Parents and Children: Education Across Generations in India*

Pushkar Maitra† and Anurag Sharma‡

November 2009

Abstract

It has been argued that one of the reasons for the uneven distributional effects of the high rates of economic growth in India has been the lack of mobility of the Indian population. In this paper we use a nationally representative data set from India to examine one aspect of mobility: that of educational attainment across generations. Specifically, we examine role of parental education on two aspects of child's educational attainment i) years of schooling attained and ii) progression across different schooling levels. We find that there has been a significant increase in educational attainment of individuals over the last 70 years, with women gaining the most in terms of increases in educational attainment. Restricting the analysis to adults (those more than 20 years old at the time of the survey), we find that when we account for the potential endogeneity of parental educational attainment, the *total* effect of parental education on years of schooling of their children is not statistically significant indicating an increase in intergenerational educational mobility. The analysis of school progression, conducted using a sample of $15 - 24$ year olds at the time of the survey, reveals that father's educational attainment has a positive and statistically significant effect on the probability of *continuing* to post-secondary school/college. Private investment (driven primarily by father's education and hence income) continues to be crucial for young adults to benefit from the opportunities offered by the Indian growth process.

JEL Classification: O12, I21, C31.

Keywords: Intergenerational Transmission of Human Capital, School Progression, Sequential Probit Model, India.

*Pushkar Maitra would like to acknowledge funding provided by an Australian Research Council Discovery Grant. We have benefitted from conversations with Hal Hill, Stephen Howes, Robert Willis and comments by seminar participants at the Australian National University, Canberra and participants at the Indian Economy and Business Update 2009. The usual caveat applies.

†Pushkar Maitra, Department of Economics, Monash University, Clayton Campus, VIC 3800, Australia. Email: Pushkar.Maitra@Buseco.monash.edu.au

‡Anurag Sharma, Centre for Health Economics, Monash University, Clayton Campus, VIC 3800, Australia. Email: Anurag.Sharma@Buseco.monash.edu.au

1 Introduction

Mobility is one of the hallmarks of the process of development. In traditional economies individuals are typically closely attached to the structure they are born into − be it the occupational structure of their parents or be it the land of their forefathers. In a modern economy however individuals are able to seek out the best for themselves in terms of their occupation, location of residence or anything else for that matter.

This is a particularly important issue in the context of India. Among developing countries India stands out in terms of the remarkably low levels of mobility (see for example Gupta, 2004; Munshi and Rosenzweig, 2009). This lack of mobility means that many sections of the society are unable to reap the benefits of the phenomenal levels of economic growth that the country has experienced over the last two decades. Indeed by a number of different measures, inequality in outcomes has actually increased over the relevant period. Part of this could be due to the fact that in a society characterized by lack of mobility, the gains from growth accrue disproportionately across the population and in particular some sections of the population are unable to take advantage of the opportunities that the growth process in the country has provided. For the benefits of the growth process to be distributed in a much more egalitarian manner, the population needs to be mobile, both vertically (in terms of increasing their levels of educational attainment across generations) and spatially (in terms of physically moving to the location that provides the best opportunities).

In this paper we re-examine the issue of differences in human capital accumulation over generations, i.e., examine the issue of vertical (or inter-generational) mobility in educational attainment. In particular we focus on the issue of the correlation between education levels of parents and children, which reflects the degree of equality of opportunity in a society (Becker and Tomes, 1986). There are several mechanisms through which parental education can affect human capital outcomes of their children. For example, maternal education can improve efficiency of human capital production leading to increasing returns, across generations, in parental human capital (Becker et al., 1990). Alternatively, one could consider the education level of mothers to be a function of their endowed human capital, which is positively correlated with that of their children (see for example Rosenzweig and Wolpin, 1994; Behrman and

Rosenzweig, 2002).

While the issue of intergenerational mobility in educational attainment has received some attention in other countries (see Behrman et al. (2000); Bourguignon et al. (2003) for a discussion of this issue in the context of Brazil and Thomas (1996) for a discussion on South Africa), the issue has received surprisingly little attention in the context of India. To the best of our knowledge the only paper that explicitly addresses this issue in the context of India is Jalan and Murgai (2008). They use data from the 1992-93 and 1998-99 National Family Health Surveys to examine inequalities in educational outcomes across groups of individuals and the perpetuation of these inequalities across generations. Our paper builds on the work by Jalan and Murgai (2008). Using a nationally representative data set for India (the Indian Human Development Survey) collected in 2005 we examine the following questions. *First*, how has the profile of educational attainment of adults (defined as at least 20 year old at the time of the survey) changed over generations? *Second*, what is the effect of parental education on the educational attainment of individuals? In this context we explicitly account for the potential endogeneity of parental educational attainment in any regression that examines the effect of parental educational attainment on the schooling of the next generation. This potential endogeneity arises from the fact that parental educational attainment might be correlated with some of the unobserved determinants of the child's schooling. Alternatively the unobserved components of the child's educational attainment might be correlated with the unobserved components of that of the parents. Both of this could result in biased estimates. *Finally* restricting the sample to young adults (aged $15 - 24$ at the time of the survey) we examine effect of parental educational attainment on the propensity to continue in school for these individuals. These individuals form an important demographic because they are the ones who need to be able to take advantage of the opportunities that are likely to come their way as a result of the growth process.

In addition to the issue of mobility (and its implications for the process of development) the issue of children's human capital accumulation is particularly crucial for India where proportion of school going children (aged between $0 - 14$ years) is projected to be 23% of total population by 2025 (compared to 18% in China and the US and 13% in Europe UN, 1999). Overall however, as Kingdon (2007) argues, the story of India's educational achievements is one of mixed success. On the one hand while India has 22% per cent of the worlds population, it has 46% of the worlds illiterates, and is home to a high proportion of the worlds out-of-school children and youth. On the other hand, it has made significant progress in raising schooling participation. Additionally it has emerged as an important player in the worldwide information technology revolution on the back of large numbers of well-educated computer-science and other graduates. However as Kingdon (2007) also argues, the base of the Indian education pyramid is weak and this has serious implications for the extent of overall human capital accumulation in the country.

Given the linkage between human capital accumulation and economic growth (Barro, 2001), an analysis of inter-generational mobility in educational attainment and also of the determinants of school progression could help predict the major trends in future economic growth in India. The results from such an analysis can help in the design of an effective policy to augment the human capital of children. There is broad consensus in the literature on the positive effects of parental education on school enrolment for both developed and developing countries. The empirical analysis in this paper goes a step further and examines the effect of parental education on the "progression" of children across different enrolment levels.

Why is school progression important? In recent years the issue of school dropouts (the inverse of school progression) has been of increasing concern to policy makers both in developed and developing countries. While there might be "valid" economic reasons for dropping out of school early, the consequences of such action can be quite severe. In developing countries children typically drop out of school because the current income requirements of the household exceed the expected returns from continuing to remain in school. Obviously this has significant long-term impacts − low educational attainment and consequently low levels of human capital accumulation, which in turn imply that future income earning opportunities are limited and lifetime incomes are low. Additionally there is an inter-generational effect: children born to parents with low levels of education are themselves more likely to end up with low levels of educational attainment.[1]

[1] However this is not only a developing country problem. Even in developed countries early school dropouts and non-completion of schooling is fast becoming a serious problem. For example in Australia (an OECD country), studies have shown that almost one third of students drop out from high school each year and most of them never gain a year 12 or equivalent qualification. High school dropouts are much more likely to be unemployed, or to have given up looking for work, or to have low lifetime incomes. They are more likely to have poor numeracy and literacy skills, which affect their productivity, social participation and decision-making. While half of the male 15 to 19-year-olds who leave school early end up with full-time work, as do 65% of 20 to 24 year olds who dropped out of school, the story is much more depressing for women. Less than 35% of

The empirical analysis of school progression uses the correlated sequential probit estimation technique (see Lillard and Willis, 1994; Pal, 2004) which has several advantages: First, it allows us to identify the children who have progressed much less than others and also to locate at what level of schooling this has happened; second, it controls for the self selection of students into the next higher level of schooling; and third, it allows us to (potentially) use different control factors for each transition. The results from this analysis therefore enable us to decompose the effect of parental education at different transitions. For example this technique enables us to identify the scenario where education of both the parents is significant for school enrolment but only father's education matters for progression to higher schooling level.

2 How has Education Increased over Generations?

This paper uses data from the Indian Human Development Survey 2005 (IHDS). This survey is a nationally representative, multi-topic survey of 41,554 households in 1,503 villages and 971 urban neighborhoods across India. Two one-hour interviews in each household covered topics concerning health, education, employment, economic status, marriage, fertility, gender relations, and social capital. Survey was conducted between November 2004 and October 2005 with a response rate of 92%. For the first part of the paper (sections $2-3$) we restrict ourselves to individuals born in 1985 or before (who are at least 20 years old at the time of the survey). The relevant sample means are presented in columns (1) and (2) in Table 1 (corresponding to Sample 1) for the rural and urban sample respectively. We want to highlight the difference in the average years of schooling by residence: the average years of schooling is 4.34 years for rural residents and 7.69 years for urban residents (the difference is statistically significant at the 1% level).

We start by presenting selected descriptive statistics on the how education has changed over generations. Figure 1 presents the average years of schooling attained by birth cohort. There is no doubt that Indians have become more educated. For urban males, the average years of schooling has increased from close to 6 years for individuals born in 1930 or before (Cohort

women in both age groups who left school early have a full-time job and more than a third of the 20 to 24 year olds are unemployed (see Spierings, 1999).

1) to 10 years for individuals born between 1981 and 1985 (Cohort 12). The corresponding increases are from 2 to 9 years for urban females, 2 to 7.5 years for rural males and 1 to 5.5 years for rural females. The decline in the average years of schooling for individuals born after 1985 is essentially due to the fact that a number of individuals belonging to these cohorts are at school at the time of the survey and have not attained the maximum level of education. To be more specific, while 3.59% of individuals born in 1985 or before are enrolled in school at the time of the survey; this percentage rises to 56.73% for individuals born over the period $1986 - 1990$ and to 92.13% for individuals born over the period $1991 - 1995$.

Mean growth rates in educational attainment are presented in the first row of Panel A in Table 2. These are the coefficient estimates from a standard OLS regression of the years of schooling on birth year.[2] Separate regressions are conducted for males and females and for those residing in rural and urban areas (i.e. we present four sets of results). The average years of education has increased by 1.2 years each decade for urban females and around 1 year each decade for rural males and females. The average years of schooling has increased by about 0.5 years per decade for urban males, though it is important to note that the average years of schooling for urban males is quite high to begin with. Figure 2 presents the lowess plots of the years of schooling on the year of birth. These lowess plots essentially tell the same story; additionally they highlight the non-linearity in the relationship between the year of birth and the years of schooling attained, something not explicitly addressed in the OLS regressions presented in Table 2, Panel A.

Panels B, C, D and E in Table 2 present the marginal effects from probit regressions for the highest level of education attained. The dependent variables in the regressions are: Enrolled in school (which takes the value of 1 if the child has any schooling and 0 otherwise), Enrolled in middle school (= 1 if the years of schooling attained by the individual more than 5 and 0 otherwise), Enrolled in post-secondary school (= 1 if the years of schooling attained by the individual is more than 10 and 0 otherwise), Enrolled in college (= 1 if the individual has ever attended college and 0 otherwise). The biggest gainers are the females with rural females gaining more than urban females. The marginal effects presented in Panel B show that the probability of enrolling in school increases by nearly 16 percentage points every decade for rural females; followed by 8.4 percentage points every decade for urban females, 7.6 percentage

[2] The sample is restricted to individuals born in 1985 or earlier.

points every decade for rural males and finally 2.6 percentage points every decade for urban males. Rural females have gained in every stage: the probability of enrolling in middle school has increased by 10.1 percentage points every decade and the probability of enrolling in post-secondary school has increased by 7.5 percentage points every decade.

Next we examine the extent of unevenness in the growth of educational attainment across generations. Table 3 essentially repeats the same analysis as in Panel A of Table 2, but this time we stratify the sample on the basis of religion (columns $2-5$) and caste (columns $6-10$). As a point of comparison we present the results for the full sample in column 1: these are the coefficient estimates presented in Table 2, Panel A.[3] There is, not surprisingly, a great deal of variation across the different religions and castes and unfortunately (particularly in the case of females) there is very little evidence of catching up by those with the lowest educational attainment. For example when we look at the sample of rural females, not only is the mean years of schooling the lowest among the Muslims, so is the growth rate. While the years of schooling for those belonging to other religions has increased on an average by 1.2 years every decade, that for the Muslims has increased by 0.8 years every decade. Given the initial difference in average years of schooling (0.27 years for Muslims compared to 0.38 years for those belonging to other religions for those born before 1930), the difference in growth rates implies that the average years of schooling attained by rural females belonging to other religions is more than double that of Muslims in 1980. Looking at the results for the different castes, it is clear that individuals belonging to scheduled tribes (ST) have fared the worst in terms of educational attainment.

3 Effect of Parental Education

Educational attainment is typically influenced by both public and private investments in education. While state policy typically drives the former, parental education is a crucial part of the latter; indeed parental education is one of the most important determinants of a child's education. In a rigid society with no mobility, parental education completely determines the educational attainment of the child. Put in another way, after controlling for other socio-

[3]Theoretically the concept of caste does not and should not exist for non-Hindus; however partly because of history and partly because of affirmative action policies aimed at certain castes, many non-Hindus appear to hold on to their caste.

economic characteristics that potentially affect educational attainment of an individual, the greater the influence of paternal and maternal education, the lower is the extent of intergenerational mobility.

