Working Paper 150

Young Lives

Accessing Higher Education in Developing Countries:

Panel Data Analysis from India, Peru, and Vietnam

Alan Sánchez and Abhijeet Singh

www.younglives.org.uk

MAY 2016

Accessing Higher Education in Developing Countries: Panel Data Analysis from India, Peru, and Vietnam

Alan Sánchez and Abhijeet Singh

ISBN 978-1-909403-67-3

A catalogue record for this publication is available from the British Library.

Funded by

UKaid
from the British people

Irish Aid
An Roinn Gnóthaí Eachtracha agus Trádála
Department of Foreign Affairs and Trade

Young Lives, Oxford Department of International Development (ODID), University of Oxford, Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK

Tel: +44 (0)1865 281751 • E-mail: younglives@younglives.org.uk

Contents

The authors

Alan Sánchez is a Research Associate at *Grupo de Análisis para el Desarrollo* (GRADE) in Lima, Peru and the Principal Investigator of Young Lives Peru. He is also a part-time lecturer at Universidad de Piura in Lima, Peru. His research has focused on the study of the short-term and long-term consequences of early-life shocks; skills formation; and, impact evaluation of social programmes. Before joining GRADE, he worked at the Economic Research Division of the Banco Central de Reserva del Peru (Central Bank of Peru).

Abhijeet Singh is a Post-doctoral researcher at UCL Department of Economics and was a Quantitative Research Officer with Young Lives. He is an applied microeconomist working on issues in development economics with a focus on topics relating to education, child nutrition and public service delivery. His current research focuses on three areas: (a) analyses of the dynamics of human capital formation through childhood (b) socially-determined inequalities in learning and health (c) reforms to improve the delivery of basic education in developing countries.

Acknowledgements

We are grateful to Nicolas Pazos for excellent research assistance. Comments are welcome at asanchez@grade.org.pe and abhijeet.singh@ucl.ac.uk

About Young Lives

Young Lives is an international study of childhood poverty, following the lives of 12,000 children in 4 countries (Ethiopia, India, Peru and Vietnam) over 15 years. **www.younglives.org.uk**

Young Lives is funded by UK aid from the Department for International Development (DFID) and co-funded by Irish Aid from 2014–16.

The views expressed are those of the author(s). They are not necessarily those of, or endorsed by, Young Lives, the University of Oxford, DFID or other funders.

Summary

We use unique individual-level panel data from India (Andhra Pradesh), Peru, and Vietnam on a cohort of individuals surveyed from the age of 8 years to 19 years to study factors affecting enrolment in higher-education in these middle-income countries. We document (a) that similar to nationally representative data, the proportion having accessed higher-education at this age is high (˜35-45 per cent); (b) that there are steep gradients in higher-education access across wealth and parental education; (c) that a substantial part of the gradient with regard to parental education is explained by parental and child aspirations for education, at 12 years of age, and previous measures of learning; (d) that in contrast, wealth gradients decline much less with the inclusion of these variables, indicating that the correlation between household economic circumstances and higher-education access is only partly due to differences in early-childhood human-capital formation; and (e) that there are important differences in terms of gender in access to levels of higher-education (favouring boys in India and girls in Vietnam) and in the association of various household and individual characteristics and parental and child aspirations with enrolment in higher-education by 19 years of age. To the best of our knowledge, this is the first such comparative longitudinal analysis of access to higher-education in developing-country settings.

1 Introduction

Education levels have risen rapidly and impressively around the world in recent decades (Barro and Lee, 2013). Enrolment in primary schooling, an explicit global policy target in the Millennium Development Goals, is near-universal in most countries (UNESCO, 2015). Access to secondary schooling, while not as high, has also seen a rapid rise and also is now an international target for the Sustainable Development Goals.

Our focus in this paper is on the subsequent, tertiary, stage of education. While higher-education has also increased rapidly in many countries in the last two decades, we know much less about the relevant determinants of access.[1] The lack of knowledge in this area contrasts with the extensive literature on this topic in developed countries.[2] And yet, due both to the rising prevalence of higher-education and the changing economic structures of many developing countries, especially those now attaining middle-income status, questions of access to higher-education are already worth considering and likely to increase in importance in the near future.

We attempt to address this gap partially for three developing countries – India, Peru, and Vietnam – using rich panel data on a cohort of individuals collected for over a decade preceding college enrolment. Specifically we focus on three related questions. First, we analyse patterns of access to higher-education, focusing in particular on inequalities in access in terms of socioeconomic characteristics and gender, to ask how inequitable is higher-education access for recent cohorts. Second, we use panel-based decompositions employing rich child-level data collected for over a decade preceding college enrolment, to establish the extent to which these inequalities apparent in early adulthood reflect household circumstances through childhood vs. intrahousehold choices or the aspirations and investments in learning by individuals. Third, we investigate the extent to which the factors affecting eventual access to higher-education vary in their effects across gender, across urban and rural areas, and across parental education.

