	5.57	4.94	3.97	4.55	1.59	3.08
1982-86	2262	6.57	5.07	4.23	4.68	1.73	3.20
1987-91	2304	8.73	4.90	5.56	4.81	2.73	3.93
ALL OBC	*12806*	*5.47*	*5.10*	*3.72*	*4.51*	*1.54*	*3.07*
Social Group: Scheduled castes/Tribes (SC/ST)							
1962-66	1514	1.70	3.33	1.14	2.64	0.32	1.44
1967-71	1738	2.48	3.89	1.59	3.25	0.46	1.78
1972-76	1847	3.10	4.14	1.97	3.49	0.63	1.98
1977-81	1866	4.16	4.71	2.65	4.16	1.02	2.64
1982-86	2137	5.25	4.80	2.96	4.24	1.06	2.61
1987-91	2361	7.16	4.94	3.53	4.41	1.57	3.13
ALL SC/ST	*11463*	*4.19*	*4.78*	*2.40*	*3.90*	*0.89*	*2.43*
Social Group: Muslims							
1962-66	630	2.45	3.67	1.96	3.75	0.77	2.07
1967-71	626	3.19	4.03	2.44	4.03	0.80	2.15
1972-76	769	3.60	4.26	2.74	3.97	1.06	2.47
1977-81	761	4.70	4.77	3.01	4.21	1.18	2.68
1982-86	931	5.71	4.76	3.54	4.42	1.60	3.04
1987-91	1107	7.18	4.85	4.22	4.61	1.99	3.46
Muslims	*4824*	*4.80*	*4.78*	*3.14*	*4.29*	*1.32*	*2.82*

Table 3: Intergenerational persistence in educational attainment among daughters

	(1) 1962-66	(2) 1967-71	(3) 1972-76	(4) 1977-81	(5) 1982-86	(6) 1987-91
Father's years of schooling	0.627***	0.584***	0.589***	0.595***	0.569***	0.535***
($\hat{\beta}$)	(0.019)	(0.017)	(0.015)	(0.014)	(0.014)	(0.013)
Father's years of schooling	0.550***	0.535***	0.542***	0.561***	0.537***	0.537***
($\hat{\rho}$)	(0.017)	(0.015)	(0.013)	(0.014)	(0.013)	(0.013)
SD in daughter's years of (σ_d)	4.548	4.663	4.899	5.085	5.123	4.969
SD in father's years (σ_f)	3.993	4.271	4.505	4.796	4.836	4.995
σ_f/σ_d	0.878	0.916	0.920	0.943	0.944	1.005
Observations	5,483	5,953	6,553	6,319	6,920	7,478
R-squared	0.303	0.286	0.294	0.315	0.288	0.289

Note: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses.

Table 4: Intergenerational persistence in educational attainment among daughters by social groups