The information on parental educational attainment is available only for a subset of the full sample. The descriptive statistics are presented in columns (3) and (4) in Table 1 (corresponding to Sample 2) for the rural and urban samples respectively. Most of the averages are similar to the full sample. The new variables here are the years of schooling attained by the father and the mother. Fathers are more educated compared to mothers in both rural and urban households (in both cases the average difference in the years of schooling attained is around 3 years) and also both fathers and mothers in urban households are more educated compared to their rural counter parts (again the difference is around 3 years at the mean).

To examine the issue of the association between parental education attainment and child educational attainment we start by presenting in Figure 3 the lowess plots from a non-parametric regression of the years of schooling attained by an individual on the years of schooling attained by the father and mother. These plots essentially capture the relationship between the educational attainment of the father or mother and that of the child using a non-parametric regression of the form:

$$\text{Years of schooling of the child} = f(\text{Years of schooling of the father or mother}) + \epsilon$$

The results are interesting. First, there is a positive correlation between parental education and children's education. Second, on an average, the educational attainment of children is greater than the educational attainment of the fathers, for fathers with less education; i.e., if the father has attained x years of schooling, the child has on an average attained more than x years of schooling $-$ the lowess plots always lie above the 45° line for fathers with low education. The slope is however generally less than 1, indicating that an additional year of schooling attained by the father is not associated with an additional year of schooling attained by the child. For more educated fathers, children have on an average less years of schooling compared to fathers. It is however worth noting that the lowess plots intersect the 45° line at a fairly high level of education attained by the father $-$ to be precise beyond 10 years of schooling. Given that the average years of schooling of rural fathers is 4.1 years and that of urban fathers is 7.1 years, this implies that for the sample under consideration children are generally more educated than the fathers. Third, while broadly the effects are similar when we look at the association between the educational attainment of the mother and that of the child, there is however a very noticeable gender effect in this case. For mothers with low educational attainment (defined as mothers having less than 3 years in the urban sample and less than 8 years in the rural sample), on an average, if the mother has attained x years of schooling, both sons and daughters have attained more than x years of schooling and the educational attainment of sons exceeds that of daughters. However for more educated mothers, on an average the educational attainment of daughters exceeds that of the sons.

This result has significant policy implications. Policy makers and economists have argued of the need to increase educational attainment of mothers as the most important way of reducing the educational gender gap. While this is true, it is also true that a threshold level of education (that is not constant by gender and or region of residence) must be attained before this policy can work.

However in trying to analyze the effect of parental educational attainment on that of the child, we face a potential endogeneity problem: the years of schooling of the mother and father are potentially endogenous arising from the fact that parental educational attainment could be correlated with some of the unobserved determinants of child's schooling[4] and failure to correct for this endogeneity would result in biased estimates. The problem is to find proper instruments that are correlated with parental educational attainment and not with the educational attainment of the child. This is a difficult problem in a non-longitudinal data set.

There are growing number of studies focussing on determinants of child schooling in developing countries. Pal (2004) analyzes child schooling data for Peruvian households and reports that parental education positively affects child schooling at primary and secondary levels, but not at post-secondary levels. Singh (1992) examines major economic aspects of demand of schooling of farm operators in Brazilian rural households and finds that parental education positively affects household demand for children's education with mothers education having larger effect than that of the father. A similar result is reported by Maitra (2003) for demand

[4] For example it might result in inter-generational transmission of values that result in the next generation attaining more schooling. Alternatively the unobserved components of child's schooling decision might be correlated with parents unobserved characteristics (one such example is genetic characteristics).

for schooling in Bangladesh. Dreze and Kingdon (2001) use data on 1143 households for rural north India to analyze the impact of school quality on school participation. They find that probability of participation increases with parental education, though mother's education does not have significant effect on male school participation. Evidence from Pakistan suggests that parental education significantly increases the education of their sons (Holmes, 2003). Unfortunately few studies have explicitly accounted for potential endogeneity of parental educational attainment. One notable exception is Lillard and Willis (1994) who explicitly account for this endogeneity using data from Malaysia.

Returning to our choice of instruments: one possibility would have been to use the grandparent's educational attainment as the relevant instrument. Lillard and Willis (1994) argue that this particular variable is likely to be strongly correlated with the educational attainment of the parents but not directly correlated with that of the grand child, who is the index individual. However there are several problems with this variable. First, this data is available for a very small subset of the population; second even when it is available, it is available only for grand parents' on the father's side. Most importantly however it is difficult to argue that this variable is uncorrelated with the educational attainment of the index child. This can happen in two ways. First, more educated grandparents are likely to be associated with richer households, which could ease considerably financial constraints within the households. To the extent financial constraints were adversely affecting child schooling, an easing of financial constraints within the household would increase educational attainment of the index child. Second, more educated grandparents would typically inculcate a culture of education within the household, thereby resulting in more education in successive generations.

We then focus on public investments in education. Ideally we would have liked to have information on school facilities at the location where the parent grew up. This is likely to have a significant effect on the educational attainment of the father and mother but would not be correlated with the child's educational attainment. Unfortunately this information is also not available. We therefore use as instruments the year of birth of the father and mother and the interaction of the year of birth with original location of the father and mother i.e., whether the father/mother originally lived in a rural or urban area (year of birth of the parent $\times$ original location of the parent).[5] We essentially assume that public investments in

[5]Similar location and time based instruments have been used by Schultz (2002) and more recently by Gau

schooling vary over time and over location and the year of birth and the original location of the parent captures that effect. The only way these two variables could really have an effect on the child's educational attainment would be through the effect on parental educational attainment. The choice of instruments is validated using the standard Sargan test.

In Table 4, we present the results from OLS and instrumental variable regressions of years of schooling of the child on year of birth, educational attainment of the father and mother and a set of other socio-economic characteristics (religion, whether the individual belongs to a poor household and also a set of state dummies).[6] Again separate regressions are computed for males and females and for rural and urban residence. Of course we have a significantly smaller sample now, since the data on parental educational attainment is available for a subset of the sample of individuals born in 1985 or before. The instrumental variable regression results show that parental education has almost no effect on children's education: in only 1 out of the 8 possible cases is the years of schooling attained by the father or mother have a statistically significant effect on the child's education − mother's education has a statistically significant effect on the educational attainment of rural males, though even here the effect is quite weak, significant only at the 10% level.

Our results corroborate those obtained by Jalan and Murgai (2008). They find that inter-generational mobility in education in India has improved significantly and consistently across generations and that mobility has improved across all major social groups and land classes. While educational gaps continue to persist across social groups and classes, the gap between social groups is actually quite small. The entire education gap is driven by difference between the rich and the poor. The implications of the results presented in Table 4 are quite interesting. One way of interpreting these results is that it is public investments in education that matters and not private investments. Educational attainment of the next generation is *not* constrained by the endowments, preferences and opportunities provided by the previous

and Smyth (2009) to estimate effect of height on wages. They use the average number of health institutions in the province in which the respondent was born at the time of birth and at age 12. They also use an index of the location (village, small town, city) in which the respondent lived when growing up. Our choice of instruments is in the spirit of that used by Schultz (2002) and Gau and Smyth (2009) but we are restricted by the lack of available data: we do not have information on the actual province in which the parents were born − all we know is whether they were born in a rural or an urban region.

[6]In an alternative specification we compute the same regressions, but instead of including dummies for religion we include dummies for caste. The results do not change qualitatively. These results are available on request.

generation and that state policy has been successful in severing this link.

There are no surprises in the coefficient estimates of the additional control variables included in Table 4. Educational attainment of Muslims is significantly lower compared to that of Hindus (the Muslim dummy is always negative and statistically significant); the Christian dummy is never statistically significant and in the rural areas individuals belonging to other religions (including Sikhs, Jains and Buddhists) have significantly lower educational attainment compared to Hindus. The Poor household dummy is always negative and statistically significant, implying that individuals in poor households (at the time of the survey) have fewer years of schooling. Interpretation of this variable could however be problematic because this variable tells us whether the household is poor or not at the time of the survey and not (necessarily) at the time when schooling decisions were made.[7]

The rest of this paper further examines this link between parental education and child education, but with one difference. We examine the effect of parental education on school progression rather than school completion. Such an analysis enables to identify factors which affect school progression and thus drop out rates. We use a restricted sample of individuals who are $15-24$ year old at the time of the survey.[8] Given the cross-sectional nature of the data means though that the information on household and socio-economic characteristics is available only at the time of the survey. It is difficult to interpret the results if we consider individuals who are far removed from the time when actual education decisions were made. By focussing on this sub-sample, we are able to also examine the role of other socio-economic characteristics (including sibling characteristics) on educational attainment.

4 A Sequential Probit Model for Educational Attainment

4.1 Methodology

We examine school progression as a relevant indicator of child schooling in India within a dynamic sequential framework. In this set up, school progression is conditional on attainment

[7]On the other hand this was not a period of rapid economic growth in India. This also means that mobility out of poverty might also not have been particularly high.

[8]Traditionally the school going age group is $6-18$. However in many developing countries children delay their initial enrolment and also continue to remain in school beyond the age of 18.

at the previous level and also self-selection into the next higher level of schooling. This is based on a correlated sequential probit model (see Lillard and Willis, 1994; Pal, 2004), which allows us to identify the children who have progressed much less than others and also to locate at what level of schooling this has happened. This is important for any assessment of policies geared to boost child schooling because it is based on a full understanding of the nature of the selection process across different levels of schooling. We will, in particular, focus on the following three levels of transition: (a) considering all sample children, whether a child gets enrolled in a primary school ($s = 1$); (b) among those enrolled in primary schools, whether a child moves on to the middle/secondary level ($s = 2$); and (c) among those enrolled in secondary schools, whether a child moves on to the post-secondary level ($s = 3$). Figure 4 presents a schematic representation of this sequential framework. While decision (a) relates to school enrolment decisions, (b) and (c) relate to school attainment, i.e., school progression from primary to middle/secondary level and that from the middle/secondary to post-secondary levels respectively. Our analysis of child school progression therefore combines both indicators of enrolment and attainment in a sequential framework. Movement from primary to middle/secondary level is conditional on the successful completion of the final year of the primary school; similarly moving from middle/secondary to post-secondary level requires one to "pass" the final year at the secondary level. We take account of the process of self-selection at each higher level as only a fraction of children successfully completing primary (or secondary) schools will move on to the secondary (or post-secondary) schools. For example, decision (b) selects those who successfully complete primary schooling and move on to middle/secondary school and (c) selects those who successfully complete middle/secondary school and move on to the post-secondary level. The framework that we use allows the determinants of selection to vary across the different transitions. In addition to the child's ability, we control for sibling composition, household resource constraint, parental preferences and some community characteristics and obtain selectivity corrected correlated sequential probit estimates of school progression.

The use of the sequential probit model also allows us to address the important issues of probability spikes and censoring. Surveys typically measure schooling attained by the years of education attained (or the highest grade completed). This leads to several problems. First, even though desired schooling might be a continuous variable, the researcher only observes only discrete years of schooling. Second, data on education attainment from developing countries is often characterized by a large mass point at zero years of education and similar probability spikes at primary and secondary school completion levels, where progress to the next level is often impeded by school fees and entrance requirements. OLS estimation is therefore inappropriate under this set up. An alternative would be to use an ordered probit/logit model of attainment of specific levels of schooling. While this approach takes into account the discreteness of the data and the probability spikes, it fails to account for the censoring in the data arising from the fact that some children are enrolled in school at the time of the survey. One can argue that the desired level of schooling equals the completed years of education for the children that are not currently enrolled in school. However for children that are currently enrolled in school, the desired years of schooling exceeds the years of completed schooling. These observations are therefore right censored. A sequential model, since it is concerned only with the probability of continuing on to a certain level (conditional on successful completion of the previous level), is able to avoid the issue of censoring. By definition it addresses the issue of probability spikes. In addition it accounts for the selection into stages, which ordered probit/logit models typically do not.

Standard modeling techniques used in most existing studies on child schooling (including those specific to India) fail to capture the specific characteristics of child schooling in India, where primary enrolment rates are high along with high failures and drop out rates. Thus school progression is a better indicator of child schooling than simply school enrolment/attainment measured by the years of schooling. We determine child school progression in dynamic framework, using sequential probit model, which is argued to be better than the corresponding ordered probit estimates. For someone at the secondary level, for example, the sequential probit model takes account of the fact that the person has completed the primary level to reach the secondary level while the ordered probit model neither takes account of the achievement at the previous level nor does it correct for any self-selection into the next higher level of schooling.

We consider four mutually exclusive levels of schooling: none, primary, middle/secondary and post-secondary/college. Define

$$\omega = \begin{cases} 0 & \text{if years of schooling is 0} \\ 1 & \text{if years of schooling is } 1-5 \text{ years} \\ 2 & \text{if years of schooling is } 6-10 \text{ years} \\ 3 & \text{if years of schooling is 10+ years} \end{cases} \tag{1}$$

The three levels of transition that we have talked about above follow directly from equation 1. The first decision is whether to attend school at all ($s = 1$); for those who attend school, the second decision is whether to continue on to middle/secondary school or stop in primary school ($s = 2$); for those who attend middle/secondary school is whether to stop there or continue to post-secondary school or college ($s = 3$).

Our primary interest is in modeling the transition from one stage to another. For each individual i we can write a probit index function at each decision node s as follows:

$$I_{si} = \beta_s X_{si} + \delta_i + u_{si}; s = 1, 2, 3 \tag{2}$$

Equation 2 is the propensity to continue in school (move from one level to the next higher level) and includes covariates (X_{si}), which vary by individual and (potentially) by decision level. For example, some variables like parental education, religion are constant across decisions, but there might be other variables (like number of siblings, state of residence) that vary across decisions. Since the data set that we use is cross-sectional, we do not include in X_{si} any variable that varies over decisions. Individual variables affect the propensity to attend school differently depending on the transition stage s, and therefore β_s is decision specific. For example it is feasible that educational attainment of the mother has a positive and statistically significant effect on the propensity to enrol in school, but not on the conditional probability of continuing on to middle/secondary school and on to post-secondary school.

