Do returns to education matter to schooling participation?

by
Geeta Kingdon
Nicolas Theopold

Abstract

While it might be expected that schooling will depend positively on the economic returns to education (ER) in the local labor market, in fact there is theoretical ambiguity about the sign of the schooling-ER relationship when households are liquidity-constrained. Whether the relationship is positive or negative depends on which effect dominates – the positive substitution effect of an increase in ER on years of education, or the negative income effect. For India, we find a positive relationship between the rate of return to education for adults in the local labor market and school attainment of girls and non-poor boys. The size of the effect of ER on years of education acquired is large for some groups. However, for poor boys the negative income effect dominates the positive substitution effect. Thus, while policy efforts to increase the rate of return to education should have a positive impact on educational attainment of girls and non-poor boys, they may worsen educational attainment of poor boys. This suggests that policies to raise labor market returns to education (for raising educational attainment) be accompanied by policies to ease liquidity constraints in poor families.

JEL classification: I20, J31, J82
Keywords: schooling attainment, market returns to education, child labor, India
Number of words: total – 8507; not counting tables and references – 5838

Corresponding author: Geeta Gandhi Kingdon, Department of Economics, University of Oxford, Manor Road, Oxford, OX1 3UQ, United Kingdom. Tel: 44-1865-271065; Fax: 44-1865-281447
Email: geeta.kingdon@economics.oxford.ac.uk

Acknowledgments: We are grateful to Kenneth Swinnerton for helpful comments on the paper and to Måns Söderbom for helpful discussions on some of the issues in the paper. Any errors are ours.

Do returns to education matter to schooling participation?

1. Introduction

Much work in education economics focuses on explaining the decisions of individuals to acquire education. This entails a comparison of the cost and benefits of education. Demand for education is hypothesized to rise with the benefits of education and to fall with its costs. There is much analysis of role of supply-side measures in reducing the *costs* of school participation, e.g. reduction of school fees, direct cash subsidies, school-construction programs to reduce travel costs and the provision of non-monetary benefits in schools, such as school meals[1]. The efficacy of supply side measures in improving the quality of schooling, in order to increase the *benefits* of education, has also been analyzed. For instance, much research focuses on the effect of class size on pupil achievement and on labor market outcomes (Angrist and Lavy, 1999; Krueger, 1999; Krueger, 2003; Case and Deaton, 1999; Card and Krueger, 1992). Arguably, one the most powerful determinants of the demand for schooling is its expected economic benefits but there is little research to test whether and how much the expected return to schooling affects individuals' demand for it. This question is particularly important in less developed countries where compulsory education laws either do not exist or are not enforced, and large sections of the child population do not participate in school.

Measurement of the economic benefits of education has a long and rich history, starting with Mincer (1974) estimating the monetary returns from an additional year of education, using a semi-log framework. A series of reviews by Psacharopoulos (1985, 1994), Psacharopoulos and Patrinos (2004) and Card (2001) document the large number of studies in the field. However,

[1] Kremer and Chen, 2002; Schultz, 2004; Miguel and Kremer, 2004; Deininger, 2003; Drèze and Kingdon, 2001; Duflo, 2001; Vermeersch, 2002.

accurate estimation of returns to education is difficult and continues to be subject of debate (Bennell, 1996; Card, 2001). Despite the methodological discussion however, Mincerian returns to education are widely used as a measure of the economic benefits of education.

The role of economic returns to education in the determination of schooling participation has received relatively little attention. Some studies include crude regional measures of monetary returns in explaining schooling participation, such as the proportion of local employment in industry (Tansel, 2002; Gungor, 2001). We are aware of only two papers that use returns to education in the local labor market to explain participation in education: an unpublished paper by Yamauchi-Kawana (1997) and a recent working paper by Gormly and Swinnerton (2003).

The object of this paper is to ask whether and how much local returns to education, as measured by the Mincer earnings function, influence educational decisions.

The paper is structured as follows: Section 2 outlines a theoretical model. Section 3 describes the estimation approach. The data are discussed in section 4. Section 5 discusses results and Section 6 concludes.

2. Theoretical and estimation issues

This paper is concerned with testing the effect of adult returns to education on the schooling of children and adolescents. The theoretical grounding for this comes from an adaptation of Baland and Robinson (2000). Gormly and Swinnerton (2003) extend the aforementioned model to create a theoretical framework of the influence of returns to education on educational outcomes at the individual level[2] .

In this two-period model, families live together in period 1, and children maintain their own households in period 2. Parents are assumed to be altruistic towards their children, meaning that they derive utility from their children's utility, but children are selfish, precluding any transfers from children to parents in period 2. Hence parents optimize over:

- $u(c_1)$, the household consumption in t=1
- $u(c_2)$, their consumption in t=2
- $dnv(c_c)$, their children's consumption in period 2
- δ, the degree of altruism towards their children

Parents' incomes are fixed at a_i. Consumption can be influenced via three parameters:

- (1-e), the amount of work that parents allocate to their children[3]
- s, the amount of savings in period 1
- b, bequests that the parents may choose to leave their children

Children's future income is determined by:

- e, the amount of education they received
- b, bequests from their parents
- θ, the return to education, which is exogenous

An adult child's wage is equivalent to a human-capital production function, $h(e,\theta)$, which is increasing and concave in its arguments. Gormly and Swinnerton (2003) provide the necessary conditions for closure of the model in more detail.

An important condition of the model is that individuals are not able to borrow to smooth their consumption between time periods, meaning that the household saving rate must be $s \geq 0$. Poor households may thus be liquidity constrained in situations where they would like to borrow to increase period 1 consumption. This yields the household optimization problem[4] :

$$\max_{(s,e,b)}[u(a_1 + n(1-e) - s) + u(a_2 + s - nb) + \delta nv(h(e;\theta) + b)] \qquad (1)$$

$s.t.$ $s \geq 0$; $b \geq 0$

---

[2]Note that the model and notation stem from Gormly and Swinnerton (2003).
[3]Which is the child's productive time not spent in education e.

It is shown that if households are not liquidity constrained, i.e. if households do not need to borrow to increase their consumption in period 1, investments in education are socially optimal and that $h_1(e;\theta)=1$. Also, the relationship between returns to education, θ, and the amount of education acquired is shown to be:

$$\frac{\partial e}{\partial \theta} = -\frac{h_{12}}{h_{11}} \tag{2}$$

For instance, if education and good adult labor market conditions or school quality are complements in the production of higher levels of human capital and therefore of higher wages (which is a plausible), then equation (2) implies that $\frac{\partial e}{\partial \theta}$>0, that is, education increases with improvement in either of the two factors.