[1] For a survey of current knowledge of higher-education in developing countries, see Kapur and Crowley (2008).

[2] For panel-based analyses of the determinants of access to higher-education in the USA and the UK, which are closest in motivation to our own work, see for instance Cameron and Heckman (1998),Cameron and Taber (2004), Cameron and Heckman (2001), Keane and Wolpin (2001), Chowdry et al. (2013), and Crawford and Greaves (2015). Additionally, several papers whose primary focus is on specific interventions at earlier school-going ages use access to higher-education as a relevant outcome variable: examples of such studies include those of school quality (Deming et al. (2014)), high-stakes performance incentives in high schools (Angrist and Lavy, 2009), Charter schools (Angrist et al., 2016), and student aid for easing borrowing constraints. See Rau et al. (2013).

These questions are of interest for several reasons, even in developing-country contexts where access may not be universal even at lower levels. First, higher-education may have a substantial effect on the future employment, wages, and security of tenure of individuals.[3] Inequality in access to higher-education, especially on grounds of pre-determined social and economic characteristics, may therefore translate into inequality in outcomes later in life and may have significant distributional consequences. The possibility of non-pecuniary benefits, such as improvements to health, further accentuates this concern.[4] Second, higher-education may have direct effects on the economic prospects of countries (see e.g. Bloom et al. 2014). Inequality in access to higher-education, when arising from factors unrelated to future worker productivity, is then an illustration of a misallocation of resources with implications not only for the individuals but also for the broader economy.[5] Finally, if higher-education has intrinsic value for individuals, at least to some degree, then inequality in access to higher-education can have direct consequences for individual welfare.

These concerns are likely to be most relevant for middle-income developing countries. Typically, in such contexts, access to schooling has already increased rapidly, the economic structure has changed, and the demand for higher-education has also risen with rising per capita incomes. These trends, moreover, are likely to continue — and the salience of this area, therefore, is likely to grow. For instance, globalisation and skill-based technical change may raise the skill premium from tertiary education and cause a polarisation in labour earnings similar to that in the USA[6]; if so, our first concern, of distributional consequences for individuals, is likely to become more germane. Similarly, skills developed in higher-education may be particularly relevant for economic growth at the stage of development in which middle-income countries find themselves; in which case, our second concern, about the macroeconomic impacts of inequality in higher-education access, may increase in relevance.

[3]These concerns are made more serious by the possibility that private economic returns to education in many developing countries may well be convex, with a year of tertiary education adding substantially more to wages than education at lower levels. See, for instance, Bigsten et al. (2000); Söderbom et al. (2006).

[4]See, for instance, Cutler and Lleras-Muney (2008) and Oreopoulos and Salvanes (2011).

[5]There is a large and rapidly growing macroeconomic literature on the misallocation of human capital. See, in particular Hsieh and Klenow (2009) and the recent summary article by Restuccia and Rogerson (2013). Apart from its direct contribution to factors of production (augmented human capital), higher-education could affect economic growth through potential spillovers into the strengths and weaknesses of state institutions, which could also affect economic outcomes (Kapur and Crowley, 2008).

[6]See, e.g., Autor et al. (2003), Acemoglu and Autor (2011), and Goldberg and Pavcnik (2007).

The data used in this paper come from the Young Lives study which has tracked a cohort of individuals born in 1994/95 over four survey rounds from 2002 to 2014 in India (Andhra Pradesh state only), Peru, and Vietnam.[7] The data are particularly suitable for this analysis. In the most recent round of data collection, individuals in this cohort were aged around 19 years and had typically either made the transition into tertiary education or dropped out of education.[8] This ensures that our information is on the most recent cohorts entering higher-education in these countries, which is very useful given relatively rapid changes in higher-education access in recent decades. Second, rich information on household circumstances, parental and child aspirations for education, detailed measures of academic achievement and ability, and household and individual investments in education allow for extensive analyses of the factors determining access. Third, the three study countries offer an excellent spread of the developing countries for which these issues are likely to become most relevant in coming decades. In all three countries, access to primary schooling is near-universal, and secondary-school access has also been rising rapidly. Further, all of these countries are middle-income countries now, with rapid rates of growth in recent decades; they are likely to need a larger pool of skilled tertiary-educated workers to sustain these growth rates in the future.

The key contribution of our paper is to present what is, to our knowledge, the first investigation to focus on this increasingly important research area using a long panel, with detailed data and comparable measures across countries with very different labour markets, educational systems, and institutonal sectors but of similar middle-income status with high economic growth in the past decade.