We also account for heterogeneity in the propensity to continue in school, assumed to capture all correlation across schooling decisions. This is represented by the residual term δ_i, which is constant across the different schooling decisions. The remaining residual terms at each decision point (u_{si}) are assumed to be independent of δ_i and of each other. Both δ_i and u_{si} are assumed to be normally distributed as follows: $u_{si} \sim N(0, 1)$ and $\delta_i \sim N(0, \sigma_\delta^2)$.

Individual i will move from level ω to level $\omega + 1$ if $I_{si} > 0$ and drop out otherwise.[9] So we

[9] A move from level ω to level $\omega + 1$ is denoted by transition level s.

can write

$$P(\omega) = \begin{cases} P[I_{1i} \leq 0] & \text{if} \quad \omega = 0 \\ P[I_{1i} > 0, I_{2i} \leq 0] & \text{if} \quad \omega = 1 \\ P[I_{1i} > 0, I_{2i} > 0, I_{3i} \leq 0] & \text{if} \quad \omega = 2 \\ P[I_{1i} > 0, I_{2i} > 0, I_{3i} > 0] & \text{if} \quad \omega = 3 \end{cases}$$

In other words, the decision to move to the next level is correlated with the previous decision and subsequent decisions are subject to selectivity with respect to the earlier decisions. Correlation across stages means that a joint marginal likelihood function is maximized for estimating the model.

It is worth noting that some individuals are enrolled in school at the time of the survey at level ω. These individuals are censored in that the highest level of schooling for these individuals is not yet known. The probability of censored schooling levels is given by the probability of the grade enrolment or higher. Also enrolment at the highest level is equivalent to attaining this level.

4.2 Specification, Data and Descriptive Statistics

The focus of our paper is to examine the intergenerational transmission of education. This is captured by looking at the effect of years of schooling attained by the father and mother on the educational attainment of the child (school progression to be more specific). We correct for the potential endogeneity of parental educational attainment in the child schooling regressions using the methodology developed by Rivers and Vuong (1988). The procedure can be described as follows. First years of schooling attained by the father and mother are regressed on the full set of explanatory variables including the instruments. This is the first stage regression. The error terms from the first stage regression (Residual Father's schooling and Residual Mother's schooling) are included as additional regressors in the second stage estimation (the sequential probit model). Joint significance of the two residuals ($LR = -2(L_R - L_U) \sim \chi^2(6)$, where L_R and L_U are the restricted and unrestricted log likelihood values) implies that the null hypothesis of exogeneity of parental educational attainment on the child's educational attainment is rejected. The estimation results show that the null hypothesis of exogeneity of parental educational attainment is strongly rejected. As before we use as instruments the year of birth of the father and mother and year of birth of the father and mother $\times$ original location of the father and mother.

The other explanatory variables that we use in the regressions include individual and household level characteristics. Individual (child level) characteristics include the age of the child and the number and composition of siblings. We compute and present separate results for males and females and also by rural or urban residence. We include both the age of the child and the square of the age of the child to account for any non-linearity in the age effect. These two age terms also allow us to account for any birth cohort effects. To examine whether there is a quantity-quality trade-off (Becker and Lewis, 1973) in educational attainment we include the number of co-resident siblings for each child. However since the age and the gender of the siblings could be important, we stratify the number of siblings by age and gender: the number of brothers and sisters in the age group $0-5, 6-10, 11-14$ and $15-24$. It has been argued that sibling composition may play an important role in a child's school participation, particularly if the child comes from a poor resource constrained household. This classification of siblings therefore takes account of whether there is competition for the limited household resources in schooling decisions. We will return to this issue below. We include the number of working age males and females (aged $25-60$) in the household; dummies for religion (Muslim, Christian and Other religion; the reference category being Hindu), a dummy to indicate if the household is below the poverty line and finally a set of province dummies to account for any unobserved public investment (including state policy) that might have an effect on child educational attainment. The estimation results show that the null hypothesis of exogeneity of parental educational attainment is strongly rejected.

The descriptive statistics are presented in Columns (5) and (6) in Table 1 (corresponding to Sample 3) for the rural and urban households respectively. Despite the fact that a large fraction of the sample (37% in the rural areas and 51% in the urban areas) are in school at the time of the survey and have essentially not completed their educational attainment, the average years of schooling are much higher compared to sample of adults: 7.57 years and 9.51 years in the rural and urban households respectively, compared to 4.34 years and 7.69 years for the sample of adults. Interestingly 63% of the rural sample of $15-24$ year olds are male. We cannot be sure if this is simply a sampling bias or the reflection of a much deeper problem.

Figure 5 gives us an indication of the extent of the problem of school progression in India. Enrolment rates among the $15-24$ year olds is quite high, ranging from 94% for urban males and females to 84% for rural females. However beyond primary school the progression rates drop considerably. Of every 100 rural females aged $15 - 24$, 84 have enrolled in primary school, 70 have continued on to middle/secondary school and finally only 19 have continued on to post-secondary school/college. The corresponding figures are 95, 85 and 36 for urban males, 94, 88 and 42 for urban females and finally 89, 75 and 20 for rural males.

4.3 Results

In Table 5 we present a subset of the results. These are maximum likelihood sequential probit estimates of (a) whether to attend school (years > 0); (b) whether to attend middle/secondary schools or stop after the primary level (years > 5) and (c) whether to move on to post-secondary levels or stop after the secondary level (years > 10). As before we stratify the sample on the basis of gender and rural and urban residence - i.e., we conduct the estimation on 4 separate sub-samples: urban male, urban female, rural male and rural female.

For each sub-sample, three sets of results are presented. Specifications 1 and 2 present the results from the restricted model where the child's education is assumed to depend only on parental education. The difference between specifications 1 and 2 is that in specification 1, parental educational attainment is assumed to be exogenous in the child schooling regressions. In specification 2 we account for the potential endogeneity of parental educational attainment. The null hypothesis of exogeneity of parental educational attainment is strongly rejected for all sub-samples, indicating that the specification where parental education is endogenous is the preferred specification.

However, one drawback of restrictive model specification is that the results might be misleading if parental education is simply a proxy for correlated economic and demographic factors that influence schooling decisions. For example, parental education might increase family income and ability to finance child's education which in turn can affect child's schooling decision. Failure to account for these other individual, household and state level characteristics could result in significant omitted variable bias (different from the potential endogeneity bias). We therefore estimate a full specification model which in addition to parental education includes a number of individual and household characteristics (as discussed above, see section 4.2). The corresponding results give us specification 3. Parental educational attainment is assumed to be endogenous in this specification.[10] Also we report only the coefficient estimates of father's and mother's education in Table 5. The full set of results are presented in the Appendix (Table A-1). Two other points are worth noting. First the additional controls are jointly statistically significant ($\chi^2(111)$, see Table 5); and second the extent of omitted variable bias is considerable (compare specifications 2 and 3 in Table 5). We come back to this issue below.

Heterogeneity in the propensity to continue in school (captured by σ_δ^2), which is assumed to capture all correlation across schooling decisions, is always statistically significant. There is a common individual level unobserved factor that affects schooling decision at every level and ignoring this common individual level unobserved heterogeneity would result in biased estimates. The extent of this bias becomes clear once we compare the effects of parental educational attainment on school progression in the three specifications.

Parental Education Effect

The issue of endogeneity of parental educational attainment turns out to be fairly crucial and it is important to note that the effects of parental educational attainment on the school progression propensities are quite different in the three specifications. In the restricted models (specifications 1 and 2), both the years of schooling attained by the father and the years of schooling attained by the mother always have a positive and statistically significant effect on the continuation probabilities. However the effects vary across the different transitions once we account for the potential endogeneity of parental education. Comparing the results from specifications 1 and 2 we find that failing to account for endogeneity of parental educational attainment results in an over estimation of the effect of the father's educational attainment on school enrolment, but an under estimation of the probability of continuing (or propensity to continue) on to the middle/secondary school and post-secondary school/college. The effect of the mother's educational attainment follows the exact opposite pattern: failing to account for endogeneity of mother's educational attainment under estimates the effect of mother's educational attainment on the probability of enrolment, but over estimates the effect on the probability of continuation to middle/secondary school level and to post-secondary

[10]The corresponding results for exogenous specification are presented in the Appendix (Table A-1).

school/college level.

The results from specification 3 are quite different. First, father's education does not have a statistically significant effect on the propensity to enrol in school; mother's educational attainment is positive and statistically significant. This result is true irrespective of the gender of the child and irrespective of the sector of residence (rural or urban). However when we compare the results from specifications 2 and 3 we find evidence of significant omitted variable bias: over estimation of the effect of the father's educational attainment and an under estimation of the effect of the mother's educational attainment.

Moving on to transition 2 (continuing to middle/secondary school, conditional on completing primary school) we find that again with the exception of urban males, the father's educational attainment does not have a statistically significant effect on the conditional probability of attending middle/secondary school. Mother's educational attainment on the other hand has a positive and statistically significant effect on the conditional probability of attending middle/secondary school, though not for urban males. Again a comparison of the estimates in the three specifications shows that with the exception of the sub sample of urban males, ignoring the control variables results in significant over estimation of the effect of the father's educational attainment on the propensity that the child attends middle/secondary effect while it results in a significant under estimation of the effect of mother's educational attainment on the corresponding probability. For the sub-sample of urban males however the effects are the opposite: ignoring control variables results in an under estimation of the effect of the father's education but an over estimation of the effect of the mother's education.

Finally when we move to transition 3 (continuing on to post-secondary education), we find that mother's educational attainment ceases to have a statistically significant effect on the probability of continuing on to post-secondary levels; on the other hand father's education is now positive and statistically significant.

It is important to note that the relative importance of the father's and mother's educational attainment changes as children move from primary to middle/secondary to post-secondary levels. We can calculate the change in probability of schooling due to an additional year of parental education. The changes in probability (computed using the estimated coefficients in Specification 3) are reported in Figure 6. An additional year of schooling attained by the mother increases the probability of school enrolment by 2.2% for urban females and 4.3% for rural females. The corresponding figures for urban and rural males are 3.2% and 4.2% respectively. Similarly the likelihood of enrolling in middle school increases by 6.7% for urban females and 8.3% for rural females with an additional year of their mothers schooling.[11] These results imply that mother's education is more important for rural females relative to urban females. Fathers education has almost no effect on school enrolment and continuation on to middle/secondary school except in the case of urban males. However, an additional year of father's education increases the probability of progression to post-secondary school/college with the effect being higher in urban areas. We come back to a possible explanation and some implications of this result later.

These results are important from a policy point of view: clearly mother's education is important at the initial levels (decision whether to attend school in the first place or the decision on whether to attend middle/secondary school, conditional on successful completion of primary school (the latter effect though breaks down for the sample of urban males). At either of these two transitions, father's educational attainment has almost no role to play. However father's educational attainment becomes crucial in the decision to continue on to post-secondary levels. This variation in effects is not evident if we do not account for endogeneity and other household socio-economic characteristics. So if the focus of policy makers is to increase school enrolment in the first place (as is the policy focus in large sections of the country), they need to target the mothers. On the other hand if the focus is to ensure that children continue on to post-secondary levels (may be to obtain technical training), policy makers need to target the fathers and to the extent that the father's educational attainment is a proxy for his income, private investment continues to be important.[12]

Other Results

The full set of results are presented in Table A-1 in the Appendix. While not the primary focus of the paper we discuss some of the more interesting results. Irrespective of the spec-

[11] These numbers have been computed using the methodology developed by Petersen (1985).

[12] This effect becomes even clearer when we compare the estimates of the years of schooling of the father and the mother in the exogenous and endogenous specifications, presented in Table A-1 in the Appendix.

ification, the sample and the level, the Poor dummy is always negative and statistically significant. This implies that children belonging to poor households (those below the poverty line) are significantly less likely to enrol in school, significantly less likely to continue on to middle/secondary school and significantly less likely to continue beyond secondary school (to post-secondary school/college). There is therefore evidence of significant resource constraints within the household, which is reflected in sub-optimal investment in schooling. Schooling has costs in India. Even apparently free government schooling has substantial costs, such as expenditure on books, stationery, travel, and school uniforms. The PROBE report (PROBE, 1999) found that, in rural north India, parents spend about Rs. 318 per year per child who attends government (i.e., tuition-free) school, so that an agricultural laborer in Bihar with three such children would have to work for about 40 days of the year just to send the children to primary school (see Kingdon, 2005).

Resource constraints within the household manifest themselves in other ways as well. The most important of which is the sibling competition or sibling rivalry effect.[13] The idea is that given resource constraints within the household, siblings compete among themselves over resources, both parental time and money resources (see for example Mulder, 1998; Greenhalgh, 1985) and this has implications for human capital accumulation within the household. Sibling gender composition has an important influence on intrahousehold resource allocation of schooling and health resources, particularly if the child comes from a poor, resource-constrained household. For example, if there is a gender bias operating at the household level, then for a given family size, it must be the case that a male child, growing up in a household with only brothers, may have fewer resources than if he were to grow up with sisters only. This is likely to be true for females as well. This would imply that the educational attainment of children depends not only on their own gender but also on whether their siblings are female or male. Hence, siblings become rivals in a competition for greater access to household resources. We find that irrespective of the gender of the child and the region of residence, an increase in the number of brothers or sisters aged $0-10$ always reduces the probability of enrolling in school, continuing on to middle/secondary school and continuing beyond secondary school for the $15-24$ year olds. The effects are in most cases statistically significant. These results are consistent with the argument that elder siblings in resource

[13] We ignore the issue of endogeneity of the number of siblings here.

constrained households have to leave school in order to take care of their siblings. What is interesting is that the results are quite similar across gender.