The $\frac{\partial e}{\partial \theta}$ term is positive if the liquidity constraint does not bind. However, if the liquidity constraint is binding, Gormly and Swinnerton show that:

$$\frac{\partial e}{\partial \theta} = -\frac{\delta n[h_{12}(e;\theta)v'(c_c)+h_1(e;\theta)h_2(e;\theta)v''(c_c)]}{\nabla} \tag{3}$$

∇<0 is the second order condition for *e* from (1).

In a liquidity constrained environment, two opposing effects influence schooling decisions: if returns to education increase, there is a substitution effect towards education instead of work, driven by relatively higher profitability of schooling vis-à-vis current work by children. However, there is also an income effect at play, due to increased lifetime earnings encouraging increased present consumption. If liquidity constrained parents cannot borrow to increase consumption today and cannot alter their own earnings, a consequence will be a negative effect on schooling of their children: they may choose to let their children work more to benefit in

[4]Note that discounting has been excluded for clarity.

period 1 from increased lifetime incomes associated with the now higher return (more profitable) education.

Hence, in unconstrained households, the relationship between returns to education and schooling participation is expected to be positive. In liquidity-constrained households, this relationship is expected to be smaller in magnitude (or even negative), and the extent to which it will be smaller will depend on the relative sizes of the substitution and income effects.

It is noteworthy that the relationship described above only holds at the household level. At the aggregate economy level, we expect the supply of labor to influence the relationship between educational attainment and educational returns. Duflo (2001) writes down an equation relating the returns to education to the supply of educated labor:

$$b_{jk} = 2\beta_1 S_j + 2\beta_2 \overline{S} + \beta_3 q_{jk} + v_j \tag{4}$$

Here, b_{jk} stands for the return to education of people from cohort k in region j, S_j for the average years of schooling in the individual's region, $\overline{S}$ for the average years of schooling in the country, and q_{jk} for a quality index.

Since an increase in average education is likely to reduce the returns to education, due to supply side effects, we expect that regions with high levels of education could experience lower returns to education due to a relatively higher supply of skilled labor. However, general equilibrium effects may negate such a phenomenon, if the supply of educated labor affects endogenous technical change, and thus affects demand for skilled labor. Foster and Rosenzweig (1996) note the possibility of increased endogenous growth due to a highly educated labor force. Papers by Nelson and Phelps (1966), Schultz (1975) and Gemmell (1996) put forward the view that high levels of education will enhance growth, and this could instigate a positive relationship between supply of educated persons and returns to education. Krueger and Lindahl (2001) and Temple (2001) explain the failure of other studies to find a positive relationship between education and economic growth in cross-country regressions and they find that when measurement error and outliers are taken account of, education does increase growth. Nevertheless, we remain agnostic on the precise relationship between education and growth, and thus on the expected relationship between a region's educational attainment and its returns to education. The possibility remains that in an educational attainment function, the educational return variable will suffer from simultaneity bias, i.e. that it will be jointly determined with educational attainment.

Another problem is that the schooling attainment equation may suffer from omitted variable bias, which (like joint-determination) is another source of endogeneity bias. Both educational returns (henceforth ER) in the local labor market and educational attainment (henceforth EDYRS) may be driven by some third unoberved factor: certain regional characteristics that are unmeasured and thus in the error term of the estimated schooling equation, may be correlated with both ER and with EDYRS. For instance, in regions that are progressive for historical regions, both EDYRS will be high and ER may also be high if such regions attract inward investment which raises the return to education in the local labor market.

In this paper, we attempt to deal with both forms of endogeneity bias, namely simultaneity bias and omitted variable bias.

3. Data description

This paper draws on data from two household surveys conducted by the National Sample Survey Organization of India: the 50^{th} and the 55^{th} round, dating from 1993-1994 and 1999-2000 respectively. They are abbreviated with 1993 and 1999 for convenience. Both rounds have employment and unemployment as their topic (National Sample Survey Organization, 1993, 1999).

Each of the rounds contains information on approximately 100,000 households covering all Indian states and sub-regions[5] . The information contained in these two data-sets overlaps to a large extent. The data include information relating to demographic factors, education, employment and earnings, and household-level information relating to social status, expenditure, principal household activity and related information.

For the measurement of the educational return (ER), three variables are of particular importance and their structure of interest: wages, hours worked and the years of education attained.

Wages are recorded in monetary units for both cash and kind income, and added together to form a total. In the questionnaires, the recall period for waged earnings is one week. Hours worked are inferred from weekly activity reports. Respondents were asked to detail the time spent in different activities over the last week. Responses were recorded in half-day units. From this, we employ a simple transformation to infer the number of hours worked, if the activity reported led to wage earnings:

$$\textit{Hours worked} = \textit{Half days reported} \times 4 \qquad (8)$$

Mincerian earnings functions take years of education as the measure of human capital accumulated. In the NSS samples, however, educational attainment is not recorded by years of education, but rather by level of education completed. Conversion from educational attainment categories to years of education is detailed in Table 1.

[5]The NSSO covers of all 78 state-regions defined by it for the 55^{th} round, however one state-region the Jhelam Valley in Jammu & Kashmir is not covered in 1993 and hence excluded from analysis in 1999.

Clearly, in this context educational attainment only serves as a proxy measure for the years of education completed. It does not take into account any repeats. This, however, is not problematic in the context, as, arguably, the education level completed captures more accurately the level of human capital accumulated than a direct measure of years spent in schooling. This view is directly supported by the human capital hypothesis. A discussion of repeats and other non-regular years of schooling is found in Groot and Oosterbeek (1994).

A second limitation associated with this method of conversion is the fact that high levels of education, such as postgraduate or doctoral studies, cannot be recorded. This implies a potential over-estimation of the returns of education, as high earnings associated with very high levels of education are effectively attributed to lower educational attainment.

Table 2 defines the variables used in estimation. Per capita household expenditure, (*pce*), and wages earned have all been deflated to 1995 prices for comparability, using CPI information from the World Development Indicators (World Bank, 2003). 77 different state-regions are contained in the sample. With regards to religion dummies, Hinduism has been chosen as the base category, due to its high prevalence in India. As to the social group variables, the base category comprises persons not belonging to the 'scheduled caste' or 'scheduled tribe' categories. An overview of summary statistics from the two data-sets is presented in Table 3.

As can be seen from Table 3, both data sets are approximately of the same size, with more than 560,000 individuals. The age distribution and proportion of wage workers is also fairly similar in both samples[6] . Mean levels of education have increased from 1993 to 1999, irrespective of the chosen decomposition of the data set. Wage earners have, on average, more than a year of education greater than those not earning a wage. Approximately one third of the

[6]Wage earners are those individuals for whom a wage is recorded and whose activity status is recorded to be wage employment.

sample resides in urban areas, although this number increases slightly between the two surveys; relatively more wage activity takes place in urban areas.