	(1) 1962-65	(2) 1966-70	(3) 1971-75	(4) 1976-80	(5) 1981-85	(6) 1986-90
Social Group= Higher Hindu Castes						
Father's years of schooling	0.527***	0.555***	0.476***	0.506***	0.537***	0.416***
($\hat{\beta}$)	(0.027)	(0.025)	(0.028)	(0.027)	(0.037)	(0.026)
Father's years of schooling	0.516***	0.563***	0.514***	0.560***	0.584***	0.504***
($\hat{\rho}$)	(0.027)	(0.026)	(0.030)	(0.030)	(0.041)	(0.031)
SD in daughter's years of (σ_d)	4.993	4.919	4.827	4.640	4.767	4.121
SD deviation in father's years (σ_f)	4.886	4.992	5.207	5.136	5.188	4.993
σ_f/σ_d	0.979	1.015	1.079	1.107	1.088	1.211
Observations	1,318	1,401	1,478	1,387	1,426	1,520
R-squared	0.266	0.318	0.264	0.313	0.342	0.254
Social Group= Other Backward Castes						
Father's years of schooling	0.554***	0.480***	0.561***	0.523***	0.524***	0.494***
($\hat{\beta}$)	(0.041)	(0.034)	(0.026)	(0.027)	(0.025)	(0.026)
Father's years of schooling	0.486***	0.437***	0.503***	0.481***	0.483***	0.484***
($\hat{\rho}$)	(0.036)	(0.031)	(0.024)	(0.025)	(0.023)	(0.025)
SD in daughter's years of (σ_d)	4.247	4.397	4.758	4.943	5.070	4.902
SD deviation in father's years (σ_f)	3.724	4.004	4.262	4.554	4.675	4.809
σ_f/σ_d	0.877	0.911	0.896	0.921	0.922	0.981
Observations	1,826	1,984	2,289	2,141	2,262	2,304
R-squared	0.236	0.191	0.253	0.232	0.233	0.234
Social Group= Scheduled Castes/Tribes						
Father's years of schooling	0.518***	0.520***	0.511***	0.599***	0.505***	0.540***
($\hat{\beta}$)	(0.051)	(0.040)	(0.035)	(0.035)	(0.033)	(0.026)
Father's years of schooling	0.410***	0.435***	0.431***	0.529***	0.446***	0.482***
($\hat{\rho}$)	(0.041)	(0.034)	(0.029)	(0.031)	(0.029)	(0.023)
SD in daughter's years of (σ_d)	3.331	3.893	4.136	4.711	4.797	4.935
SD deviation in father's years (σ_f)	2.640	3.251	3.490	4.163	4.237	4.408
σ_f/σ_d	0.793	0.835	0.844	0.884	0.883	0.893
Observations	1,514	1,738	1,847	1,866	2,137	2,361
R-squared	0.168	0.189	0.186	0.280	0.199	0.232
Social Group= Muslims						
Father's years of schooling	0.504***	0.451***	0.454***	0.515***	0.498***	0.523***
($\hat{\beta}$)	(0.053)	(0.047)	(0.047)	(0.041)	(0.037)	(0.031)
Father's years of schooling	0.516***	0.452***	0.423***	0.454***	0.463***	0.497***
($\hat{\rho}$)	(0.054)	(0.047)	(0.043)	(0.037)	(0.034)	(0.029)
SD in daughter's years of (σ_d)	3.667	4.028	4.256	4.774	4.755	4.853
SD deviation in father's years (σ_f)	3.750	4.030	3.971	4.212	4.416	4.613
σ_f/σ_d	1.023	1.001	0.933	0.882	0.929	0.951
Observations	630	626	769	761	931	1,107
R-squared	0.266	0.204	0.179	0.206	0.214	0.247

Note: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses.

Table 5: Decomposition of persistence measured by correlation (ρ)