The effects are however quite varied if there is an increase in the number of siblings aged $11-24$. Having an additional brother in this age group reduces the probability of enrolling in school, continuing on to middle/secondary school and continuing beyond secondary school. This is true for both boys and girls. However having an additional sister in this age group actually increases the corresponding probabilities at every level. It is worth noting however that the effects of having an additional sister aged $11-24$ is much weaker for continuation to middle/secondary school and continuation beyond secondary school. The results then imply that having a sister around the same age is actually beneficial for both boys and girls. In this case appears to be no sibling competition effect $-$ rather we seem to find evidence of what could be termed as the sibling synergy effect. The results are consistent with the hypothesis that older children may subsidize the education of younger siblings by contributing to family time or budget. At transition 2 (continuation to middle/secondary school), only older sisters in the age group of $15-24$ years significantly increase the probability that the younger brothers and sisters continue on to middle/secondary school. The effect of older sisters on school enrolment is independent of the gender composition of the siblings. However, at transition 3 (continuation to post-secondary school/college) the effect of older sisters (in the age group of $15-24$) remains significant only for siblings of the same gender. This leads to two interesting observations. First, school enrolment is affected by both categories of older sisters: the $11-15$ year olds who are more likely to contribute to family time and the $15-24$ year olds who are more likely to contribute to family budget, whereas progression across schooling levels is significantly affected by only one category of older sisters; ones who are more likely to contribute financially to family budget. Second, females are more likely to continue beyond middle/secondary school if they have an elder sister in the age group of $15-24$: again a reflection of the sibling synergy effect. There is therefore a positive spill over effect from having children attend school, due to economies of scale in child costs and/or from the development of a culture of schooling within the household influenced by social forces. The spill over essentially results in a form of positive externality in children's educational attainment, which the parents will try to internalize through their schooling decisions for subsequent (other) children. Thus rather than the standard sibling competition effect we find a sibling synergy effect, making it optimal for parents to educate a greater number of children.[14]

The effect of an additional working age male is never statistically significant. However, an additional working age female in the household has a significant and positive effect on school enrolment and progression of all children in the household. This is not surprising and is partly a reflection of the fact that in developing countries one of the most important reasons as to why children drop out of school is that they need to help with household chores. Having an additional working age woman in the household clearly makes things easier. Our results show significant heterogeneity in schooling among children from different religions. Muslim, Christian and other religion children have reduced likelihood of enrolment in school, relative to Hindu students. Even conditional on enrolment children from Muslim households are less likely to progress beyond primary school. Conditional on enrolment however the results are more positive for Christian children and children belonging to other religions.

5 Concluding Remarks

Investment in human capital through education is universally recognized as an essential component of economic development. While education endows individuals with the means to enhance their skills, knowledge, health and productivity, it also enhances the economy's abil-

[14]While the issue of sibling competition effect is fairly well established result in the context of developed countries (see for example Behrman et al. (1989) and Conley and Glauber (2006) for the US and Goux and Maurin (2004) for France), evidence using data from developing countries is quite mixed. Evidence from Thailand (Knodel et al., 1990) and Brazil (Psacharopoulos and Arriagada, 1989) suggest that there is a negative relationship between the number of siblings and educational attainment. In the case of Vietnam (Anh et al., 1998) the relationship is negative for families with six or more children and the effects are quite small once other families are controlled for. In the context of Indonesia, Malarani (2004) finds that the relationship between family size and schooling was positive or neutral for the older cohorts, but for the younger cohorts there is a negative relationship. Qian (2009), using data from China, provides causal evidence of benefit of an additional sibling to first born children. She finds that for one-child families, an additional child significantly increases the school enrolment of the first born. Using data from Taiwan, Parish and Willis (1993) and Greenhalgh (1985) find that where children with relatively more older sisters have higher schooling rates. Shavit and Pierce (1991), using data from Israel find that for the richer Jews, family size has a negative relationship with educational attainment of children, while there is a positive relationship between family size and educational attainment of children in the poorer Muslim households. In African countries, however, there is no negative effect and in fact the opposite is true: educational attainment has a positive relationship with the number of siblings (see Gomes (1984) for evidence from Kenya; Chernichovsky (1985) for evidence from Botswana; and Cornwell et al. (2007) for evidence from South Africa). The explanation for this positive relationship typically involves households in Africa (and indeed the poor Muslim communities in Israel) drawing on a large kinship network beyond the immediate family, which reduces the costs (financial, emotional and time) associated with additional children.

ity to develop and adopt new technology. Not surprisingly therefore, increasing education levels is an important policy concern in most countries.

The primary purpose of this paper is to look at the association between parental education and children's education. Educational attainment of any individual depends on both private and public investments. Parental educational attainment is the prime component of the former. In the extreme case where the society is completely rigid, children's educational attainment is determined completely by parental educational attainment. All that matters is private investment. But that also means that the distributional impacts of growth could be negative, since a large section of the population will not be able to accumulate the human capital necessary to reap the benefits of the process of economic growth.

There are three main results of this paper. First, we find that in India there has been a significant increase in educational attainment of individuals over the last 70 years or so with women gaining the most in terms of increases in educational attainment. Second, restricting the sample to adults (those more than 20 years old at the time of the survey), we find that when we account for the potential endogeneity of parental educational attainment, it is public investments in education that matters and not private investments. Educational attainment of the next generation is *not* constrained by the endowments, preferences and opportunities provided by the previous generation and that state policy has been successful in severing this link. Ignoring the potential endogeneity of parental educational attainment significantly over estimates the effect of parental educational attainment on the educational attainment of adults. Finally, the sequential probit analysis of school progression (for the $15-24$ year olds at the time of the survey) shows that the effect of parental education on school progression varies over the different stages. Mother's education is important at the initial transitions − decision whether to attend school in the first place or the decision on whether to attend middle/secondary school, conditional on successful completion of primary school (the latter effect though surprisingly breaks down for the sample of urban males). At either of these two transitions, father's educational attainment has almost no role to play. However father's educational attainment becomes crucial in the decision to continue on to post-secondary levels. This variation in effects is not evident if we do not account for endogeneity of parental education. So if the focus of policy makers is to increase school enrolment in the first place (as is the policy focus in large sections of the country), they need to target the mothers. On the other hand if the focus is to ensure that children continue on to post-secondary levels (may be to obtain technical training), policy makers need to target the fathers. These results are a distinct improvement over the existing univariate probit or ordered probit estimates of the highest educational attainment as they allow us to account for selection at the different levels and also shows (i) how different characteristics affect school progression differently at the different stages and (ii) how the same set of factors might affect different levels of schooling in a different manner.

Controlling for endogeneity weakens the direct effect of fathers education on school enrolment and on continuation beyond primary school schooling and mothers education on continuation beyond middle/secondary school. However conditional on the child continuing beyond primary school, father's educational attainment has a significant effect on progression to post-secondary school/college. This result is consistent with the existing hypothesis in the literature that the direct effect of fathers education results from its impact on the household income and resources (market factors). Thus the decision to send a child to post-secondary school/college is primarily determined by household income. In contrast the direct effect of mothers education is associated with non-market factors such as awareness about education and quality of time she spends with children. Our result of positive and significant effect of mothers education at initial stages of schooling is consistent with the above hypothesis.

Over the long run (and after accounting for the potential endogeneity of parental educational attainment) we find that there has been a significant increase in intergenerational educational mobility in India. These results also imply that it is public investments in education and not private investments that matters. Educational attainment of the next generation is *not* constrained by the endowments, preferences and opportunities provided by the previous generation and that state policy has been successful in severing this link. This is a positive result and is not particularly surprising. Over the last 70 years the rate of growth of educational infrastructure has far exceeded the rate of growth of parental educational attainment. It is therefore logical that the effect is dominated by the supply side variable: a clear case of *if you build it they will come*. The main result from school progression analysis is that father's educational attainment has a positive and statistically significant effect on the probability of continuing to post-secondary school/college. To the extent that father's educational attainment is a proxy for his income, these results seem to suggest that the probability of continuing to post-secondary schooling is positively affected by the father's income (which one would think is the primary determinant of private investment in schooling).[15] Given that a large part of the Indian growth process over the last 2 decades has been driven by the highly skilled Information Technology sector, continuation to post-secondary school/college (in order to accumulate the necessary technical skills) is crucial to benefit from the opportunities offered by the Indian growth process. In this respect private investment (driven primarily by father's income) continues to be crucial.

[15] It is worth noting that the "poor" dummy is always negative and statistically significant: it is the father's income that matters and not overall household income. This is not surprising. In India where most households are not nuclear, it is likely that income from different sources and accruing to different individuals have different uses; child's education is the responsibility of the father.

References

Anh, T. S., J. Knodel, D. Lam, and J. Friedman (1998). Family Size and Childrens Education in Vietnam. *Demography 35*(1), 57 – 70.

Barro, R. (2001). Human capital and growth. *American Economic Review*, 12–17.

Becker, G., K. Murphy, and R. Tamura (1990). Human capital, fertility, and economic growth. *Journal of Political Economy*, 12–37.

Becker, G. and N. Tomes (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, 1–39.

Becker, G. S. and H. G. Lewis (1973). Interaction between quantity and quality of children. *Journal of Political Economy 81*(2), S279 – S288.

Behrman, J., N. Birdsall, and M. Szekely (2000). *Intergenerational Mobility in Latin America: Deeper Markets and Better Schools Make a Difference*, pp. 135 – 167. Washington DC: Brookings Institution.

Behrman, J. and M. Rosenzweig (2002). Does increasing women's schooling raise the schooling of the next generation? *American Economic Review 92*(1), 323–334.

Behrman, J. R., A. Pollak, and P. Taubman (1989). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy 97*(2), 389 – 419.

Bourguignon, F., F. Ferreira, and M. Menendez (2003). Inequality of outcomes and inequality of opportunities in brazil. Technical Report 3174, World Bank Policy Research Working Paper.

Chernichovsky, D. (1985). Socioeconomic and Demographic Aspects of School Enrolment and Attendance in Rural Botswana. *Economic Development and Cultural Change 33*, 319 – 332.

Conley, D. and R. Glauber (2006). Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility. *Journal of Human Resources XLI*(4), 722 – 735.

Cornwell, K., B. Inder, P. Maitra, and A. Rammohan (2007). Household composition and schooling of rural south african children: Sibling synergy and migrant effects. Technical report, Monash University, Australia.

Dreze, J. and G. Kingdon (2001). School Participation in Rural India. *Review of Development Economics 5*(1), 1–24.

Gau, W. and R. Smyth (2009). Health Human Capital, Height and Wages in China. *Journal of Development studies Forthcoming*.

Gomes, M. (1984). Family Size and Educational Attainment in Kenya. *Population and Development Review 10*, 647 – 660.

Goux, D. and E. Maurin (2004). The Effects of Overcrowded Housing on Childrens Performance at School. *CEPR Discussion Paper # 3818*.

Greenhalgh, S. (1985). Sexual stratification: The other side of" growth with equity" in east Asia. *Population and Development Review*, 265–314.

Gupta, D. (2004). *Caste*. Oxford: Oxford University Press.

Holmes, J. (2003). The determinants of completed schooling in pakistan: Analysis of censoring and selection bias. *Economics of Education Review 22*, 249 – 264.

Jalan, J. and R. Murgai (2008). Intergenerational mobility in education in india. Technical report, Paper Presented at the Indian Statistical Institute, Delhi.

Kingdon, G. G. (2005). Where has all the bias gone? Detecting gender bias in the Intra-household Allocation of Educational Expenditure. *Economic Development and Cultural Change 53*(2), 409 – 452.

Kingdon, G. G. (2007). The progress of school education in India. *Oxford Review of Economic Policy 23*(2), 168 – 195.

Knodel, J., N. Havanon, and W. Sittirai (1990). Family Size and the Education of Children in the Context of Rapid Fertility Decline. *Population and Development Review 16*, 31 – 62.

Lillard, L. and R. J. Willis (1994). Intergenerational educational mobility: Effects of family and state in malaysia. *Journal of Human Resources 29*(4), 1126–1167.

Maitra, P. (2003). Schooling and educational attainment: Evidence from bangladesh. *Education Economics 11*(2), 129 – 153.

Malarani, V. (2004). Family size and educational attainment in indonesia: A cohort perspective. Technical report, California Center for Population Research, University of California, Los Angeles.

Mulder, M. (1998). Brothers and sisters. *Human Nature 9*(2), 119–161.

Munshi, K. and M. R. Rosenzweig (2009). Why is mobility in india so low? social insurance, inequality and growth. Technical report, Brown University.

Pal, S. (2004). Child schooling in peru: Evidence from a sequential analysis of school progression. *Journal of Population Economics 17*(4), 657–680.

Parish, W. and R. Willis (1993). Daughters, education, and family budgets Taiwan experiences. *Journal of Human Resources*, 863–898.

Petersen, T. (1985). A comment on presenting results from logit and probit models. *American Sociological Review*, 130–131.

PROBE (1999). *Public Report on Basic Education in India.* Delhi: Oxford University Press.

Psacharopoulos, G. and A. M. Arriagada (1989). The Determinants of Early Age Human Capital Formation: Evidence from Brazil. *Economic Development and Cultural Change 37*, 683 – 708.

Qian, N. (2009). Quantity-quality and the one child policy:the only-child disadvantage in school enrollment in rural china. Technical report, NBER Working Papers #14973, National Bureau of Economic Research, Inc.

Rivers, D. and Q. H. Vuong (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics 39*, 347 – 366.

Rosenzweig, M. and K. Wolpin (1994). Are there increasing returns to the intergenerational production of human capital? Maternal schooling and child intellectual achievement. *Journal of Human Resources*, 670–693.