Women are under-represented among wage earners and attain lower levels of education. Especially at low levels of per capita household expenditure (bottom decile), there is a notable gap in average educational attainment between females and males, though the size of this gender gap has fallen substantially over time: in 1993, the education attainment gap between the genders was about 0.95 years of education, which reduced to 0.58 years by 1999. At high expenditure levels (top decile) though, this gender gap is much smaller, with 0.26 years in 1993 and a reversal to women attaining 0.09 more years of education by 1999[7] . These results are a useful starting point for gender-based results presented in section 4.

Real wages increased by approximately 15% between the two time periods. However, deflated per capita household expenditure decreased slightly between the years, owing to an increase in average household size between the years. Lastly, the demographic composition of the sample is very similar between the two surveys and closely resembles that reported in demographic data on India (CIA, 2004).

4. Estimation and results

Estimation of the influence of Mincerian returns on schooling participation is carried out in a two-stage process. In the first stage, regional rates of returns to education, (ER), are estimated using Mincerian earnings functions. The second stage comprises individual-level estimation, as well as aggregate (state-region-level) estimation, of educational attainment for age, (EDYRS). Key to the

[7]It has been verified that this result is not due to differing age distributions.

analysis is the high degree of heterogeneity in educational attainment in different regions of India[8]

4.1 "First Stage" Earnings Function Estimation

In the first stage, an earnings equation is estimated. The Mincer specification, as outlined previously, is used as follows:

$$w = \alpha + \beta X + \gamma Y + \sum_{i=2}^{i=77} \delta_i sr_i + \sum_{i=2}^{i=77} \eta_i sr_i \times e$$

X is a vector of individual characteristics, Y a vector of social and demographic characteristics, sr_i is a dummy variable for the state-region the *i*th individual lives in and $sr_i \times e$ is an interaction term of the years of education, (e), and the state-region dummy variable, (sr). Table 2 defines the variables used.

The use of state-region dummies, (sr_i), and of the interaction variable between state-region and educational attainment, ($sr_i \times e$), allows calculation of state-regional returns to education:

$$er_1 = \beta_e$$

$$er_i = \beta_e + \beta_{sr_i \times e} \quad \text{for } 2 <= i <= 77$$

Whilst the variation in ER is driven by differences in the slope of the earnings function, as recorded by the sum of β_e and $\beta_{sr_i \times e}$, the inclusion of state-region dummies is also important: it controls for differences in the intercept of the earnings function, i.e. for differences in wage *levels* across state-regions.

[8] In the 1999 sample, the highest average school attainment for people younger than 21 is in Kerala with an average of 6.75 years and the lowest average school attainment is in Sambalpur in Orissa, with an average school attainment of 2.62 years.

Estimation including the state-region variables and their interaction generates a regression function with 166 explanatory variables. The size of the data-sets makes this viable, with about 60,000 wage earners of age 21 and above in each year's sample (see Tables 4 and 5).

A source of concern in earnings function estimation is that of sample selectivity bias: the sample of people earning a wage may not be a random draw from the adult population. Using variables that determine participation of a person in the waged labor force but do not influence the conditional level of wages, a selection equation is estimated and its results used to correct the estimation of the earnings function (Heckman, 1979). The binary selection variable 'wage earner' (or *we*) takes value 0 if an individual is not earning a wage and value 1 if she is earning a wage.

The credibility of the Heckman procedure depends on the extent to which good identifying variables are available that can be excluded from the wage equation but affect selection into waged work. The data-sets yield three variables that may explain participation in the waged labor force, but not affect wages conditional on being in the labor force: LAND-OWNER, NUM-65 and CHILD-10 are used to control selection. Household demographic characteristics, such as the number of elderly aged 65 and above (NUM-65), and number of child dependants (CHILD-10), are likely to play a role in individuals' choice about labor force participation and type of employment undertaken. For instance, in households with a large number of dependants, working-age adults are more likely to seek and accept flexible forms of work, such as self-employment or informal or casual employment rather than wage work. Similarly, land ownership, (LAND-OWNER), is likely to be associated with the likelihood of working-age adults seeking wage employment: in households that own land, adults are more likely to pursue self-employment. Hence, the first stage selection equation contains all wage equation variables (except hours worked) and the three exclusion restrictions outlined above.

The sample of earners in the wage equation is limited to ages 21 and above. This precludes overlap between the observations included in earnings function calculation and those included in educational attainment functions to be estimated in the second stage.

Detailed estimation results are presented in Tables 4 and 5. Both estimations are adjusted for cluster effects at the village level and use heteroscedasticity-robust estimators, as this proved an issue in preliminary estimations. Results for the robust estimators can be considered efficient due to the large sample sizes in both time periods.

The variables used in the first stage probit for identifying the selectivity term, λ, are LAND-OWNER, NUM-65 and CHILD-10. They are valid exclusion restrictions, as they show strong association with selection into waged work and are theoretically justified above. λ is significant at the 1% level in the earnings functions for both years in Tables 4 and 5.

In the first stage probit of wage work participation, all coefficients exhibit the expected signs except for those the social indicators of schedule caste and schedule tribe. A possible explanation is that members of scheduled castes and tribes are less likely to have capital to start self-employment, thus explaining the higher likelihood of low caste members to be wage earners.

However, an inspection of the coefficients of the earnings functions shows that there is little difference between selectivity-bias corrected and OLS estimates. This fact is also confirmed by results in table 7, which shows that, for each time period, the two competing specifications show very little difference in mean and extreme values of ER. Consequently, we choose OLS results for further analysis.

Earnings function results presented in Tables 4 and 5 omit the coefficients for the 77 state-region variables and the 77 interaction variables for space reasons. The coefficient of ER reported here is that for state-region 21, the dry areas of Gujarat (the base category). Its value is not representative for mean returns to education in India. R^2 values of the OLS earnings functions are reassuring, with values of 0.54 for the 1993 data-set and 0.67 for the 1999 data-set. Also, except for the Buddhist religion dummy, coefficients exhibit highly significant t-values.

In the OLS earnings functions, all variables exhibit expected signs. The age-earnings relationship derived from AGE and AGESQ predicts earnings to peak at the age of 50 in 1993 and the age of 52 in 1999, ceteris paribus. This conforms to human capital theories of increased productivity due to experience being offset by age-driven productivity losses later in life. Female wage disadvantage stands at around 30%, but decreases between the two time periods. Marital status and urban location show strong association with wages earned, again conforming with expected magnitudes and directions. Lastly, the data suggests that caste discrimination in waged work is still an issue, although wage losses associated with belonging to a scheduled caste or tribe decrease considerably between the years.