We document (a) that similar to nationally representative data, the proportion having accessed higher-education at this age is high (˜35-45 per cent)[9]; (b) that there are steep gradients in higher-education access across wealth and parental education; (c) that a substantial part of the gradient with regard to parental education is explained by parental and child aspirations for education, at 12 years of age, and previous measures of learning; (d) that in contrast, wealth gradients decline much less with the inclusion of these variables, indicating that the correlation between household economic circumstances and higher-education access is only partly due to differences in early-childhood human-capital formation; and (e) that there are

[7]The Young Lives study also collects data on individuals in the same birth cohort in Ethiopia. We have not used data on Ethiopia in the present analysis since a substantial portion of the sample are yet to complete secondary education and because our focus here is restricted to middle-income countries.

[8]The precise distribution is presented in Section 3.

[9]For approximate comparison, the initial participation rate for 17-30 year olds in the UK was about 47% in 2013/14 (Department of Business, Innovation and Skills, 2015).

important gender differences in access to levels of higher-education (favouring boys in India and girls in Vietnam) and in the association of various household and individual characteristics and parental and child aspirations with enrolment in higher-education by 19 years of age. To the best of our knowledge, this is the first such comparative longitudinal analysis of access to higher-education in developing-country settings.

The rest of this paper is divided as follows: Section 2 presents, as a background for our panel analysis, an analysis of the trends in the access of men and women to higher education across successive cohorts in these three countries, using nationally representative datasets; Section 3 presents details about the Young Lives data and sample and presents descriptive statistics of access to higher-education in the sample across various dimensions; Section 4 presents regression-based analyses of the factors affecting access; Section 5 investigates heterogeneity in the associations by sex, urban/rural location, and parental education; and Section 6 presents our conclusions.

2 Background

In this section, we foreground our later analysis of access of recent cohorts to higher-education by looking at (a) a basic description of the educational systems in the three countries, with a focus on schooling which acts as a feeder into the higher-education system, to understand the arrangement of the schooling systems and the typical pattern of enrolment and dropping-out for recent cohorts in these settings and (b) trends in changes to higher-education access across decades in these settings for men and women.

2.1 Educational context of the study countries

Higher-education enrolment is an advanced stage of the formal education process which begins much earlier in childhood. Understanding patterns of access to higher-education requires, therefore, at least a minimal understanding of the organisation of primary and secondary schooling in the relevant context. To describe these patterns we utilise the latest nationally representative household survey that we could find freely available. For India, this is the India Human Development Survey - II (2012), for Peru the National Household Survey (2010), and for Vietnam the 2010-2011 round of the Multiple Indicator Cluster Survey (MICS).

Figure 1 depicts access to education in recent years among individuals between the ages of 5 and 18 years in each country, using data from recent nationally representative surveys.[10] Red vertical lines indicate the official starting ages for primary and secondary education. Although the starting age for the primary level is common in these countries (6 years), the length of each educational level is not. In particular, elementary schooling is longer in India (8 years) compared with Peru and Vietnam (6 and 5 years respectively). At the same time, the secondary level in India is shorter. This balances things out compared to Vietnam, but not with respect to Peru. Assuming no repetition, a child requires 12 years to complete school education in India and Vietnam, compared with 11 years in Peru.

In all three countries, enrolment rises to nearly 100 per cent within a year of the starting age for formal schooling. In Peru, and to a lesser degree Vietnam, it stays at these near-universal levels for the duration of primary schooling; in India, children begin to drop out of education at younger ages, with this decline particularly pronounced from about the age of 10 years. In all countries, the decline in enrolment is steeper at secondary-school ages, with significant differences in enrolment across adjacent years.

Gender differences in enrollment appear differentially across countries. In India, girls seem more at risk of dropping out from the primary-school age onwards, and the gap between male and female enrolment grows throughout the educational trajectory. In Vietnam, there is little evidence of a gender-based difference in enrolment at primary school ages, but a noticeable gap favouring girls emerges in secondary schooling, with the gender gap in enrolment being greatest at the end of secondary schooling. In Peru, by contrast, there seems to be no evidence of gender gaps throughout schooling. There is some indication of a modest gap favouring girls in access to higher-education which seems to exist only at that stage and is small in magnitude.

Finally, in talking of higher-education, it is worth providing some background notes also about the quality of education which schools impart in the three countries. The quantity of education differs widely in the three countries. In the 2012 PISA assessment round, which aims to provide international comparative assessments of the mathematics, reading, and science skills of 15-year-old students, Vietnamese students score very highly (above, for example, the UK and the USA) while Peru ranked the last among the 65 countries covered (OECD, 2013); India had not participated in the 2012 round, but in the previous round (2009) the learning

[10]This figure shows the percentage of enrolled children and young adults of different ages in the national survey for each country.

levels of Indian students were found to be significantly lower than these in Peru and second-lowest among the countries then surveyed, higher only than Kyrgyztan (Walker, 2011). Similar ordinal rankings across the three countries are demonstrated in the common tests administered by the Young Lives surveys: as Singh (2014) documents, while there are some gaps already apparent at 5 years of age between these countries, before children have enrolled in formal schooling, these gaps widen significantly by the age of 8 years, with divergence from Vietnam being accounted for almost entirely by the much higher productivity of primary schooling in producing learning gains. In this paper, we will not be able to address questions of what individuals are learning in higher-education, or indeed their degree of preparedness for college, but it is worth keeping in mind that the ability of students to thrive in college may be very different across these three settings, reflecting the differential productivity of the national schooling systems.