Schultz, T. (2002). Wage gains associated with height as a form of health human capital. *American Economic Review*, 349–353.

Shavit, Y. and J. L. Pierce (1991). Sibship Size and Educational Attainment in Nuclear and Extended Families: Arabs and Jews in Israel. *American Sociological Review 56*, 321 – 330.

Singh, R. (1992). Underinvestment, low economic returns to education, and the schooling of rural children: some evidence from Brazil. *Economic Development and Cultural Change*, 645–664.

Spierings, J. (1999). A crucial point in life: learning, work and young adults. australia's young adults: the deepening divide. Technical report, Dusseldorp Skills Forum: 5-27.

Thomas, D. (1996). Education across generations in south africa. *American Economic Review, Papers and Proceedings 86*(2), 330 – 334.

UN (1999). *World population ageing 1950-2050.* UN.

Table 1: Descriptive Statistics

Variable	Sample 1		Sample 2		Sample 3	
	Rural	Urban	Rural	Urban	Rural	Urban
	(1)	(2)	(3)	(4)	(5)	(6)
Brahmin	0.0453	0.0914	0.0499	0.0931	0.0451	0.0798
Other backward caste	0.4051	0.3776	0.4133	0.3734	0.4054	0.3781
Scheduled caste	0.2093	0.1559	0.1842	0.1463	0.2076	0.1677
Scheduled tribe	0.1002	0.0326	0.0884	0.0312	0.1099	0.0357
Other caste	0.2400	0.3425	0.2642	0.3560	0.2321	0.3386
Hindu	0.8266	0.7708	0.8157	0.7506	0.7967	0.7421
Muslim	0.0929	0.1564	0.0998	0.1726	0.1179	0.1891
Christian	0.0284	0.0346	0.0280	0.0339	0.0291	0.0302
Other religion	0.0521	0.0382	0.0564	0.0429	0.0563	0.0386
Poor	0.2083	0.1892	0.1908	0.1820	0.2053	0.1871
Jammu and Kashmir	0.0176	0.0251	0.0272	0.0378	0.0260	0.0320
Himachal Pradesh	0.0418	0.0216	0.0426	0.0227	0.0398	0.0200
Uttarkhand	0.0114	0.0105	0.0085	0.0117	0.0127	0.0114
Punjab	0.0440	0.0462	0.0512	0.0559	0.0486	0.0511
Haryana	0.0552	0.0178	0.0598	0.0168	0.0628	0.0188
Delhi	0.0024	0.0553	0.0033	0.0562	0.0032	0.0704
Bihar	0.0355	0.0330	0.0333	0.0371	0.0347	0.0393
Jharkhand	0.0191	0.0278	0.0206	0.0256	0.0222	0.0305
Rajasthan	0.0605	0.0598	0.0573	0.0574	0.0631	0.0647
Chattisgarh	0.0327	0.0194	0.0300	0.0203	0.0326	0.0191
Madhya Pradesh	0.0822	0.0443	0.0801	0.0456	0.0803	0.0452
Northeast	0.0236	0.0209	0.0262	0.0265	0.0309	0.0211
Assam	0.0245	0.0205	0.0304	0.0247	0.0293	0.0237
West Bengal	0.0421	0.0741	0.0359	0.0672	0.0398	0.0607
Orissa	0.0549	0.0406	0.0531	0.0406	0.0517	0.0381
Gujarat	0.0468	0.0598	0.0442	0.0557	0.0482	0.0556
Maharashtra	0.0848	0.0845	0.0916	0.0851	0.0757	0.0850
Andhra Pradesh	0.0497	0.0560	0.0354	0.0442	0.0399	0.0458
Karnataka	0.1094	0.0806	0.1220	0.0753	0.1098	0.0718
Kerala	0.0415	0.0487	0.0440	0.0516	0.0330	0.0361
Tamil Nadu	0.0311	0.0772	0.0263	0.0664	0.0264	0.0651
Uttar Pradesh	0.0890	0.0762	0.0770	0.0758	0.0891	0.0945
Years of Schooling	4.3410	7.6860	4.3352	7.6872	7.5765	9.5019
Male	0.5015	0.5042	0.5011	0.5038	0.6315	0.5759
Age	40.5411	39.2559	40.5415	39.2589	18.4693	18.7128
Years of Schooling: Father			4.0894	7.6002	4.6327	7.8618
Years of Schooling: Mother			1.4763	4.3785	1.9965	5.1872
Brothers $0-5$					0.2315	0.1278
Sisters $0-5$					0.2116	0.1070
Brothers $6-10$					0.2468	0.1576
Sisters $6-10$					0.2164	0.1450
Brothers $11-14$					0.3384	0.2732
Sisters $11-14$					0.3104	0.2431
Brothers $15-24$					0.6920	0.7020
Sisters $15-24$					0.6985	0.6787
Males $25-60$					1.2451	1.2215
Females $25-60$					1.2003	1.1609
Male 60+					0.1644	0.1108
Female 60+					0.1667	0.1152
Currently in school					0.3759	0.5088

Sample 1: Sample of Adults, born 1985 or before
Sample 2: Sample of Adults with available parental education, born 1985 or before
Sample 3: Sample of Children, born $1981-1990$

Table 2: Levels and Growth of in Educational Attainment of Adults

	Urban Male (1)	**Urban Female** (2)	**Rural Male** (3)	**Rural Female** (4)
Panel A				
Years of Schooling (OLS)	0.0528***	0.1219***	0.1025***	0.1011***
	(0.0021)	(0.0020)	(0.0012)	(0.0012)
Sample Mean	8.8035	6.5496	5.6779	2.9961
Observations	21959	21592	40196	39954
Panel B				
Enrolled in School (Marginal Effects)	0.0026***	0.0084***	0.0076***	0.0159***
	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Sample Mean	0.8727	0.7492	0.3481	0.2163
Observations	29973	29079	56875	55601
Panel C				
Enrolled in Middle School (Marginal Effects)	0.0024***	0.0055***	0.0069***	0.0101***
	(0.0001)	(0.0002)	(0.0001)	(0.0003)
Sample Mean	0.6979	0.5605	0.2295	0.1369
Observations	22800	18486	34718	21782
Panel D				
Enrolled in Post-Secondary School (Marginal Effects)	0.0022***	0.0070***	0.0046***	0.0075***
	(0.0003)	(0.0004)	(0.0003)	(0.0005)
Sample Mean	0.6948	0.4988	0.1465	0.0647
Observations	16452	12102	20051	10217
Panel E				
Enrolled in College (Marginal Effects)	-0.0017***	0.0007	-0.0007	-0.0007
	(0.0005)	(0.0008)	(0.0007)	(0.0016)
Sample Mean	0.4041	0.2557	0.0573	0.0241
Observations	6127	3457	4304	1239

Notes:
$^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$
Standard errors clustered at the household level
Adults defined as born in 1985 or earlier

Table 3: Levels and Growth of in Educational Attainment of Adults: by Religion and Caste

	All	**Hindu**	**Muslim**	**Christian**	**Other Religion**	**Brahmin**	**OBC**	**SC**	**ST**	**Other Caste**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Urban Males										
Growth per year	0.0528***	0.0571***	0.0563***	0.0472***	0.0545***	0.0397***	0.0710***	0.0895***	0.0755***	0.0378***
	(0.0021)	(0.0024)	(0.0049)	(0.0100)	(0.0098)	(0.0052)	(0.0033)	(0.0051)	(0.0121)	(0.0034)
Sample mean	8.80	9.17	6.49	9.94	10.06	11.59	8.11	7.11	8.24	9.65
Sample size	21959	16945	3463	730	821	2010	8300	3408	706	7535
Urban Females										
Growth per year	0.1219***	0.1254***	0.1255***	0.1070***	0.1437***	0.1349***	0.1332***	0.1344***	0.1316***	0.1153***
	(0.0020)	(0.0023)	(0.0047)	(0.0102)	(0.0090)	(0.0058)	(0.0030)	(0.0047)	(0.0126)	(0.0035)
Sample mean	6.55	6.78	4.42	9.08	8.21	9.22	5.75	4.55	6.26	7.66
Sample size	21592	16625	3347	777	843	1972	8143	3380	715	7382
Rural Males										
Growth per year	0.1025***	0.1058***	0.0897***	0.0720***	0.1081***	0.0983***	0.1076***	0.1133***	0.0974***	0.1003***
	(0.0012)	(0.0014)	(0.0040)	(0.0074)	(0.0050)	(0.0054)	(0.0019)	(0.0025)	(0.0040)	(0.0025)
Sample mean	5.68	5.72	4.89	7.36	5.57	8.97	5.70	4.61	4.04	6.65
Sample size	40196	33180	3775	1109	2132	1796	16216	8463	4037	9684
Rural Females										
Growth per year	0.1011***	0.1033***	0.0823***	0.1019***	0.1203***	0.1499***	0.1029***	0.0874***	0.0754***	0.1176***
	(0.0012)	(0.0012)	(0.0036)	(0.0075)	(0.0052)	(0.0051)	(0.0018)	(0.0023)	(0.0033)	(0.0024)
Sample mean	3.00	2.92	2.40	6.20	3.53	5.01	2.89	2.03	1.89	4.10
Sample size	39954	33070	3669	1171	2044	1837	16255	8316	3994	9552

Notes:
$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$
Standard errors clustered at the household level
Adults defined as born in 1985 or earlier

Table 4: Parental Education and Children's Education. OLS and Instrumental Variable Regression

	Urban Male		**Urban Female**		**Rural Male**		**Rural Female**	
	OLS	**IV Reg**	**OLS**	**IV Reg**	**OLS**	**IV Reg**	**OLS**	**IV Reg**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Year of Birth	-0.0463***	-0.0241	0.0570***	0.0660**	0.0153***	0.0154	0.1518***	0.1756***
	(0.0073)	(0.0395)	(0.0221)	(0.0334)	(0.0059)	(0.0111)	(0.0163)	(0.0250)
Father: Years of Schooling	0.3332***	0.6197	0.2708***	-0.1949	0.3831***	-0.0955	0.3914***	0.2562
	(0.0146)	(0.8680)	(0.0265)	(0.5581)	(0.0106)	(0.1882)	(0.0267)	(0.3147)
Mother: Years of Schooling	0.1308***	-0.4940	0.2047***	0.4051	0.1524***	0.7752*	0.2470***	0.1374
	(0.0130)	(1.2969)	(0.0222)	(0.4928)	(0.0149)	(0.4305)	(0.0317)	(0.2503)
Muslim	-1.5454***	-2.2743**	-1.0155***	-1.4475**	-1.0505***	-1.0397***	-1.1199***	-1.3564***
	(0.1579)	(1.0710)	(0.3221)	(0.6671)	(0.1517)	(0.2114)	(0.3791)	(0.4177)
Christian	-0.8051***	-0.0212	-0.2756	-0.1304	-0.1417	-0.7105	0.6057	0.7845
	(0.2581)	(1.4372)	(0.3373)	(0.4151)	(0.2803)	(0.6737)	(0.4178)	(0.5022)
Other religion	-0.3544	0.1583	-0.2407	0.1678	-0.7811***	-0.8757***	-0.8772*	-0.9177*
	(0.2286)	(0.9329)	(0.2979)	(0.5219)	(0.2338)	(0.2739)	(0.5188)	(0.5134)
Poor	-1.5356***	-2.2232***	-1.8189***	-3.0193***	-1.6817***	-2.0718***	-1.8829***	-2.4918***
	(0.1603)	(0.8297)	(0.2961)	(1.1384)	(0.1154)	(0.2224)	(0.2879)	(0.6672)
Constant	98.7989***	55.7278	-105.6707**	-120.2461*	-23.6804**	-22.2034	-296.7307***	-342.6647***
	(14.3543)	(78.4342)	(43.7773)	(65.8116)	(11.6766)	(22.0450)	(32.2313)	(48.2645)
Observations	5789	5789	1886	1886	11515	11515	2078	2078
Hansen J-statistic		0.048		0.799		0.538		0.905
Sargan Statistic†		0.092		0.811		0.724		1.178

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Standard errors clustered at the household level

Regressions include state dummies. Reference category: household resides in Uttar Pradesh

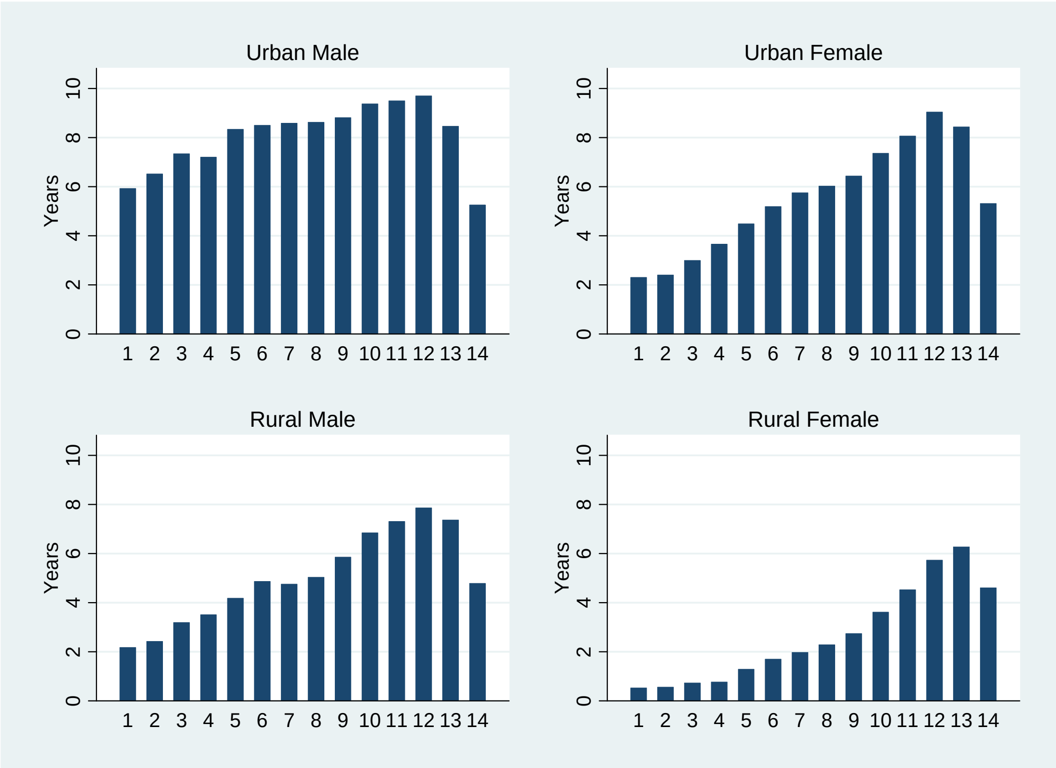
† : Sargan Statistics computed assuming IID error structure