4.2 "Second Stage" Estimation of Educational Attainment

In the second stage of the estimation process, the effect of educational returns on schooling participation are estimated on two levels: firstly, individual schooling attainment functions are estimated and secondly, state-region level average schooling attainment functions are estimated, to aggregate results at the level of the regional labor markets.

<u>Individual-level analysis</u>

The first and most intuitive way is to estimate educational attainment functions at the individual level. For this, the sample is limited to ages 5 to 20, driven by the regular school enrolment age in India being 5, and to preclude overlap between the individuals included in the estimation of wage functions and those included in attainment functions. Since many persons aged 5-20 will still be in school, we include dummy variables for each age from 6 to 20; this means that effectively, we are modeling years of schooling for age, as in Case and Deaton (1999), who also examine the determinants of educational attainment.

Table 6 presents the educational attainment functions. The equations are estimated for ages 5 to 20, jointly for both genders and all household per capita expenditure levels[9]. Estimates at the individual level are conducted using a heteroscedasticity robust and cluster-robust estimator.

The first two columns contain the individual-level educational attainment functions. The main variable of interest is the return to education in the state-region, which is derived from the OLS earnings equations in tables 4 and 5.

Columns 1 and 2 present estimates of educational attainment in 1999 using OLS and instrumental variable approaches. Regional returns to education in 1993 serve as an instrument for regional returns to education in 1999. For the 1993 *ER* to be a valid instrument for the 1999 *ER*, it must be the case that 1993 local *ER* is correlated to the 1999 local *ER* (which is the case) and it must not be in the error term of the schooling equation in 1999[10].

EDYRS increases with age, as expected, except for cohorts of ages 18 and 20. Actual *EDYRS* as a proportion of possible *EDYRS* are expected to decrease with age[11], due to dropping-out of school at higher ages. The discrete nature of the *EDYRS* variable leads to slight deviations

---

[9] Educational attainment and household per capita expenditure (*pce*) are very likely to be co-determined. If a child drops out of school and earns a wage, this will impact directly on household expenditure. Instrumentation would be the appropriate control strategy here but the data does not yield an instrument of acceptable quality. Thus, it is noted that the coefficient of *pce* will be biased downward. This does not impact our analysis in a central way though, as the focus here is on the effect of schooling returns on educational attainment, rather than the effect of household expenditure per capita on educational attainment, and because *pce* and ER are unlikely to be highly correlated.

[10] See Footnote 12.

from the expected pattern, as, for example, between ages 12 and 13. The use of age dummies normalizes EDYRS estimation, making the estimation result independent of age.

The individual level relationship between *ER* and *EDYRS* is positive and significant, using an OLS estimator. The size of the coefficient implies that if the return to education in the local labor market increases from one SD below to one SD above the mean return to education across state-regions, years of education acquired increases by approximately 0.2 years, though as we will see in Table 8, the size of effect of *ER* on *EDYRS* is much greater for certain population groups than others. However, using an instrumental variable approach, the relationship becomes smaller and is only significant at the 10% level, due to the larger standard error. The point estimates of the returns to education variable in the OLS and IV columns are not significantly different, though. A Wald test shows that the null that the coefficients on *ER* in the OLS and IV columns are equal could not be rejected at the 5% level.

<u>Aggregate-level analysis</u>

Individual educational attainment functions discussed above are not able to capture aggregate outcomes. Whilst we expect to find a positive relationship between *ER* and *EDYRS* at the individual level (in households that are not liquidity-constrained), at the aggregate level, the relationship may be weaker or negative, owing to supply effects: high levels of educational attainment in a state-region may signal an imminent influx of skilled labor into the regional labor market, leading to lower returns to education. This would cause *EDYRS* and *ER* to be simultaneously determined.

---

[11]An index of actual to potential attainment, (*ae*), holding other variables constant, can be calculated with a simple formula: $ae = \frac{\beta_{AGE_i}}{AGE_i - 5}$.

The approach used to control for this is instrumental variables. For a variable to be a valid instrument, it must be highly correlated with the variable it instruments for, and must not be correlated with the error term of the equation of main interest. In the case of the variable ER_{1999}, its lagged value ER_{1993} fulfils both criteria: the variables are well correlated, and, by definition, ER_{1993} will not be correlated with time-variant effects that occur between the years[12], though we cannot adequately control for time-invariant relationship between *ER* in 1999 and 1993. For that, we have used state-region fixed effects, exploiting the panel aspect of our data.

Data is aggregated separately for each year at the state-region level, the unit of the regional labor market chosen for the estimation of the earnings function. With 77 state-regions in India, an average state-region contained about 13 million inhabitants in 1999. Thus, state-regions are of sufficient size to represent local labor markets, whereby migration takes place within the state-region, but is less likely to take place across state-regions.

Variable values are the means for each state-region and each year. This yields:

$$\bar{e}_j = \alpha + \beta\bar{X}_j + \gamma er_j + \delta t + \varepsilon_j$$

Here, the average level of education in the j^{th} region, ($\bar{e}_j$), depends on a vector of averaged personal and demographic characteristics, ($\bar{X}_j$), the return to education in the region, (er_j), and a time dummy variable, (t), to control for increases in schooling participation between the two years. In comparison to individual level attainment estimation, the vector of variables used in estimation was reduced firstly to preserve degrees of freedom, owing to the relatively low number of observations (only 77 per year), and secondly due to variables failing to add explanatory power

[12] ER_{1993} will not be correlated with shocks that occur after 1993 and which may affect both ER_{1999} and $EDYRS_{1999}$.

to the estimation. Thus, the variables included in estimation are the state-regional means of age, (*age*), of education level of the household heads, (*hh*), of household per capita expenditure, (*pce*), and of the dummy variable for urban location, (*ur*), capturing the share of urban population in a state-region.

Columns 3-5 of table 6 present results at the state-region level. As expected, the relationship between returns to education and educational attainment at the state-region level is smaller than in the individual attainment function; in fact, it turns out to be of negative sign, but estimates suffer from low levels of significance.

The state-region fixed-effects estimator in column 4 controls for unobserved factors, using the panel aspect of the two data-sets. The point estimate for the returns to education variable is very similar to the OLS estimate. When we estimate the state-region level educational attainment equations separately for males and females and for poor and non-poor samples (not reported), the point estimates of OLS and fixed effects estimators do not differ significantly either. Since the fixed effects estimator is a powerful control for endogeneity and its introduction does not alter the OLS coefficient on *ER*, we can reject the idea that unobserved heterogeneity across state-regions is affecting results.