	Daughter-stage attended	Father-stage attended	1962-66	1967-71	1972-76	1977-81	1982-86	1987-91
1	D:No education	F:No education	0.249	0.210	0.173	0.149	0.113	0.062
2	D:Primary	F:No education	0.057	0.055	0.060	0.049	0.047	0.037
3	D:Middle	F:No education	0.031	0.035	0.035	0.043	0.047	0.041
4	D:Secondary	F:No education	0.023	0.026	0.033	0.038	0.046	0.053
5	D:College	F:No education	0.002	0.002	0.003	0.003	0.006	0.010
6	*Total contribution to the correlation coefficient of the group of sons with not educated father*		*0.362*	*0.329*	*0.304*	*0.283*	*0.259*	*0.203*
7	D:No education	F:Primary	*0.033*	*0.031*	*0.028*	*0.025*	*0.019*	*0.014*
8	D:Primary	F:Primary	0.015	0.015	0.018	0.014	0.013	0.012
9	D:Middle	F:Primary	0.012	0.013	0.014	0.017	0.017	0.018
10	D:Secondary	F:Primary	0.013	0.015	0.020	0.023	0.025	0.035
11	D:College	F:Primary	0.002	0.002	0.003	0.004	0.006	0.013
12	*Total contribution to the correlation coefficient of the group of sons with Primary attended father*		*0.075*	*0.077*	*0.084*	*0.082*	*0.080*	*0.092*
13	D:No education	F:Middle	0.011	0.012	0.011	0.011	0.009	0.006
14	D:Primary	F:Middle	*0.008*	*0.009*	*0.010*	*0.008*	*0.008*	*0.007*
15	D:Middle	F:Middle	0.008	0.009	0.010	0.013	0.013	0.012
16	D:Secondary	F:Middle	0.013	0.015	0.020	0.025	0.028	0.033
17	D:College	F:Middle	0.004	0.004	0.005	0.007	0.010	0.019
18	*Total contribution to the correlation coefficient of the group of sons with Middle attended father*		*0.045*	*0.049*	*0.057*	*0.063*	*0.068*	*0.077*
19	D:No education	F:Secondary	0.007	0.008	0.008	0.009	0.006	0.004
20	D:Primary	F:Secondary	0.007	0.008	0.009	0.009	0.008	0.006
21	D:Middle	F:Secondary	*0.010*	*0.012*	*0.012*	*0.017*	*0.016*	*0.014*
22	D:Secondary	F:Secondary	0.020	0.025	0.032	0.047	0.045	0.053
23	D:College	F:Secondary	0.011	0.011	0.015	0.023	0.028	0.050
24	*Total contribution to the correlation coefficient of the group of sons with Secondary attended father*		*0.056*	*0.064*	*0.075*	*0.105*	*0.103*	*0.127*
25	D:No education	F:College	0.000	0.000	0.000	0.000	0.000	0.000
26	D:Primary	F:College	0.000	0.001	0.001	0.001	0.000	0.001
27	D:Middle	F:College	0.001	0.001	0.001	0.002	0.002	0.001
28	D:Secondary	F:College	*0.004*	*0.006*	*0.008*	*0.010*	*0.009*	*0.011*
29	D:College	F:College	0.004	0.005	0.008	0.011	0.013	0.024
30	*Total contribution to the correlation coefficient of the group of sons with College attended father*		*0.009*	*0.013*	*0.018*	*0.023*	*0.025*	*0.037*
31	**Correlation Coefficient**		**0.547**	**0.531**	**0.538**	**0.557**	**0.535**	**0.535**

Note: The continuous years of schooling is grouped to refer attended stages of schooling. No education: 0 years; Primary: 1-5 years; Middle: 6-8 years; Secondary: 9-12 years; and College: 13 -16 years.

Intergenerational Mobility among Daughters: Evidence from India

Mehtabul Azam

Oklahoma State University & IZA

Online Appendix

(Mother-daughter persistence)

Table A1: Intergenerational persistence in educational attainment between mothers and daughters

	(1) 1962-66	(2) 1967-71	(3) 1972-76	(4) 1977-81	(5) 1982-86	(6) 1987-91
Mother's years of schooling ($\hat{\beta}$)	1.030*** (0.030)	0.936*** (0.025)	0.865*** (0.020)	0.814*** (0.017)	0.772*** (0.014)	0.640*** (0.013)
Mother's years of schooling ($\hat{\rho}$)	0.549*** (0.016)	0.538*** (0.014)	0.532*** (0.012)	0.548*** (0.011)	0.544*** (0.010)	0.528*** (0.010)
SD in daughter's years of (σ_d)	4.537	4.689	4.931	5.111	5.136	4.975
SD deviation in mother's years (σ_m)	2.417	2.695	3.035	3.440	3.618	4.101
σ_m/σ_d	0.533	0.575	0.615	0.673	0.704	0.824
Observations	5,531	6,000	6,637	6,412	7,123	7,985
R-squared	0.299	0.292	0.286	0.303	0.295	0.282

Table A2: Decomposition of persistence measured by correlation (ρ)