Table 5: Parental Education and Children's Education. Estimates from a Sequential Probit Regression of Educational Attainment

	Male Urban			Female Urban			Male Rural			Female Rural		
	Spec 1	**Spec 2**	**Spec 3**	**Spec 1**	**Spec 2**	**Spec 3**	**Spec 1**	**Spec 2**	**Spec 3**	**Spec 1**	**Spec 2**	**Spec 3**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(19)	(11)	(12)
School Enrolment												
Years of Schooling:	0.1630 ***	0.1163 ***	-0.0477	0.1657 ***	0.1201 ***	-0.0181	0.1663 ***	0.1236 ***	-0.0131	0.1680 ***	0.1244 ***	-0.0132
Father	(0.0067)	(0.0261)	(0.0898)	(0.0066)	(0.0254)	(0.1079)	(0.0069)	(0.0260)	(0.0967)	(0.0070)	(0.0260)	(0.0948)
Years of Schooling:	0.1589 ***	0.3342 ***	0.3639***	0.1560 ***	0.3274 ***	0.3428***	0.1587 ***	0.3239 ***	0.3170***	0.1569 ***	0.3211 ***	0.3170***
Mother	(0.0129)	(0.0296)	(0.0975)	(0.0128)	(0.0287)	(0.1183)	(0.0133)	(0.0293)	(0.1050)	(0.0133)	(0.0293)	(0.1028)
Continuing to Middle/Secondary School												
Years of Schooling:	0.1506 ***	0.1990 ***	0.2356***	0.1456 ***	0.1808 ***	-0.0008	0.1582 ***	0.2066 ***	0.0886	0.1575 ***	0.1951 ***	-0.0098
Father	(0.0049)	(0.0200)	(0.0771)	(0.0048)	(0.0195)	(0.0845)	(0.0049)	(0.0197)	(0.0816)	(0.0050)	(0.0200)	(0.0807)
Years of Schooling:	0.1418 ***	0.2293 ***	0.0582	0.1418 ***	0.2299 ***	0.2698***	0.1540 ***	0.2503 ***	0.2521***	0.1536 ***	0.2545 ***	0.3408***
Mother	(0.0073)	(0.0216)	(0.0830)	(0.0073)	(0.0209)	(0.0915)	(0.0076)	(0.0214)	(0.0884)	(0.0075)	(0.0216)	(0.0879)
Continuing to Post-Secondary/College												
Years of Schooling:	0.1135 ***	0.2546 ***	0.2499***	0.1109 ***	0.2460 ***	0.1735**	0.1171 ***	0.2624 ***	0.2193***	0.1168 ***	0.2585 ***	0.1688**
Father	(0.0039)	(0.0155)	(0.0697)	(0.0039)	(0.0154)	(0.0689)	(0.0040)	(0.0158)	(0.0718)	(0.0040)	(0.0158)	(0.0720)
Years of Schooling:	0.0924 ***	0.0561 ***	0.0216	0.0919 ***	0.0554 ***	0.0831	0.0966 ***	0.0609 ***	0.0670	0.0960 ***	0.0619 ***	0.1186
Mother	(0.0042)	(0.0149)	(0.0732)	(0.0042)	(0.0148)	(0.0722)	(0.0043)	(0.0152)	(0.0755)	(0.0043)	(0.0152)	(0.0757)
σ^2_δ	1.0728 ***	1.0370 ***	1.0434***	1.0568 ***	1.0255 ***	1.0219***	1.1217 ***	1.0797 ***	1.0922***	1.1146 ***	1.0750 ***	1.0858***
	(0.0154)	(0.0153)	(0.0160)	(0.0155)	(0.0153)	(0.0162)	(0.0152)	(0.0150)	(0.0158)	(0.0153)	(0.0151)	(0.0159)
ln-L	-28112.02	-27756.61	-25064.83	-28087.68	-27762.15	-24972.29	-28431.91	-28062.43	-25332.71	-28415.57	-28056.52	-25307.72
Endogeneity of Parental Education†		710.8166***	43.0322***		651.0582***	25.2598***		738.965***	47.6914***		718.1008***	42.6954***
Significance of Additional Controls‡			5383.56***			5579.72***			5459.44***			5497.60***

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$
All standard errors are Huber-corrected.
$^{\dagger}LR = -2(L_R - L_U) \sim \chi^2(6)$
$^{\ddagger}LR = -2(L_R - L_U) \sim \chi^2(111)$
Regressions include state dummies. Reference category: household resides in Uttar Pradesh
Spec. 3: Full set of results presented in Table A-1

Figure 1: Average Years of Schooling by Cohort of Birth



Cohort 1 = Year of Birth $\leq$ 1930	Cohort 2 = Year of Birth $>$ 1930 & Year of Birth $\leq$ 1935
Cohort 3 = Year of Birth $>$ 1935 & Year of Birth $\leq$ 1940	Cohort 4 = Year of Birth $>$ 1940 & Year of Birth $\leq$ 1945
Cohort 5 = Year of Birth $>$ 1945 & Year of Birth $\leq$ 1950	Cohort 6 = Year of Birth $>$ 1950 & Year of Birth $\leq$ 1955
Cohort 7 = Year of Birth $>$ 1955 & Year of Birth $\leq$ 1960	Cohort 8 = Year of Birth $>$ 1960 & Year of Birth $\leq$ 1965
Cohort 9 = Year of Birth $>$ 1965 & Year of Birth $\leq$ 1970	Cohort 10 = Year of Birth $>$ 1970 & Year of Birth $\leq$ 1975
Cohort 11 = Year of Birth $>$ 1975 & Year of Birth $\leq$ 1980	Cohort 12 = Year of Birth $>$ 1980 & Year of Birth $\leq$ 1985
Cohort 13 = Year of Birth $>$ 1985 & Year of Birth $\leq$ 1990	Cohort 14 = Year of Birth $>$ 1990 & Year of Birth $\leq$ 1995

Figure 2: Lowess Plot of Years of Schooling by Year of Birth

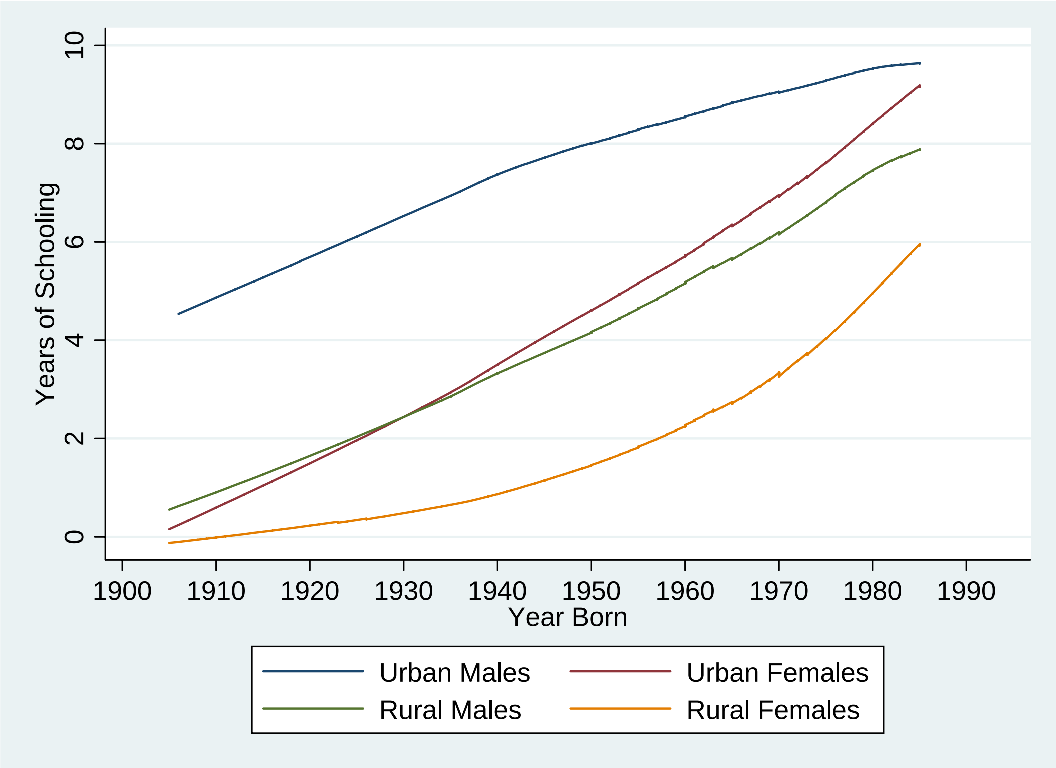


Figure 3: Lowess Plot of Years of Schooling by Years of Schooling Attained by Father and Mother