2.2 Trends in access to higher-education

Figure 2 presents by year of birth the proportion of men and women in the nationally representative data who had attended higher-education.[11] In all three countries, the percentage of individuals who have accessed higher-education is relatively high for recent cohorts, ranging between 30 and 40 per cent for cohorts born in the 1990s.

In all countries, this proportion has risen over the recent decades, but the timing and pace of increase seems different across the countries: in Peru, we observe a steady increase at a near-constant rate for cohorts born mid-1950s onwards for a period of about 25 years; in Vietnam, this increase begins later, evident mostly for cohorts born after the 1970s but then increasing at a relatively rapid rate; finally, for India, we see a slow but steady increase throughout the period covered by Figure 1, but with the pace of increase having increased sharply for cohorts born after the 1980s.

This figure also reveals interesting patterns by gender. In all three countries, men born in earlier cohorts (1950s and 1960s) were more likely to have had higher-education than women in their cohorts. Subsequently, however, the patterns diverge between India, on the one hand, and Peru and Vietnam on the other. In

[11]Note that these are individuals who are surveyed in the relevant nationally representative survey. This is not a random subsample of individuals born in the relevant year, due to mortality (and resulting attrition in the data). Mortality risk varies by age and gender. The magnitude of the problem that this poses is likely to be more severe for older birth cohorts. However, we think the magnitude of this problem is not of great importance for the description of trends, which is all we seek to achieve through this figure.

India, this pro-male pattern in access to higher-education seems to have widened steadily for cohorts, born before 1980 and then stayed constant in the period since. In contrast, Peru and Vietnam display an extensive period of convergence between genders and, in recent cohorts a marked reversal of the gender gap: for current cohorts, access to higher-education is higher for women than for men, with this reversal being particularly pronounced in Vietnam.[12] The levels of enrolment in higher-education, and the differences in enrolment by gender, will be a key reference point for us when analysing the Young Lives data in Sections 4 and 5.

3 Data and descriptive statistics

3.1 Data

Data used in this survey come from Young Lives, a longitudinal study of childhood poverty which has collected data on two cohorts of children – born in 1994/95 and 2001/02 – since 2002 in Ethiopia, India (Andhra Pradesh), Peru, and Vietnam.[13] In this paper, we use data on the 1994/95 cohort from India, Peru, and Vietnam.[14] In each country individuals from the older cohort were sampled following a similar multi-stage sampling procedure. First, 20 clusters were selected in each country. Then, in each cluster around 50 households that had at least one child aged 7 to 8 were enrolled (between 25 and 50 households in the case of Peru). Although the samples are not statistically representative of the national populations, comparisons with nationally representative datasets show that they are informative of a large range of living-standard conditions in each of the selected countries (Escobal and Flores, 2008; Kumra, 2008; Nguyen, 2008).

[12]The larger proportion of girls entering higher-education in Peru and Vietnam is similar to that in many OECD countries and illustrates a broader trend in developing countries, which is particularly pronounced in Latin America and Southeast Asia. This is summarised in Grant and Behrman (2010), who look at enrolment in schooling for a wide age range, using nationally representative data from 38 developing countries.

[13]Over the period of this study, the state of Andhra Pradesh (with a population of 84 million people in 2011) was bifurcated into Telangana and Andhra Pradesh states in 2014. Throughout this paper, when referring to Andhra Pradesh, we mean the undivided state as it existed until 2014. In terms of enrolment and learning outcomes, Andhra Pradesh is typically close to all-India averages (see e.g. Pratham 2015). In the paper we will often refer to results for 'India' or 'the Indian sample'; readers are requested to keep in mind that the sample is exclusively based in this one state.

[14]The younger cohort, born in 2001/02, were aged only 12 in 2013/14 and hence not of interest for this particular paper. The reasons for excluding Ethiopia have already been mentioned (see footnote 4).

The first survey took place in 2002, immediately after the households were enrolled, with further rounds of data collection in 2006/7, 2009/10, and 2013/14. The timing of the different rounds and the ages of individuals in the sample at these points are provided in Figure 3. Attrition rates across the four rounds are relatively low over this 12-year period of data collection: 4.3 per cent in India, 10.3 per cent in Peru and 11.3 per cent in Vietnam. In the case of Peru and Vietnam, attrition increased by around 5 percentage points between the third and fourth rounds as a consequence of high migration rates and increases in refusal rates among the young people, many of whom have moved and become economically independent.