Analysis at the state-region level changes the sign of the *ER* - *EDYRS* relationship, however the estimate is not statistically significant. The fact that a positive coefficient on *ER* is present in all individual-level experiments conducted suggests that simultaneity *does* affect results at the aggregate level: it seems that the higher the supply of educated workers in a region, the lower the returns to education in that region, and that this negative supply-side factor undermines our ability to pick up any positive effect that returns to education may otherwise have on educational attainment. When this is addressed using an IV procedure, in the final column of

Table 6, the *ER* – *EDYRS* relationship turns positive and is of approximately the same size as in the IV column of the individual-level results in Table 6.

In summary, the results show evidence of a small positive influence of returns to education on educational attainment at the individual level. At the aggregate level, these results are much weaker. This may be attributable to negative supply-side effects, or caused by low power of estimation due to the small number of observations at the aggregated level.

5. The Effects of Liquidity Constraints in Educational Attainment

In the Indian context, liquidity constraints may affect male and female schooling decisions differently. Thus, analysis of the effects of *ER* on *EDYRS* by gender and for poor and non-poor households yields more detailed insight into the role of liquidity constraints and gender bias in educational attainment.

We repeat the experiments presented in Table 6, subdivided by gender and for different quantiles of household per capita expenditure (*pce*)[13]. Female educational attainment functions are based on female returns to education and male attainment functions on male returns, since female returns to education are more likely to be important for girls' schooling decisions, and male returns more important for boys. Detailed results are presented in Table 8.

The variable set used is identical to that in Table 6 but we omit the presentation of age, religion, social status and location dummies for brevity. Estimation uses the instrumental variable approach. Controls for heteroscedasticity and cluster effects are introduced, too, as before.

Gender analysis yields three striking insights in table 8: firstly, male schooling participation in the poorest households (the bottom 10 deciles) is affected by the income effect

[13] It should not be presumed that our per capita expenditure categories are exogenous. However, there is no clear way – with cross-section data – of addressing the potential endogeneity of *pce* category.

predicted by the model in section 2, as the negative significant coefficient on *ER* shows. To test this result for robustness, a probit equation of enrolment was estimated for children who cannot be affected by the income effect: school enrolment of children of age 5 and 6 in households where no household member has received any education should only be affected by the substitution effect. Neither the child nor any other household member will be subject to an income effect if *ER* increases. The large size of the sample allows the estimation of such an enrolment probit[14] . The coefficient of *ER* suggests that there is a positive association between returns to education and enrolment in this subgroup, thus reconfirming the hypothesis that in liquidity constrained households, the effect of *ER* on male schooling participation is affected by the income effect. This result is also found in Gormly and Swinnerton (2003) and Edmonds (2004).

Secondly, for females of the same income group, the relationship is equally large, but positive in sign. This suggests that in India, male children with some education have better possibilities of earning waged income than otherwise equivalent female children, i.e. the male opportunity cost of education is higher. For young females, this opportunity cost is smaller or absent, as their choices are more between domestic work and going to school. There is some support in the data for the notion that girls are less likely than boys to do market work in India: in the 5-20 age group, 7.5% of boys but only less than 3% of girls are in waged work[15]. Hence, for girls, the positive substitution effect of higher *ER* dominates any negative income effect and they exhibit a positive overall relationship between *EDYRS* and *ER*.

Thirdly, the data suggest that the monetary cost of education poses a barrier to education for both boys and girls in very poor households. For example, for girls, the size of the *ER* coefficient increases significantly between the bottom decile and the 10-25th quantile. A Wald test

[14] Results of this estimation are presented in table 8.
[15] This is not incompatible with the existence of pro-male bias in education in India. Kingdon (2005) finds evidence

on the null hypothesis $H_0: \beta_{ER_{0-10}} = \beta_{ER_{10-25}}$ is significant at the 6% level, suggesting that schooling participation responds to *ER* more in quantiles 10th to 25th than in the bottom decile. Thus, the data suggest that monetary costs do pose a barrier to female schooling participation in the poorest households. For males too the relationship becomes positive at higher income groups, implying a stronger effect of the opportunity cost of education at low levels of household income.

It is noteworthy that for the gender groups individually, the absolute effect of *ER* becomes sizable: in the female sub-sample "10-25": if *ER* increases from 1 SD below to 1 SD above mean *ER*, education attainment increases by 1 whole year. Given that mean educational attainment of girls in this *pce* group is 2.7 years (Table 8), a 1-year increase is a very substantial increase in years of schooling.

To summarize, results in Table 8 show that for the poorer parts of the population, returns to education play a more major part in educational decisions than for the richer part. Female educational decisions respond in the way theory predicts, with changes in the size of coefficients suggesting that the cash cost of education may act as a barrier to education for the females in the poorest households: female *EDYRS* responds less to labor market incentives in the bottom decile than in the 10th to 25th quantile. Poor male children's educational decisions exhibit a negative relationship with *ER* suggesting that boys have a higher opportunity cost of education, which plays out particularly in liquidity-constrained households. In areas where ER is higher, boys in poor households are withdrawn from school to take advantage of the higher return to their (existing) levels of schooling. In other words, the (negative) income effect of ER is greater for boys than for girls.

---

of significant gender bias in household education expenditure allocations in India.

6. Conclusion

We find that the Mincerian return to education for adults in the local labor market influences schooling decisions of young people in India. The results are robust to omitted variable bias but are affected by simultaneity, especially at the aggregate level of the state-region.

At the individual level, we find strong relationships between monetary returns and educational decisions. For females, the relationship is positive and mostly highly statistically significant though the cost of attending school still acts as a barrier to schooling for poor females. The data suggest that for poor males, education has a significant opportunity cost, causing the relationship between educational returns and schooling participation to become negative.

These results suggest that schooling decisions are influenced not only by household income and taste for education, and by availability and quality of schools, but also by the prevailing economic returns to education in the local labor market. However, labor market policies to raise the returns to education to encourage schooling participation could lead to unintended effects: poor males may acquire less education than otherwise, due to the negative income effect prevailing in a liquidity constrained situation. Thus, in order for labor market incentives to work in the intended direction, policies aimed at raising the monetary returns to education must be complemented by policies to alleviate liquidity constraints and to reduce opportunity costs of schooling for poor households, such as a policy of school-attendance-contingent cash subsidies.

The results here offer a preliminary insight into the role of economic returns in schooling decisions. Our understanding would benefit from further analysis of smaller geographical subunits than the state-region, allowing for alternative "labor market boundaries" and from more explicit modeling and detection of liquidity constraints. This suggests promising avenues for future research in this area of economics of education.

References

Angrist, J. D. and Krueger, A. B. (1992). Estimating the payoff to schooling using the vietnam-era draft lottery. *NBER Working Paper*, (W4067).

Angrist, J. D. and Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 114(2):533–575.

Baland, J.-M. and Robinson, J. A. (2000). Is child labor inefficient? *Journal of Political Economy*, 108(4):663–679.