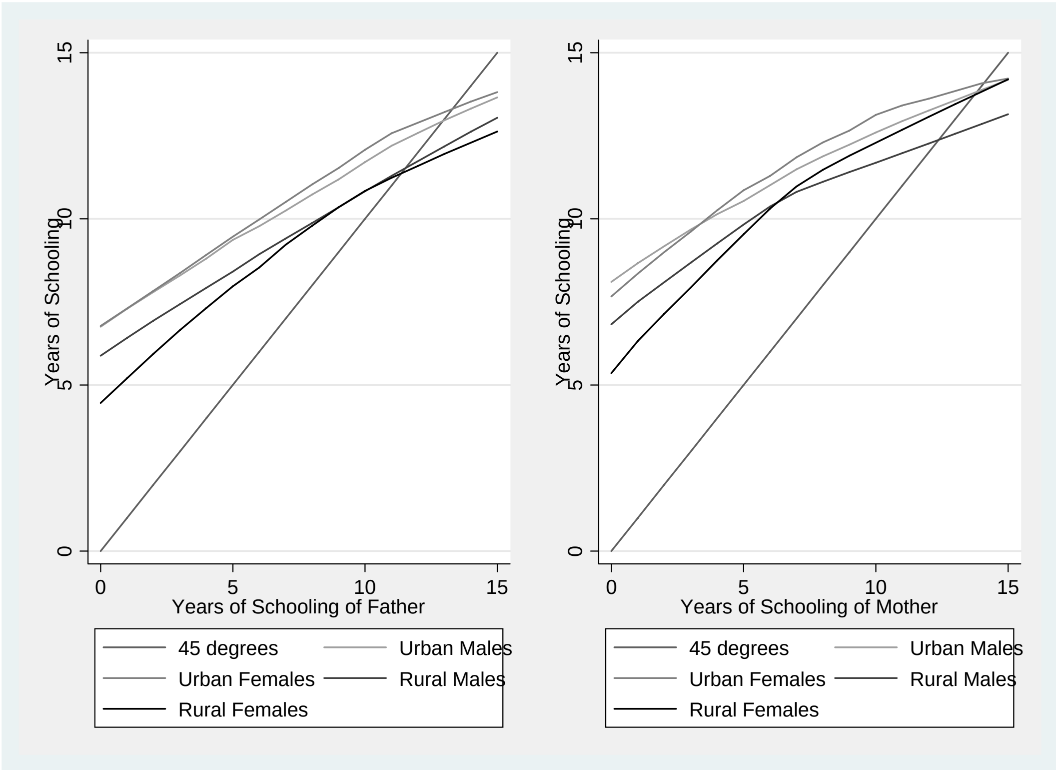


Figure 4: Schematic Representation of the Sequential Framework

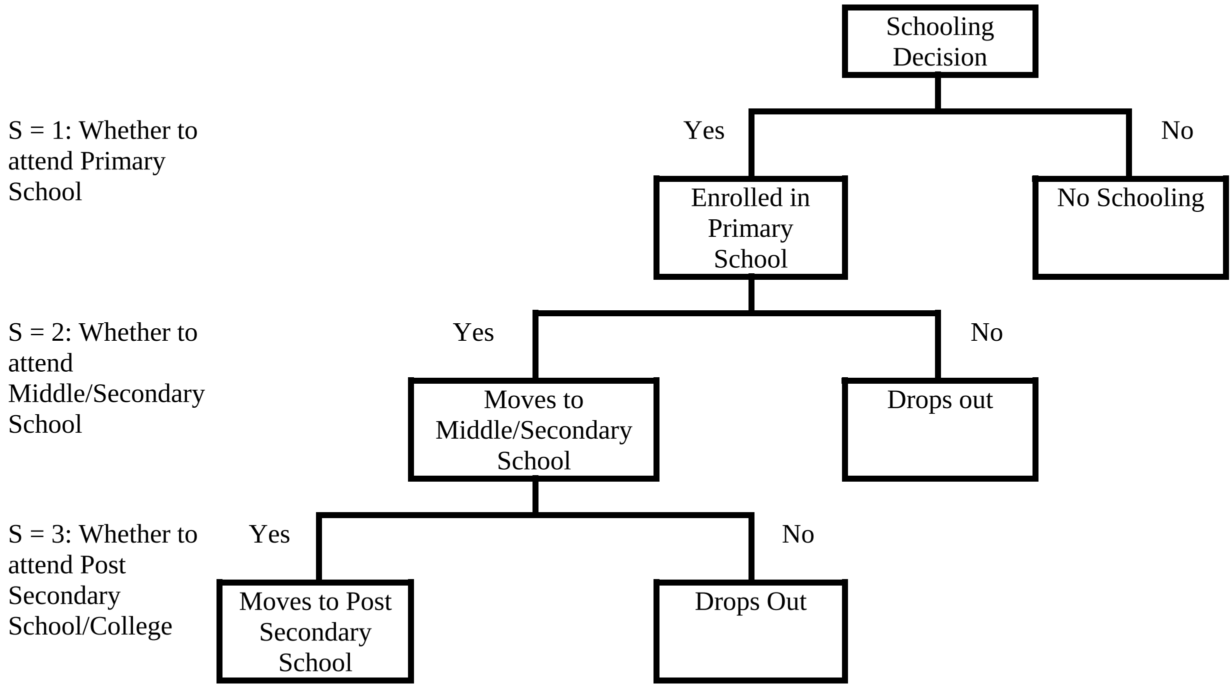


Figure 5: School Progression Rates

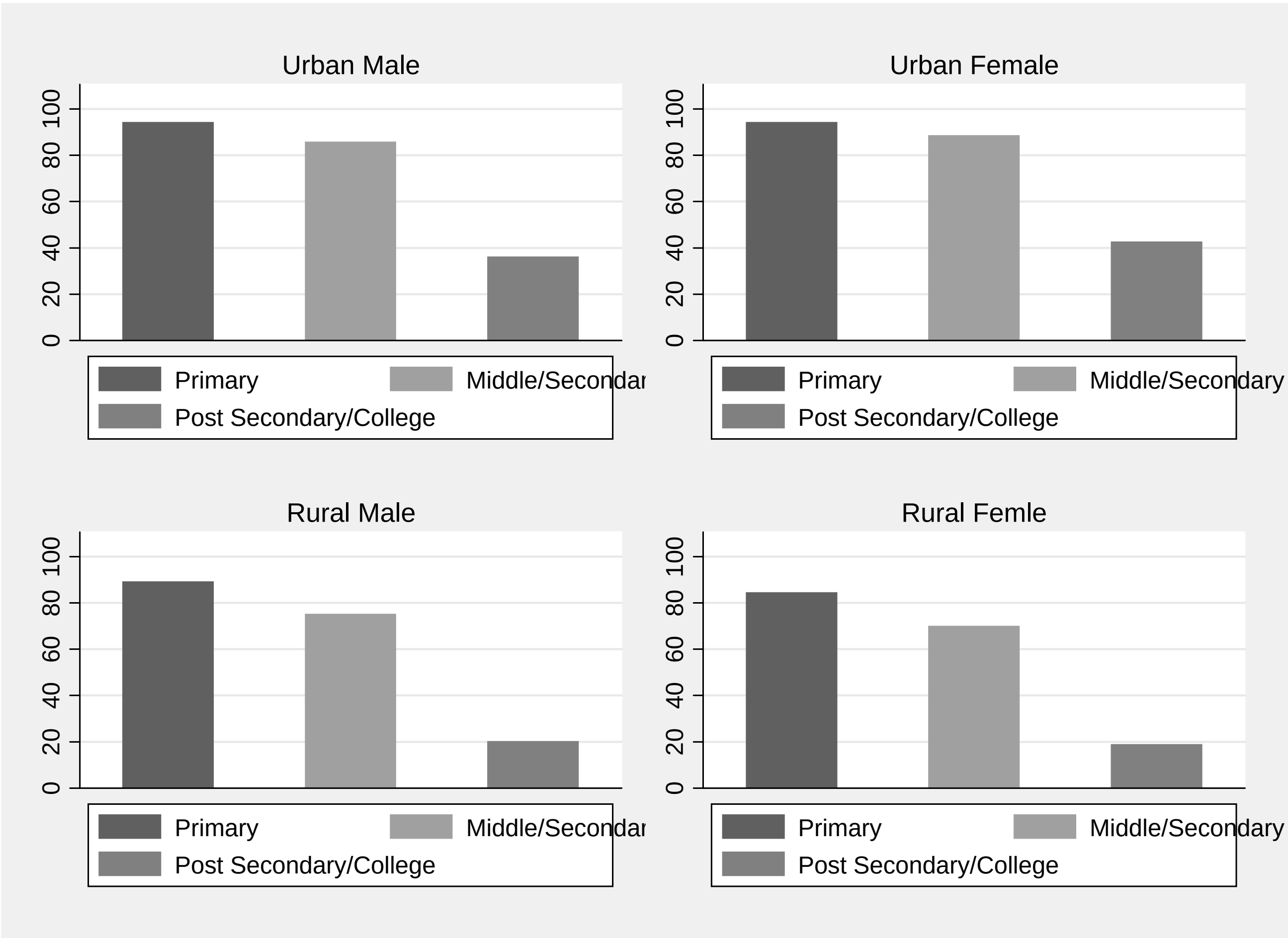
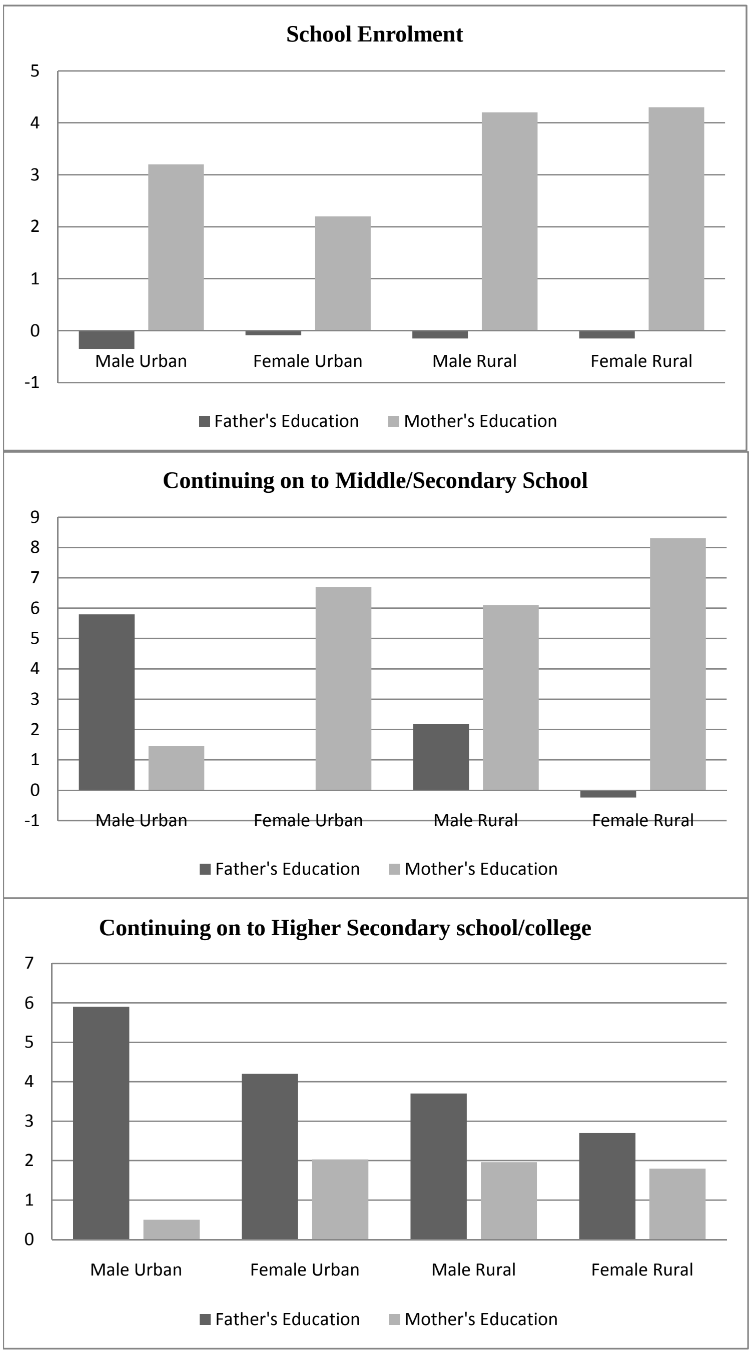


Figure 6: Changes in the Probability of schooling



A.1 Appendix

Table A-1: Sequential Probit Regression of Educational Attainment. Full Set of Estimates

	Male Urban		Female Urban		Male Rural		Female Rural	
	Exog.	**Endog.**	**Exog.**	**Endog.**	**Exog.**	**Endog.**	**Exog.**	**Endog.**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
School Enrolment								
Constant	4.0620***	4.3542***	4.0270***	4.2522***	4.2218***	4.5363***	4.2511***	4.5583***
	(0.8659)	(0.8871)	(0.8564)	(0.9410)	(0.8945)	(0.9211)	(0.8884)	(0.9152)
Age	-0.3076***	-0.2855***	-0.3021***	-0.2837***	-0.3158***	-0.3006***	-0.3203***	-0.3049***
	(0.0921)	(0.0922)	(0.0909)	(0.0910)	(0.0952)	(0.0951)	(0.0945)	(0.0943)
Age Square	0.0075***	0.0070***	0.0072***	0.0068***	0.0076***	0.0073***	0.0077***	0.0074***
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Years of Schooling: Father	0.1540***	-0.0477	0.1547***	-0.0181	0.1573***	-0.0131	0.1576***	-0.0132
	(0.0064)	(0.0898)	(0.0064)	(0.1079)	(0.0065)	(0.0967)	(0.0066)	(0.0948)
Years of Schooling: Mother	0.1239***	0.3639***	0.1226***	0.3428***	0.1245***	0.3170***	0.1209***	0.3170***
	(0.0127)	(0.0975)	(0.0126)	(0.1183)	(0.0132)	(0.1050)	(0.0137)	(0.1028)
Brothers Aged 0 – 5	-0.1453***	-0.1460***	-0.1650***	-0.1547***	-0.1580***	-0.1676***	-0.1544***	-0.1595***
	(0.0373)	(0.0408)	(0.0374)	(0.0466)	(0.0381)	(0.0419)	(0.0379)	(0.0417)
Sisters aged 0 – 5	-0.0988***	-0.1022**	-0.1142***	-0.1096**	-0.1032***	-0.1124***	-0.1085***	-0.1159***
	(0.0367)	(0.0402)	(0.0364)	(0.0517)	(0.0384)	(0.0418)	(0.0385)	(0.0419)
Brothers aged 6 – 10	-0.2940***	-0.3079***	-0.2844***	-0.2891***	-0.2975***	-0.3144***	-0.2977***	-0.3136***
	(0.0349)	(0.0383)	(0.0353)	(0.0392)	(0.0367)	(0.0408)	(0.0372)	(0.0409)
Sisters aged 6 – 10	-0.0904**	-0.0804**	-0.0875**	-0.0735*	-0.0856**	-0.0804**	-0.0878**	-0.0838**
	(0.0359)	(0.0371)	(0.0360)	(0.0412)	(0.0371)	(0.0384)	(0.0379)	(0.0393)
Brothers aged 11 – 14	-0.1300***	-0.0910**	-0.1214***	-0.0845**	-0.1325***	-0.1037***	-0.1316***	-0.1013**
	(0.0349)	(0.0380)	(0.0348)	(0.0416)	(0.0363)	(0.0392)	(0.0363)	(0.0394)
Sisters aged 11 – 14	0.0531	0.0868**	0.0582	0.0865*	0.0619	0.0870**	0.0588	0.0879**
	(0.0374)	(0.0405)	(0.0383)	(0.0498)	(0.0391)	(0.0421)	(0.0396)	(0.0429)
Brothers aged 15 – 24	-0.1354***	-0.1188***	-0.1716***	-0.1525***	-0.1640***	-0.1542***	-0.1744***	-0.1596***
	(0.0245)	(0.0255)	(0.0243)	(0.0315)	(0.0252)	(0.0257)	(0.0256)	(0.0260)
Sisters aged 15 – 24	0.0742***	0.1166***	0.1015***	0.1371***	0.0990***	0.1319***	0.1054***	0.1411***
	(0.0262)	(0.0300)	(0.0260)	(0.0299)	(0.0270)	(0.0306)	(0.0279)	(0.0314)
Males aged 25 – 60	0.0378	0.0720	0.0365	0.0607	0.0322	0.0608	0.0420	0.0693
	(0.0406)	(0.0447)	(0.0406)	(0.0484)	(0.0429)	(0.0469)	(0.0410)	(0.0448)
Females aged 25 – 60	0.2008***	0.2147***	0.2128***	0.2086***	0.2080***	0.2259***	0.2095***	0.2237***
	(0.0502)	(0.0557)	(0.0506)	(0.0775)	(0.0540)	(0.0595)	(0.0534)	(0.0585)
Muslim	-0.4139***	-0.4550***	-0.4073***	-0.4359***	-0.4401***	-0.4802***	-0.4493***	-0.4865***
	(0.0602)	(0.0660)	(0.0591)	(0.0661)	(0.0610)	(0.0670)	(0.0634)	(0.0684)
Christian	-0.4145**	-0.6510***	-0.3576*	-0.5872**	-0.4450**	-0.6246***	-0.4459**	-0.6320***
	(0.1859)	(0.2099)	(0.1975)	(0.2331)	(0.2002)	(0.2276)	(0.2034)	(0.2278)
Other Religion	-0.4421***	-0.5141***	-0.4577***	-0.5208***	-0.4737***	-0.5397***	-0.4596***	-0.5270***

Continued . . .