Table 1 presents information on some basic characteristics of the sample in each country. In India and Vietnam, the sample is predominantly rural, while in Peru a majority of the sample lives in urban areas.[15] Peru and Vietnam both have a significantly higher level of maternal education: the proportion of mothers in our sample with no formal education is about 60 per cent in India but about 10 per cent in Vietnam and Peru; most mothers in Vietnam and Peru report having at least secondary education. A similar picture is also evident for paternal education, although levels in all countries are higher than for maternal education, especially so in India. These differences are less stark when comparing a simple wealth index which aggregates ownership of durable goods, access to services, and housing quality; this probably reflects the basic measure of prosperity that such measures capture. The Young Lives survey also collected anthropometric data on these individuals in each round of the survey; at the age of 8 years, it is evident that the mean height-for-age z-scores (constructed with reference to the WHO 2006 standards) are significantly below the mean of the international reference distribution. In each of the countries, reflecting the sample design, the survey individuals were aged just under 19 years on average.

An aspect of the Young Lives data that is very attractive is that questions were asked of children and their primary caregiver (nearly always the mother) about their aspirations for education for the index child.[16] As is evident from the Table 1, both parental and child aspirations are high: most parents/children respond with

[15]National statistics classify about 70 per cent of the population in Peru as urban, with about 30 per cent of the total population concentrated in the capital city, Lima. This is in stark contrast to the other two countries.

[16]Educational aspirations were mesaured through the following question: *Imagine you had no constraints and could study for as long as you liked, or go back to school if you have already left. What level of formal education would you like to complete?* An analogous question was asked of the caregiver (typically, the mother) of the individual. Both own and parental educational aspirations were measured for the first time when the individual was aged 11 to 12, and this is the observation that we use.

aspirations of going to university. In the 2006 and 2009 rounds, the survey also collected scores for the test in mathematics and receptive vocabulary which were administered to all sample children, regardless of current enrolment status.[17] In this analysis we use the measure of such aspirations as collected in the 2006 round of the survey, when the sample children were aged 12 years on average.[18]

3.2 Descriptive statistics of higher-education enrollment

Table 2 presents information on the current enrolment status of our sample. In the Indian sample, around 43 per cent of the sample are already enrolled or have been enrolled in higher education, and about 46 per cent in Peru and 36 per cent in Vietnam.[19] Within those individuals never enrolled in higher education, most of

While aspirations are collected in terms of the precise grade of schooling or university, in our analysis we reduce this information to a dichotomous variable which equals zero if the respondent aspired to an education up to secondary level or lower, and 1 if they aspired to higher-education.

[17]In the analysis, we normalise the test scores to have a mean of zero and the standard deviation of one within each country sample. This enables an easier interpretation of coefficients on test scores in regression results but comes at the cost of incomparability of test scores across countries. An international comparison of learning levels is not our purpose here; readers interested may however want to see Singh (2014), where such comparisons are presented and analysed.

[18]We choose the measure from this age, rather than from the later 2009 round, since a vast majority of the sample were still enrolled in school, in contrast to the 2009 round, when about 22 per centof the children in India and Vietnam and 15 per cent of the children in Peru had already dropped out of school. While the question as posed does incorporate the possibility of eliciting a response even when individuals have already dropped out of school, it is possible that this measure is contaminated by the experience of having dropped out of school already.

In the 2006 round, aspirations were not collected in India for children who had dropped out of schooling. For this sub-group (˜10 per cent of the sample), we use the aspirations variable collected at 15 years of age as the measure of aspirations. Although this strategy is not ideal, given that aspirations by the age of 15 are likely to be much more influenced by school performance than at the age of 12 (thus more informative of objective expectations than aspirations), we believe this strategy to be superior to the alternative of excluding these observations from the analysis altogether. This is because in this group of children both aspirations for higher-education and actual enrolment in higher-education are likely to be lower than for the rest of the children. Therefore, excluding these individuals from the analysis would lead to sample-selection bias. In the case of Vietnam and Peru the number of children who did not answer the aspiration questions in 2006 is very small (5 and 3 students), thus in this case no further imputation was made.

[19]This figure includes all individuals who have enrolled in higher-education up to the survey round in 2013/14, thus including those individuals who had enrolled in higher-education but potentially dropped out by the time of the survey as well as, exceptionally, any individuals who had completed their post-secondary studies. We make this choice since our focus in this paper is to assess factors determining initial access to higher-education and not on completion of higher-education; given the currently enrolled status of most of the individuals in our sample who had at any time accessed higher-education, we cannot (yet) distinguish between those who have enrolled but may not complete higher-education and those who eventually get degrees or other post-secondary qualifications. If we consider only those who were enrolled at the time of the last survey round (2013), around per cent of the Indian and Peruvian samples were enrolled, and 35 per cent in Vietnam; the difference from the higher figure quoted above results primarily from drop-out in higher-education that had already occurred by 2013/14. Among those who are

them have already dropped out of school at this stage. In India and Peru, only 10 per cent of the total sample are still enrolled in school education. In Vietnam the figure of those still enrolled in school is higher, at about 18 per cent. Most sample individuals who are still enrolled but are not in higher-education are in the last year of secondary schooling.[20] As in the national data, there are significant gender gaps in enrolment favouring young men in India and young women in Vietnam; however no evidence of a gender gap is observed in the Peruvian sample at this level of agreggation.