	Daughter- stage attended	Mother-stage attended	1962-66	1967-71	1972-76	1977-81	1982-86	1987-91
1	D:No education	M:No education	0.281	0.248	0.200	0.173	0.134	0.072
2	D:Primary	M:No education	0.077	0.080	0.085	0.070	0.067	0.051
3	D:Middle	M:No education	0.049	0.058	0.057	0.070	0.077	0.066
4	D:Secondary	M:No education	0.042	0.049	0.064	0.076	0.089	0.107
5	D:College	M:No education	0.004	0.004	0.006	0.008	0.013	0.027
6	*Total contribution to the correlation coefficient of the group of sons with not educated father*		*0.453*	*0.438*	*0.412*	*0.397*	*0.381*	*0.324*
7	D:No education	M:Primary	0.015	0.013	0.013	0.013	0.009	0.006
8	D:Primary	M:Primary	0.010	0.009	0.011	0.009	0.008	0.007
9	D:Middle	M:Primary	0.010	0.011	0.011	0.014	0.014	0.012
10	D:Secondary	M:Primary	0.017	0.017	0.022	0.028	0.029	0.034
11	D:College	M:Primary	0.005	0.004	0.006	0.008	0.010	0.019
12	*Total contribution to the correlation coefficient of the group of sons with Primary attended father*		*0.058*	*0.055*	*0.062*	*0.071*	*0.070*	*0.079*
13	D:No education	M:Middle	0.002	0.002	0.002	0.003	0.002	0.001
14	D:Primary	M:Middle	0.002	0.002	0.003	0.003	0.002	0.002
15	D:Middle	M:Middle	0.003	0.004	0.004	0.006	0.006	0.005
16	D:Secondary	M:Middle	0.009	0.012	0.014	0.019	0.021	0.023
17	D:College	M:Middle	0.006	0.007	0.008	0.011	0.015	0.025
18	*Total contribution to the correlation coefficient of the group of sons with Middle attended father*		*0.020*	*0.027*	*0.031*	*0.041*	*0.046*	*0.056*
19	D:No education	M:Secondary	0.000	0.000	0.000	0.001	0.000	0.000
20	D:Primary	M:Secondary	0.000	0.000	0.001	0.001	0.000	0.001
21	D:Middle	M:Secondary	0.001	0.001	0.002	0.003	0.002	0.003
22	D:Secondary	M:Secondary	0.004	0.006	0.010	0.014	0.013	0.019
23	D:College	M:Secondary	0.006	0.008	0.012	0.016	0.020	0.036
24	*Total contribution to the correlation coefficient of the group of sons with Secondary attended father*		*0.011*	*0.016*	*0.025*	*0.034*	*0.036*	*0.059*
25	D:No education	M:College	0.000	0.000	0.000	0.000	0.000	0.000
26	D:Primary	M:College	0.000	0.000	0.000	0.000	0.000	0.000
27	D:Middle	M:College	0.000	0.000	0.000	0.000	0.000	0.000
28	D:Secondary	M:College	0.000	0.000	0.000	0.001	0.001	0.001
29	D:College	M:College	0.001	0.002	0.002	0.005	0.006	0.010
30	*Total contribution to the correlation coefficient of the group of sons with College attended father*		*0.001*	*0.002*	*0.003*	*0.005*	*0.007*	*0.011*
31	**Correlation Coefficient**		**0.544**	**0.538**	**0.532**	**0.548**	**0.540**	**0.529**

Note: The continuous years of schooling is grouped to refer attended stages of schooling. No education: 0 years; Primary: 1-5 years; Middle: 6-8 years; Secondary: 9-12 years; and College: 13 -16 years.