Banerjee, B. and Knight, J. B. (1985). Caste discrimination in the Indian urban labor market. *Journal of Development Economics*, 16(3):293–311.

Barro, R. J. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, 106(2):407–43.

Becker, G. S. (1964). *Human Capital: a theoretical and empirical analysis, with special reference to education*. Number 80. National Bureau of Economic Research.

Benhabib, J. and Spiegel, M. M. (1994). he role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34(2):143–74.

Bennell, P. (1996). Rates of return to education: Does the conventional pattern prevail in sub-saharan africa? *World Development*, 24(1):183–199.

Blundell, R., Dearden, L., and Sianesi, B. (2001). Estimating returns to education: Models, methods and results. *Centre for Economics of Education, LSE, Discussion Paper*, (16).

Card, D. (1999). *Handbook of Labor Economics*, volume 3, chapter The Causal Effect of Education on Earnings. Elsevier, Amsterdam.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160.

Case, A. and Deaton, A. (1999). School inputs and educational outcomes in south africa. *Quarterly Journal of Economics*, 114(3):1047–84.

CIA (2004). World Factbook. Electronic Source.

Deaton, A. (1997). *The Analysis of Household Surveys*. John Hopkins University Press for the World Bank, Baltimore and London.

Deininger, K. (2003). Does cost of schooling affect enrolment by the poor? Universal primary education in Uganda. *Economics of Education Review*, 22(3):291–305.

Drèze, J. and Kingdon, G. G. (2001). School participation in rural India. *Review of Development Economics*, 5(1):1–24.

Drèze, J. and Murthi, M. (2001). Fertility, education and development: Evidence from India. *Population and Development Review*, 27(1):33–63.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Edmonds, E. V. (2004). Does illiquidity alter child labor and schooling decisions? Evidence from household responses to anticipated cash transfers in South Africa. NBER Working Paper 10265, January.

Foster, A. D. and Rosenzweig, M. R. (1996). Technical change and human-capital returns and investments: Evidence from the green revolution. *American Economic Review*, 86(4):931–953.

Garen, J. (1984). Returns to schooling: A selectivity bias approach with a continuous variable. *Econometrica*, 52(5):1199–1218.

Gemmell, N. (1996). Evaluating the impacts of human capital stocks and accumulation on economic growth: Some new evidence. *Oxford Bulletin of Economics and Statistics*, 58(1):9–28.

Glewwe, P. (1999). *The Economics of School Quality Investments in Developing Countries*. St. Martin's Press, New York.

Gormly, S. and Swinnerton, K. A. (2003). The effect of adult returns to schooling on children's school enrolment, theory and evidence from south africa. U.S. Department of Labor.

Greene, W. H. (2003). *Econometric Analysis*. Prentice Hall International, London, 5th edition.

Groot, W. and Oosterbeek, H. (1994). Earnings effects of different components of schooling; human capital versus screening. *The Review of Economics and Statistics*, pages 317–321.

Gungor, A. D. (2001). The gender gap in turkish education: A regional approach. *Middle-East Technical University Studies in Development*, 28(3-4):339–360.

Hanushek, E. (1995). Interpreting recent research on schooling in developing countries. *World Bank Research Observer*, 10(2):227–246.

Harbison, R. and Hanushek, E. (1992). *Educational Performance of the Poor: Lessons from Rural Northeast Brazil*. Oxford University Press.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Johnston, J. and DiNardo, J. (1997). *Econometric Methods*. McGraw-Hill International Editions, 4th edition.

Kingdon, G. G. (2005). Where has all the bias gone? Detecting gender bias in the household allocation of educational expenditure in India. *Economic Development and Cultural Change,* 53(2).

Kremer, M. and Chen, D. L. (2002). Income distribution dynamics with endogenous fertility. *Journal of Economic Growth*, 7(3):227–58.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114(2):497–532.

Krueger, A. B. and Lindahl, M. (2001). Education for growth: Why and for whom? *Journal of Economic Literature*, 39(4):1101–36.

Krueger, A. B. (2003) Economic considerations and class-size, *Economic Journal*, 113, F34-F63.

Lockheed, M. and Verspoor, A. (1991). *Improving Primary Education in Developing Countries*. New York.

Miguel, E. and Kremer, M. (2004). Worms: Education and health externalities in kenya. *Econometrica*, 72(1):159–217.

Mincer, J. (1974). *Schooling Experience and Earnings*. Columbia University Press, New York.

Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics*, 32(3):385–97.

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics,* 72(2):334–38.

National Sample Survey Organization (1993). Socio-economic household survey, 50^{th} round: Employment and unemployment. Survey.

National Sample Survey Organization (1999). Socio-economic household survey, 55^{th} round: Employment and unemployment. Survey.

Nelson, R. R. and Phelps, E. S. (1966). Investments in humans, technological diffusion and economic growth. *American Economic Review*, 56(1):69–75.

Pritchett, L. (2001). Where has all the education gone? *World Bank Economic Review*, 15(3):367–391.

Psacharopoulos, G. (1973). *Returns to Education: an International Comparison*. Jossy Bass-Elsevier, San Francisco.

Psacharopoulos, G. (1981). Returns to education: an updated international comparison. *Comparative Education*, 17:321–41.

Psacharopoulos, G. (1985). Returns to education: a further international update and implications. *Journal of Human Resources*, 20(4):583–604.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, 22(9):1325–1343.

Psacharopoulos, G. and G. Patrinos (2004). Returns to investment in education: A further update, *Education Economics*, 12(2), pp. 111-134.

Rouse, C. E. (1999). Further estimates of the economic return to schooling from a new sample of twins. *Economics of Education Review*, 18(2):149–157.

Schultz, P. T. (2004). School subsidies for the poor: Evaluating the Mexican Progresa poverty program. *Journal of Development Economics*.

Schultz, T. W. (1975). The value of the ability to deal with equilibria. *Journal of Economic Literature*, 13(1):827–46.

Tansel, A. (2002). Determinants of school attainment of boys and girls in turkey: Individual, household and community factors. *Economics of Education Review*, 21(5):455–470.

Temple, J. R. W. (2001). Generalizations that aren't? Evidence on education and growth. *European Economic Review*, 45(4-6):905–18.

World Bank (2003). World development indicators. Electronic source, World Bank.

Yamauchi-Kawana, F. (1997). Information, neighborhood effects and investment in human capital: Learning schooling returns in a dynamic context. *mimeo, University of Pennsylvania*.