The presence of some school-going individuals at this age complicates our analysis: presumably some, but not all, of these individuals will eventually enrol in higher education, but this is not yet observed. Given that our interest lies in understanding the factors associated with higher-education attendance, in the analysis that follows we will take a conservative approach and focus only on those individuals who have already enrolled or were enrolled in higher-education: our key (binary) dependent variable will be defined to equal 0 if an individual is never enrolled or enrolled in school education, and 1 otherwise.[21] In addition, for the current analysis we do not distinguish between attending university and attending technical / post-secondary vocational institutions. This is both because such distinction is not material to our core purpose of studying access to tertiary education and because the precise definition of what qualifies as technical / vocational education varies across the countries in our sample.

Table 3 presents initial descriptive statistics on the characteristics of those individuals who have had higher education. In all countries, sample individuals in urban areas are more likely to have been enroled in higher-education than those in rural areas. Higher-education access increases with the wealth of the household, maternal education, and paternal education, with a pronounced gradient in all these dimensions. It is also more prevalent among the first-born individuals. These patterns are observed in all countries. As in Table 2, there are gender gaps in enrolment favouring young men in India and young women in Vietnam. Interestingly, however, these gender gaps seem to be concentrated largely in rural areas, among the less educated and the less wealthy. Finally, it is interesting to

currently enrolled in higher-education, 60 per cent are enrolled in private institutions in India, 63 per cent in Peru and 11 per cent in Vietnam.

[20]Of those still attending school, 87 per cent in Vietnam, 61 per cent in India and 60 per cent in Peru are in the last year of secondary school.

[21]Strictly, this may be interpreted as a measure of whether an individual has made it to higher-education at the right age. However, our purpose is to analyse access to higher-education (and not necessarily access at the right age). In the appendix, we also present results assuming that all individuals currently enrolled in secondary schooling will also go to college.

observe the existence of a pro-male bias in rural areas and among the less wealthy in Peru, something that is not observed at the aggreggated level.

3.3 Aspirations for higher-education

Aspirations for education, and for higher-education in particular, are central to our analysis. We focus on aspirations measured in 2006, when the child was aged 12 years, a point of transition from elementary to secondary level. In Table 4 we report parental aspirations for the child and the child's own aspirations at that moment of time.[22] In our country samples, child aspirations for higher education are the highest in Peru (91 per cent), followed by Vietnam (76 per cent) and India (64 per cent). A similar pattern is observed for parental aspirations, which are highest in Peru (94 per cent) followed by Vietnam (78 per cent) and India (70 per cent). Parental aspirations are only slighly higher than child's own aspirations. It is also imporant to note that there are some differences in aspirations by gender that tend to reflect of the gender gaps eventually observed in higher-education at the age of 19: pro-male in India and pro-female in Vietnam. In the case of Peru, a pro-female bias is also observed. The gender gap in child aspirations is larger in India (14 percentage points), followed by Vietnam and Peru (5 and 4 percentage points).[23] A similar pro-male and pro-female gap is observed for parental aspirations in India and Vietnam (16 and 4 percentage points, respectively), whereas in the case of Peru the pro-female bias reverts to a small and most likely insignificant pro-male bias (2 percentage points)[24].

4 Factors affecting access to higher-education at 19

4.1 Empirical specifications

Patterns presented in Table 3 are useful, but only partially indicative of the factors affecting access to higher-education in our data. In this section, we go further and

[22]In the case of India, there were 104 children who were not attending school in 2006 who did not have to answer the question about aspirations for education. For this group of children we decided to consider the aspirations for education reported in 2009 (when they were aged 15 years) instead.

[23]The proportion of boys (girls) who aspire to higher-education is 75 per cent (63 per cent) in India, 89 per cent (93 per cent) in Peru, and 74 per cent (79 per cent) in Vietnam.