**Figure A1: Probability of daughters education conditional on mothers' education**
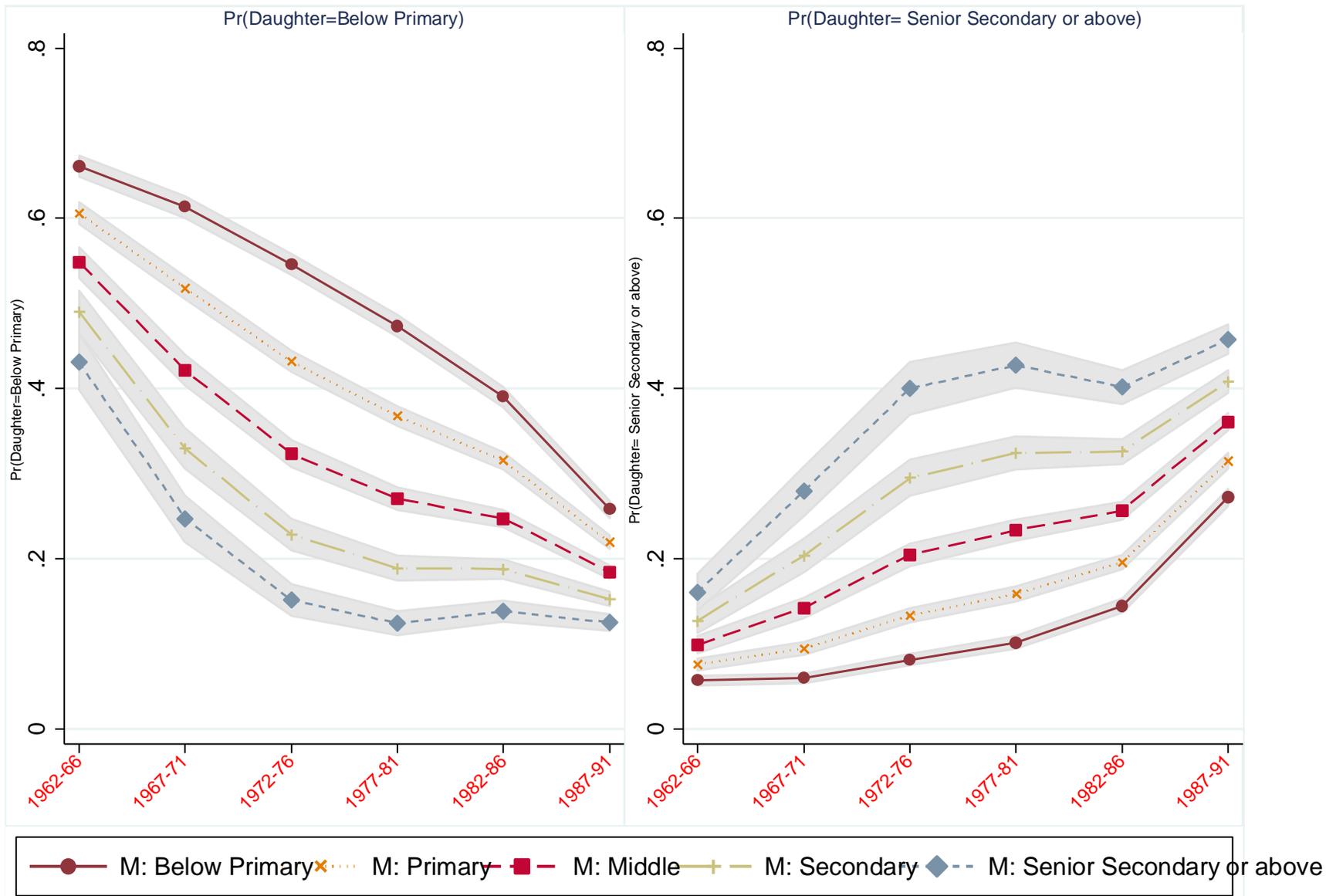
Figure A2: Probability of daughters achieving Below Primary conditional on mothers' education by caste
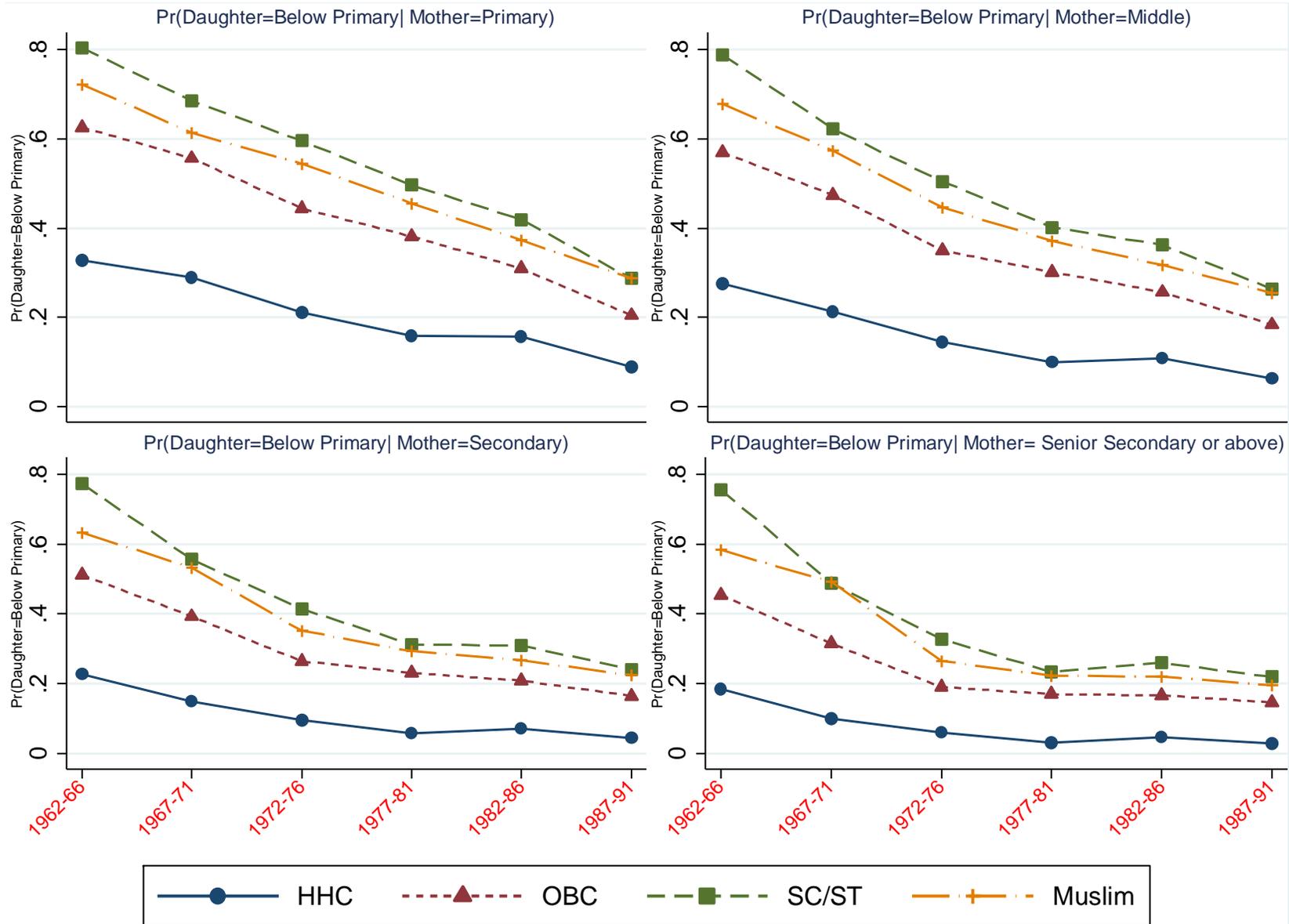
Figure A3: Probability of daughters achieving Post-Secondary conditional on mothers' education by caste

Pr(Daughter=Senior Secondary or above| Mother=Primary)

Pr(Daughter=Senior Secondary or above| Mother=Middle)

Pr(Daughter=Senior Secondary or above| Father=Secondary)

Pr(Daughter=Senior Secondary or above| Father= Senior Secondary or above)

HHC   OBC   SC/ST   Muslim