Table 1: Transformation of education coding to years of education

Educational attainment code	Imputed years of education
Not literate	0
Literate through attending NFEC/AEC, TLC or others	1
Literate, but below primary	3
Primary	5
Middle	8
Secondary	10
Higher secondary	12
Graduate and above	15

Note: NFEC = Non Formal Education Centre, TLC = Total Literacy Campaign, AEC = Alternative Education Centre

Table 2: Variables used in estimation

Variable Name	Abbreviation	Definition
Personal Variables:		
AGE	a	Age of individual in years
AGESQ	a^2	Square of AGE
EDYRS	e	Number of years of education, as defined in table 1
LN-WAGES	w	ln(Weekly total wage)
HOURS	hr	Hours worked, as defined in (8)
AGE_i	age_i	Dummy variable for age i
FEMALE	f	Gender dummy: male=0, female=1
MARRIED	m	Marital status dummy: never married=0; married, divorced, widowed=1
Demographic Variables:		
HH-EDUC	he	EDYRS of the designated head of household
HH-EXP	pce	Household per capita expenditure over the last month
CHILD-10	$ch10$	Number of children aged 10 or younger in the household
NUM-65	$num65$	Number of individuals aged 65 or older in the household
LAND-OWN	lo	Dummy: household owns land=1, does not own land=0
SR_i	sr_i	Regional dummy: state-region
$\mathrm{SR}_i{}'e$	$sr_i{}'e$	State-region and EDYRS interaction variable
URBAN	ur	Location dummy: rural=0, urban=1
REL-*	rel_i	Religion dummies: Muslim, Christian, Sikh, Jainist, Buddhist Hinduism omitted as base category
SCH-TRIBE	st	Scheduled tribe dummy
SCH-CASTE	sc	Scheduled caste dummy
Calculated Variable:		
ER	er	Local rate of return to education in the state-region (education coefficient as calculated in (6)

Table 3: Summary Statistics for NSS Datasets

Variable	NSS 1993	(s.d.)	NSS 1999	(s.d.)
Size of dataset				
Individuals in data-set	564,695		588,525	
Individuals aged 21 or older	294,616		313,486	
Individuals aged 5 to 20	203,345		214,498	
Wage earners aged 21 or older	59,421		68,629	
Mean education levels (in years)				
Whole sample	3.880	(4.359)	4.358	(4.549)
Age 21 and above	4.428	(4.898)	5.121	(5.132)
Wage earners age 21 and above	5.824	(5.495)	6.345	(5.477)
Age 21 and above not earning a wage	4.067	(4.663)	4.702	(4.940)
Age 5 to 20	4.105	(3.425)	4.345	(3.421)
Female age 5 to 20, bottom 10%*	1.609	(2.396)	2.142	(2.587)
Males age 5 to 20, bottom 10%*	2.556	(2.799)	2.726	(2.781)
Female age 5 to 20, top 10%*	6.086	(3.753)	6.784	(3.737)
Males age 5 to 20, top 10%*	6.344	(3.624)	6.718	(3.611)
Demographic composition				
Share living in urban areas	35.9%		38.1%	
Share of females in the sample	47.2%		47.4%	
Female share of wage earners	22.8%		22.5%	
Urban share of wage earners	46.9%		48.8%	
Share of Hindus in sample	78.1%		77.4%	
Share of Muslims in sample	11.2%		12.6%	
Share of Christians in sample	6.0%		5.1%	
Share of Sikhs in sample	2.3%		2.5%	
Share of Jains in sample	0.3%		0.4%	
Share of scheduled tribe members in sample	11.1%		11.4%	
Share of scheduled caste members in sample	14.8%		16.2%	
Economic variables				
Per capita monthly household expenditure	457.6	(529.2)	438.9	(358.1)
Average weekly wage earned (1995 prices)	348.5	(412.5)	400.8	(891.3)
Returns to education for wage earners (aged 21 or older)	7.81%	(1.90%)	8.34%	(1.46%)

Note: Per capita household expenditure and average weekly wages have been deflated to 1995 prices.
* Top and bottom 10% rank in distribution of household expenditure per capita

Table 4: Earnings function: 1993 Sample

	OLS		Heckman Correction			
	Earnings fn.		Earnings fn.		Selection fn.	
Variable	*Coeff.*	*s.e.*	*Coeff.*	*s.e.*	*Coeff.*	*s.e.*
EDYRS	0.0824	0.0037	0.0822	0.0029	-0.0009	0.0030
HOURS	0.0324	0.0004	0.0324	0.0002		
AGE	0.0555	0.0021	0.0477	0.0023	0.0867	0.0015
AGESQ	-0.0006	0.00002	-0.0005	0.0000	-0.0011	0.0000
FEMALE	-0.3157	0.0081	-0.2445	0.0151	-0.8333	0.0058
URBAN	0.2226	0.0123	0.1971	0.0082	0.1912	0.0064
MARRIED	0.1666	0.0123	0.1594	0.0104	0.1431	0.0098
REL-MUSL	-0.0231	0.0168	-0.0176	0.0109	-0.0387	0.0096
REL-CHRIST	0.0190	0.0235	0.0107	0.0152	0.1166	0.0151
REL-SIKH	0.0615	0.0338	0.0741	0.0272	-0.1314	0.0244
REL-JAIN	0.0650	0.0763	0.1197	0.0571	-0.6523	0.0449
REL-BUDDH	-0.0191	0.0320	-0.0427	0.0284	0.3268	0.0286
SCH-TRIBE	-0.0385	0.0201	-0.0571	0.0125	0.2323	0.0111
SCH-CASTE	-0.0447	0.0131	-0.0829	0.0106	0.4846	0.0075
LAND-OWNER					-0.3776	0.0076
NUM-65					-0.0926	0.0045
CHILD-10					-0.0632	0.0019
Intercept	1.9348	0.0478	2.214	0.0648	-1.7454	0.0338
λ			-0.114	0.0213		
N	73753		73753		358276	
R^2	0.5421					

Table 5: Earnings function: 1999 Sample

	OLS		Heckman Correction			
	Earnings fn.		Earnings fn.		Selection fn.	
Variable	*Coeff.*	*s.e.*	*Coeff.*	*s.e.*	*Coeff.*	*s.e.*
EDYRS	0.0774	0.0042	0.0776	0.0022	-0.0057	0.0029
HOURS	0.0352	0.0002	0.0352	0.0001		
AGE	0.0645	0.0016	0.0578	0.0018	0.098	0.0015
AGESQ	-0.0006	0.0000	-0.0001	0.0000	-0.001	0.0000
FEMALE	-0.2873	0.0074	-0.225	0.0124	-0.940	0.0057
URBAN	0.2261	0.0086	0.210	0.0057	0.104	0.0065
MARRIED	0.1724	0.0100	0.1675	0.0074	0.116	0.0095
REL-MUSL	-0.0152	0.0113	-0.008	0.0078	-0.084	0.0093
REL-CHRIST	0.0392	0.0167	0.0347	0.0114	0.058	0.0153
REL-SIKH	0.0804	0.0266	0.0952	0.0189	-0.183	0.0227
REL-JAIN	0.1417	0.0693	0.183	0.0432	-0.655	0.0446
REL-BUDDH	0.0127	0.0331	0.0173	0.0210	-0.074	0.0280
SCH-TRIBE	-0.0051	0.0117	-0.0184	0.0088	0.223	0.0106
SCH-CASTE	-0.0127	0.0080	-0.0404	0.0076	0.449	0.0070
LAND-OWNER					-0.3591	0.0069
NUM-65					-0.0808	0.0040
CHILD-10					-0.0457	0.0016
Intercept	1.7194	0.0386	1.939	0.0499	-1.7517	0.0342
λ			-0.0920	0.0167		
N	86251		86251		338129	
R^2	0.6707					