[24]The proportion of parents that aspire for higher-education for their boys (girls) is 78 per cent (62 per cent) in India, 95 per cent (93 per cent) in Peru, and 76 per cent (80 per cent) in Vietnam. That girls' aspirations for their own education are higher than their parents' aspirations for them in these data has previously been noted and analysed in Dercon and Singh (2013).

explore the partial correlations of a set of household and individual characteristics with higher-education access at 19 years. We focus in particular on three sets of factors: potential household-level determinants such as wealth and rural location; potential drivers of intra-household choices of which child to send to college, such as sex, birth order, and parental aspirations; and potential individual determinants of reaching higher-education, in particular measures of human capital from earlier in childhood (such as test scores and nutrition) and individuals' own aspirations in childhood for completing higher education. For all time-varying factors, most importantly the aspirations and human-capital measures, we use observations from when the individual was aged about 12 years and decisions about when to drop out of education had not been realised for most of the sample.

Our particular interest is in understanding two specific sets of relationships: (a) to what extent does higher-education access depend on household socioeconomic status and fixed individual characteristics such as gender and birth order; and (b) to what extent is it related to the fomation of human capital in childhood, at earlier stages of education.

We use linear probability models to measure these relationships. We make this choice, instead of non-linear limited dependent variable models such as probit or logistic regressions, because of the ease in interpretaton of coefficients as marginal effects, because we will be using some site-level fixed effects in our analysis, and mostly because our primary aim is to evaluate the partial correlations and not to forecast probabilities; we will, however, show in the Appendix that this choice does not guide our results.

Our regression specification is as follows:

$$
\begin{aligned}
Y_{ij,19} &= \alpha + \beta_1.X_{ij} && (1)\\
&\quad +\beta_1.ParentalAsp_{ij,12} + \beta_2.ChildAsp_{ij,12} && (2)\\
&\quad +\beta_3.PPVT_{ij,12} + \beta_4.Math_{ij,12} && (3)\\
&\quad +\lambda_j + \epsilon_{ij} && (4)
\end{aligned}
$$

The dependent variable $Y_{is,19}$ is a binary variable taking the value of one if individual i in cluster s is enrolled in higher-education in the 2013/14 round (aged 19 on average). The analysis is carried out step-wise, as indicated in the equation above. The first specification (Equation 1) regresses higher-education enrolment on a vector of background characteristics X_{ij} that contains the following variables:

mother´s and father´s level of education (no formal education is the omitted group), household location in the distribution of the wealth index by terciles at the age of 7-8 years (the lowest tercile is the omitted group), area of location (rural or urban) at the age of 7-8 years, height-for-age at the age of 7-8 years, birth order of the individual, number of siblings, age in years, and gender.[25]

Equation (2) further adds two variables - caregivers' and individuals' own aspirations of accessing higher-education - to the variables already in X_{ij}; both caregiver and child aspirations are included here as binary variables equalling one if the respondent had expressed an aspiration for post-secondary education. Coefficients on the aspirations variable, in this set-up, may be interpreted as the correlation of these variables with higher-education, having accounted for the difference in aspirations across different wealth groups, different levels of parental education, and rural/urban areas, etc. These estimates allow us to answer whether the aspirations variables add any further information to our analysis, and whether aspirations function as a channel for realising the observed differences in higher-education attendance across, say, different wealth groups or sexes.

In Equation (3), we further add test scores at the age of 12 on mathematics and receptive vocabulary to the specification. Two issues are central in the interpretation of these results. The first is the interpretation of the coefficients on learning themselves: they should be interpreted as the extent to which learning accumulation by 12 predicts higher-education access at 19; they may be suggestive of the effects of skills gained through primary education on long-term outcomes (in this case, higher-education), but such a causal interpretation cannot straightforwardly be sustained in the specification above, since lagged achievement also proxies for individual-specific heterogeneity and unobserved determinants which affect higher-education directly and are also correlated with test scores.

The second issue relates to the interpretation of the other coefficients in the regression. If lagged achievement acts as a summary statistic for fixed ability

[25]The reason for using rural location from 7-8 years, rather than from later in life, is that some families may move to urban areas in order to be closer to secondary schools and colleges. This would complicate the interpretation of the regression coefficient on rural location. The reason for using wealth from 7-8 years of age is subtly different. Children who drop out of education by 12 or 15, the next age group for which information is available, are themselves more likely to be contributing economically to the household through paid work, work on the family farm or business, or domestic chores and care activities that free up other members of the household for remunerative employment: thus wealth at 15, and even more so at 19, may well reflect the consequences of having dropped out and not enrolling in higher-education. Finally, the reason for choosing the height-for-age z-scores from 7-8 years of age is that the next round of anthropometric measurements (at 12 years) are made inherently noisier by the onset of puberty for some members of the cohort.