Table A-1 (continued)

	Male Urban		Female Urban		Male Rural		Female Rural	
	Exog.	**Endog.**	**Exog.**	**Endog.**	**Exog.**	**Endog.**	**Exog.**	**Endog.**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.1280)	(0.1307)	(0.1281)	(0.1322)	(0.1404)	(0.1444)	(0.1353)	(0.1399)
Poor	-0.4933***	-0.5909***	-0.4900***	-0.5479***	-0.5278***	-0.6274***	-0.5213***	-0.6160***
	(0.0491)	(0.0820)	(0.0471)	(0.0874)	(0.0502)	(0.0862)	(0.0501)	(0.0856)
Residual Father's Schooling		0.2017**		0.1725		0.1709*		0.1709*
		(0.0901)		(0.1095)		(0.0971)		(0.0950)
Residual Mother's Schooling		-0.2424**		-0.2230*		-0.1941*		-0.1983*
		(0.0970)		(0.1175)		(0.1047)		(0.1025)
Continuation to Middle/Secondary School								
Constant	-0.6418	-0.9857	-0.7815	-0.5686	-0.4466	-0.5530	-0.4250	-0.2896
	(0.7227)	(0.7418)	(0.7219)	(0.7671)	(0.7275)	(0.7525)	(0.7263)	(0.7510)
Age	0.0710	0.0658	0.0807	0.0972	0.0397	0.0511	0.0382	0.0570
	(0.0766)	(0.0768)	(0.0768)	(0.0771)	(0.0772)	(0.0774)	(0.0770)	(0.0773)
Age squared	-0.0016	-0.0016	-0.0016	-0.0019	-0.0007	-0.0010	-0.0007	-0.0011
	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)
Years of Schooling: Father	0.1422***	0.2356***	0.1357***	-0.0008	0.1497***	0.0886	0.1480***	-0.0098
	(0.0049)	(0.0771)	(0.0049)	(0.0845)	(0.0050)	(0.0816)	(0.0050)	(0.0807)
Years of Schooling: Mother	0.1115***	0.0582	0.1125***	0.2698***	0.1234***	0.2521***	0.1214***	0.3408***
	(0.0074)	(0.0830)	(0.0074)	(0.0915)	(0.0077)	(0.0884)	(0.0079)	(0.0879)
Brothers aged 0 – 5	-0.1438***	-0.0962***	-0.1433***	-0.1459***	-0.1638***	-0.1223***	-0.1609***	-0.1362***
	(0.0339)	(0.0372)	(0.0346)	(0.0451)	(0.0343)	(0.0378)	(0.0345)	(0.0379)
Sisters aged 0 – 5	-0.1146***	-0.0674*	-0.0815**	-0.0910*	-0.1215***	-0.0799**	-0.1275***	-0.1063***
	(0.0337)	(0.0366)	(0.0344)	(0.0489)	(0.0347)	(0.0376)	(0.0345)	(0.0374)
Brothers aged 6 – 10	-0.2659***	-0.2295***	-0.2730***	-0.2840***	-0.2804***	-0.2561***	-0.2857***	-0.2807***
	(0.0323)	(0.0348)	(0.0320)	(0.0350)	(0.0329)	(0.0360)	(0.0330)	(0.0359)
Sisters aged 6 – 10	-0.0918***	-0.0761**	-0.1058***	-0.1002***	-0.1036***	-0.0823**	-0.1089***	-0.0937***
	(0.0338)	(0.0347)	(0.0327)	(0.0376)	(0.0343)	(0.0355)	(0.0342)	(0.0355)
Brothers aged 11 – 14	-0.0717**	-0.0713**	-0.0918***	-0.0691*	-0.0987***	-0.0688**	-0.0967***	-0.0561*
	(0.0307)	(0.0332)	(0.0303)	(0.0371)	(0.0314)	(0.0340)	(0.0311)	(0.0337)
Sisters aged 11 – 14	0.0436	0.0374	0.0117	0.0305	0.0312	0.0508	0.0265	0.0604*
	(0.0324)	(0.0348)	(0.0326)	(0.0457)	(0.0336)	(0.0365)	(0.0335)	(0.0365)
Brothers aged 15 – 24	-0.1173***	-0.1109***	-0.1465***	-0.1379***	-0.1449***	-0.1238***	-0.1543***	-0.1276***
	(0.0221)	(0.0231)	(0.0214)	(0.0301)	(0.0223)	(0.0231)	(0.0223)	(0.0230)
Sisters aged 15 – 24	0.0655***	0.0626**	0.1062***	0.1299***	0.0980***	0.1245***	0.1089***	0.1501***
	(0.0230)	(0.0266)	(0.0238)	(0.0282)	(0.0238)	(0.0273)	(0.0241)	(0.0280)
Males aged 25 – 60	0.0669*	0.0406	0.0374	0.0595	0.0618	0.0599	0.0718**	0.0905**
	(0.0357)	(0.0391)	(0.0357)	(0.0438)	(0.0379)	(0.0416)	(0.0363)	(0.0396)
Females aged 25 – 60	0.2373***	0.1824***	0.2597***	0.2670***	0.2593***	0.2188***	0.2558***	0.2399***
	(0.0436)	(0.0472)	(0.0440)	(0.0693)	(0.0459)	(0.0498)	(0.0449)	(0.0486)
Muslim	-0.4777***	-0.4304***	-0.4819***	-0.5097***	-0.4652***	-0.4464***	-0.4603***	-0.4709***
	(0.0539)	(0.0582)	(0.0524)	(0.0573)	(0.0548)	(0.0592)	(0.0568)	(0.0605)
Christian	-0.2668*	-0.2341	-0.2293	-0.3855**	-0.3074**	-0.4526***	-0.3003**	-0.5254***

Continued ...

Table A-1 (continued)

	Male Urban		Female Urban		Male Rural		Female Rural	
	Exog.	**Endog.**	**Exog.**	**Endog.**	**Exog.**	**Endog.**	**Exog.**	**Endog.**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.1427)	(0.1636)	(0.1476)	(0.1730)	(0.1471)	(0.1719)	(0.1472)	(0.1704)
Other Religion	-0.2226**	-0.2064**	-0.2402**	-0.2854***	-0.2821**	-0.3186***	-0.2723**	-0.3347***
	(0.1013)	(0.1043)	(0.1028)	(0.1066)	(0.1103)	(0.1143)	(0.1069)	(0.1109)
Poor	-0.5964***	-0.4568***	-0.5904***	-0.6613***	-0.6027***	-0.5379***	-0.5932***	-0.6152***
	(0.0434)	(0.0720)	(0.0418)	(0.0725)	(0.0440)	(0.0745)	(0.0438)	(0.0738)
Residual Father's Schooling		-0.0960		0.1368		0.0584		0.1558*
		(0.0773)		(0.0854)		(0.0817)		(0.0808)
Residual Mother's Schooling		0.0495		-0.1589*		-0.1343		-0.2245**
		(0.0831)		(0.0914)		(0.0885)		(0.0879)
Continuation to Post-Secondary School/College								
Constant	-30.7870***	-31.1957***	-30.4924***	-30.7388***	-31.2802***	-31.6097***	-31.2791***	-31.5224***
	(0.8204)	(0.8365)	(0.8156)	(0.8339)	(0.8436)	(0.8595)	(0.8411)	(0.8565)
Age	2.7461***	2.7345***	2.7230***	2.7200***	2.7853***	2.7749***	2.7871***	2.7822***
	(0.0827)	(0.0828)	(0.0822)	(0.0825)	(0.0849)	(0.0851)	(0.0847)	(0.0848)
Age squared	-0.0645***	-0.0644***	-0.0639***	-0.0639***	-0.0654***	-0.0652***	-0.0655***	-0.0654***
	(0.0021)	(0.0021)	(0.0021)	(0.0021)	(0.0021)	(0.0021)	(0.0021)	(0.0021)
Years of Schooling: Father	0.1227***	0.2499***	0.1180***	0.1735**	0.1269***	0.2193***	0.1261***	0.1688**
	(0.0045)	(0.0697)	(0.0045)	(0.0689)	(0.0046)	(0.0718)	(0.0046)	(0.0720)
Years of Schooling: Mother	0.0963***	0.0216	0.0957***	0.0831	0.1001***	0.0670	0.0992***	0.1186
	(0.0048)	(0.0732)	(0.0047)	(0.0722)	(0.0050)	(0.0755)	(0.0049)	(0.0757)
Brothers aged $0-5$	-0.1961***	-0.1328***	-0.1950***	-0.1488***	-0.2061***	-0.1445***	-0.2049***	-0.1471***
	(0.0393)	(0.0417)	(0.0402)	(0.0524)	(0.0400)	(0.0425)	(0.0404)	(0.0428)
Sisters aged $0-5$	-0.1113***	-0.0477	-0.1030***	-0.0630	-0.1195***	-0.0569	-0.1198***	-0.0631
	(0.0363)	(0.0389)	(0.0376)	(0.0525)	(0.0369)	(0.0398)	(0.0367)	(0.0395)
Brothers aged $6-10$	-0.1905***	-0.1403***	-0.1870***	-0.1543***	-0.2029***	-0.1558***	-0.2039***	-0.1639***
	(0.0377)	(0.0399)	(0.0384)	(0.0411)	(0.0387)	(0.0412)	(0.0383)	(0.0407)
Sisters aged $6-10$	-0.0796*	-0.0614	-0.0798*	-0.0641	-0.0797*	-0.0604	-0.0839*	-0.0667
	(0.0424)	(0.0425)	(0.0427)	(0.0471)	(0.0434)	(0.0437)	(0.0434)	(0.0437)
Brothers aged $11-14$	-0.0549*	-0.0580*	-0.0596**	-0.0567	-0.0617**	-0.0580*	-0.0644**	-0.0524
	(0.0298)	(0.0320)	(0.0299)	(0.0375)	(0.0306)	(0.0331)	(0.0305)	(0.0329)
Sisters aged $11-14$	0.0073	-0.0002	-0.0039	-0.0062	0.0025	0.0002	-0.0017	0.0051
	(0.0361)	(0.0378)	(0.0371)	(0.0551)	(0.0378)	(0.0403)	(0.0372)	(0.0393)
Brothers aged $15-24$	-0.0467**	-0.0434*	-0.0615***	-0.0575*	-0.0536**	-0.0470*	-0.0598**	-0.0471*
	(0.0227)	(0.0238)	(0.0230)	(0.0347)	(0.0235)	(0.0249)	(0.0232)	(0.0243)
Sisters aged $15-24$	0.0205	0.0123	0.0373*	0.0347	0.0356	0.0342	0.0410*	0.0484*
	(0.0212)	(0.0249)	(0.0217)	(0.0295)	(0.0221)	(0.0260)	(0.0220)	(0.0258)
Males aged $25-60$	0.0394	0.0033	0.0280	0.0063	0.0389	0.0069	0.0437	0.0213
	(0.0303)	(0.0328)	(0.0308)	(0.0393)	(0.0321)	(0.0347)	(0.0318)	(0.0342)
Females aged $25-60$	0.2508***	0.1814***	0.2530***	0.2037***	0.2589***	0.1925***	0.2601***	0.2021***
	(0.0380)	(0.0415)	(0.0383)	(0.0572)	(0.0394)	(0.0435)	(0.0391)	(0.0428)
Muslim	-0.3454***	-0.2775***	-0.3437***	-0.3028***	-0.3442***	-0.2802***	-0.3350***	-0.2849***

Continued ...

Table A-1 (continued)

	Male Urban		Female Urban		Male Rural		Female Rural	
	Exog.	**Endog.**	**Exog.**	**Endog.**	**Exog.**	**Endog.**	**Exog.**	**Endog.**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.0563)	(0.0602)	(0.0556)	(0.0604)	(0.0578)	(0.0617)	(0.0586)	(0.0623)
Christian	-0.1179	-0.0747	-0.1017	-0.1111	-0.1200	-0.1182	-0.1092	-0.1558
	(0.0987)	(0.1222)	(0.0999)	(0.1229)	(0.1006)	(0.1254)	(0.1005)	(0.1252)
Other Religion	-0.0782	-0.0609	-0.0831	-0.0827	-0.1006	-0.0967	-0.0931	-0.1032
	(0.0814)	(0.0832)	(0.0808)	(0.0827)	(0.0840)	(0.0858)	(0.0844)	(0.0860)
Poor	-0.5309***	-0.3501***	-0.5265***	-0.4151***	-0.5484***	-0.3813***	-0.5336***	-0.4016***
	(0.0529)	(0.0741)	(0.0522)	(0.0741)	(0.0543)	(0.0764)	(0.0546)	(0.0760)
Residual Father's Schooling		-0.1300*		-0.0571		-0.0952		-0.0455
		(0.0699)		(0.0691)		(0.0720)		(0.0722)
Residual Mother's Schooling		0.0701		0.0091		0.0279		-0.0246
		(0.0736)		(0.0726)		(0.0759)		(0.0761)
σ_δ^2	1.0483***	1.0434***	1.0239***	1.0219***	1.0983***	1.0922***	1.0908***	1.0858***
	(0.0161)	(0.0160)	(0.0163)	(0.0162)	(0.0159)	(0.0158)	(0.0160)	(0.0159)
ln-L	-25086.35	-25064.83	-24984.94	-24972.29	-25356.56	-25332.71	-25329.07	-25307.72
Endogeneity of Parental Education†	43.0322***		25.2598***		47.6914***		42.6954***	

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

All standard errors are Huber-corrected.

† : $LR - -2(L_R - L_U) \sim \chi^2(6)$

Regressions include state dummies. Reference category: household resides in Uttar Pradesh