Table 6: Educational Attainment Functions: Full Sample Results

	Individual-level results		*State-region-level results*		
	OLS	IV	OLS	FE	IV
ER	4.7100***	2.6764*	-2.9398	-3.1700	2.1742
	(5.36)	(1.88)	(1.48)	(1.27)	(0.31)
FEMALE	-0.4255***	-0.4252***	4.2873**	-2.7361	1.9540
	(31.60)	(31.59)	(2.61)	(1.43)	(0.60)
URBAN	0.5473***	0.5486***	-0.0189	0.1850	0.0688
	(18.62)	(18.64)	(0.06)	(0.32)	(0.14)
HH-EXP	0.0010***	0.0010***	-0.0000	-0.0005	0.0005
	(8.07)	(8.03)	(0.08)	(0.75)	(0.62)
HH-EDUC	0.1564***	0.1565***	0.4204***	0.3954***	0.2655***
	(51.71)	(51.80)	(10.08)	(11.31)	(3.22)
SCH-TRIBE	-0.6510***	-0.6352***			
	(14.54)	(13.81)			
SCH-CASTE	-0.5709***	-0.5726***			
	(20.44)	(20.43)			
AGE6	0.7137***	0.7145***			
	(37.53)	(37.58)			
AGE7	1.1739***	1.1748***			
	(59.77)	(59.78)			
AGE8	1.3662***	1.3672***			
	(74.45)	(74.48)			
AGE9	1.5601***	1.5618***			
	(74.77)	(74.88)			
AGE10	1.9197***	1.9210***			
	(96.64)	(96.63)			
AGE11	2.4787***	2.4804***			
	(103.25)	(103.26)			
AGE12	2.7739***	2.7754***			
	(118.14)	(118.06)			
AGE13	3.6040***	3.6050***			
	(122.92)	(122.96)			
AGE14	4.0823***	4.0841***			
	(130.74)	(130.65)			
AGE15	4.4084***	4.4092***			
	(127.92)	(127.92)			
AGE16	4.9538***	4.9550***			
	(132.12)	(132.10)			
AGE17	5.6633***	5.6641***			
	(134.68)	(134.68)			
AGE18	5.1922***	5.1936***			
	(129.26)	(129.25)			
AGE19	5.8428***	5.8439***			
	(117.44)	(117.51)			
AGE20	4.9675***	4.9687***			
	(110.05)	(110.07)			
AGE			1.0243***	0.3460***	1.1805***
			(8.24)	(3.50)	(9.05)
Constant	-0.1335	0.0368	-12.3079***	-0.2515	-12.8935***
	(1.55)	(0.28)	(8.47)	(0.15)	(7.69)
N	217834	217834	154	154	77
R^2	0.44	0.44	0.81	0.67	0.82

Robust t-stats in parentheses; * significant at 10%;** significant at 5%; *** significant at 1% levels
Note: OLS Individual-level regression uses 1999 data. In IV equations, ER_{1999} is instrumented with ER_{1993}.
Aggregate regressions use data from both time periods.

Table 7: Summary of the Returns to Education coefficient under different specifications

Data-Set	*Specification*	*Mean*	*Minimum*	*Maximum*
1993	Heckman	7.65%	2.68%	11.44%
1993	OLS	7.81%	2.82%	11.49%
1999	Heckman	8.34%	4.57%	12.10%
1999	OLS	8.34%	5.18%	11.93%

Table 8: Individual Education Attainment Functions: Gender Sub-Sample Liquidity Constraint Assessment

				Female Sub-Sample			
Row title: *pce* quantile	All	0-10	10-25	25-50	50-75	75-100	Enrolment
ER	11.358***	15.319**	30.663***	19.756***	15.975***	1.798	38.008**
	(4.92)	(2.57)	(5.88)	(3.96)	(3.79)	(0.57)	(2.29)
HH-EXP	0.001***	-0.0001	0.006***	0.007***	0.003***	0.0002***	0.002
	(5.69)	(0.11)	(2.51)	(7.54)	(6.14)	(3.51)	(0.47)
HH-EDUC	0.175***	0.162***	0.145***	0.162***	0.159***	0.142***	
	(41.00)	(17.25)	(18.54)	(30.30)	(32.04)	(31.19)	
N	102556	11976	12866	25327	26004	26383	431
R^2	0.40	0.15	0.11	0.25	0.36	0.54	0.02
Mean *EDYRS*	4.225	2.142	2.672	3.422	4.480	6.021	
				Male Sub-Sample			
Row title: *pce* quantile	All	0-10	10-25	25-50	50-75	75-100	Enrolment
ER	-3.875***	-16.930***	-6.139	3.163	7.140***	3.555*	13.784
	(2.64)	(3.69)	(1.45)	(1.14)	(3.09)	(1.82)	(0.96)
HH-EXP	0.001***	0.001	0.009***	0.004***	0.003***	0.0002***	0.001
	(9.12)	(0.95)	(3.98)	(4.43)	(6.94)	(4.52)	(0.23)
HH-EDUC	0.142***	0.128***	0.140***	0.136***	0.132***	0.122***	
	(47.93)	(14.22)	(19.31)	(28.15)	(31.19)	(30.94)	
N	115278	12222	13892	28279	29713	31172	457
R^2	0.47	0.23	0.29	0.35	0.45	0.60	0.03
Mean *EDYRS*	4.549	2.740	3.348	3.983	4.828	6.048	

Robust t-statistics in parentheses; *, **, and *** signify statistical significance at 10%, 5%, and 1% respectively. Note: 0-10 contains observations from households in the lowest decile of per capita expenditure (*pce)*, 10-25 from *pce* quantiles 10th to 25th, and so on

The "Enrolment" column shows coefficients from an IV probit estimation of enrolment on a sample of children aged 5 and 6 years old in households in the bottom decile of *pce* and where nobody has received any education. The value of R^2 in the Enrolment column is that of the pseudo R^2 measure.