and past investments in human capital, and sufficiently absorbs unobserved individual-specific heterogeneity, then the other coefficients may be interpreted causally as policy effects. This latter condition is the fundamental assumption underlying value-added models of achievement where controlling for lagged achievement is assumed to provide conditional exogeneity. Such models have recently been found to be unbiased in comparison with various experimental and quasi-experimental estimates,[26] including for predicting higher-education enrolment and the type of college attended (see Deming et. al. 2014). This is encouraging because, as pointed out clearly by Todd and Wolpin (2003, 2007), value-added estimates can still be biased by both measurement error and unobserved heterogeneity that affects trends. While recognising that such an interpretation may be tenable, we do not seek here to claim full conditional exogeneity of the coefficients estimated; our goal here is not the estimation of specific policy effects, and, despite many encouraging results assessing whether controlling for lagged achievement can provide reasonable estimates of the same, it is always possible that residual sources of omitted-variables bias exist.[27] Our more modest claim, rather, is that the ability to control for lagged achievement reduces substantially the bias due to unobserved individual-specific heterogeneity. Further, we interpret the estimates on the other coefficients as being results conditional on human-capital divergence that had already emerged by middle school and thus reflecting the divergence that emerges in the latter stages of education.

In Equation (4) we finally add a vector of fixed effects for the initial communities in which the individuals were sampled in, effectively restricting comparisons to those between individuals in the same initially sampled cluster. Standard errors in all regressions are clustered by the initial community in which individuals were sampled. Regressions are estimated separately for each country sample.

[26]See, e.g., Angrist et al. (2013); Chetty et al. (2014); Deming (2014); Deming et al. (2014); Kane and Staiger (2008); Kane et al. (2013); Andrabi et al. (2011) and Singh (2015). The latter two studies cited are in developing countries, with the rest being based on data from the USA.

[27]In Singh (2015), for instance, one of us has previously demonstrated using Young Lives data from India that some of the structural assumptions implicit in the cumulative-effects value-added model are, in fact, violated in the data (pp. 26-28). This issue is of secondary importance in that particular application, since the magnitude of the violation is small and the resultant estimates of the policy effect being evaluated (private-school effects) are not substantively biased and not statistically distinguisable from experimental estimates. While these results, and similar other results from various contexts, are encouraging as a matter of practice, any attribution of causal impact for the coefficients here will require a much more careful analysis of the possibility of residual bias than we present here.

4.2 Results

Results from the above exercise are presented in Table 5. The results from the first specification, in all countries, mostly reaffirm patterns that were evident in the descriptive statistics. In both India and Vietnam, we continue to see important gender differences in higher-education attendance: conditional on these various characteristics, girls in India seem about 11 percentage points less likely to engage in higher-education than boys; in Vietnam, the differential is reversed and girls are about 7 percentage points more likely to attend college. In all countries, there is a pronounced gradient with respect to household wealth: individuals in the middle and upper terciles of wealth are 10 and 21 percentage points more likely to enrol in higher-education in India than those in the lower third of the wealth distribution; in Peru, the upper tercile is 18 percentage points more likely to enrol; in Vietnam, individuals in the middle and upper terciles are 10 and 16 percentage points more likely to enrol. There is also a pronounced gradient with respect to maternal education: in India, adding 20(27) percentage points to the probability of higher-education access if mothers have secondary(higher) education as opposed to no schooling; adding about 19(41) percentage points in Vietnam for maternal secondary(higher) education; and in Peru, about 18 percentage points for maternal higher-education. Conditional on maternal education and various other socioeconomic characteristics, father's education much less consistently has a significant association with higher-education access for the sample. In this simple specification, we see little evidence of differences by birth order in Peru, but being the third child born (or more) reduces enrolment in higher-education by 8 and 10 percentage points in India and Vietnam, respectively. The total number of siblings has a significant negative association in Peru, but this is relatively modest in magnitude.

The one dimension in which the results from the linear probability models differ significantly from the bivariate patterns is in the specific context of rural areas. Whereas in all countries we had shown that the overall probability of having enrolled in higher-education is lower for rural areas, in the regression specification the coefficient for rural location invariably has a positive coefficient which is large and statistically significant for Vietnam: this suggests that the lower probability of higher-education in rural areas is intimately tied to the lower wealth and lower parental education in these areas but that, conditional on these characteristics, individuals with the same socioeconomic characteristics are more likely to transition

to higher-education from rural areas. This is an intriguing pattern and one that we shall explore more fully in the next section.

Looking at column (2) in each country, where we additionally include aspirations of the caregiver and the child, we see some interesting patterns emerge. The coefficients on these aspirations are invariably positive, and, in all countries, children's reported aspirations from the age of 12 are a significant predictor of higher-education attendance at 19. In India and Vietnam, additionally, the caregiver's reported aspirations also retain a significant positive coefficient. The inclusion of aspirations in India halves the negative coefficient on the female dummy variable; this indicates that lower parental and child aspirations are a channel through which gender bias manifests itself in the eventually lower higher-education attendance for Indian girls. In no country does the addition of aspirations cause much of a decline in the wealth gradient for higher-education access, which suggests that the observed gaps by wealth do not reflect 'aspirations gaps' for different economic groups. This is notable, since the 'aspirations poverty traps' have drawn attention in the theoretical literature (see e.g. Ray 2006; Mookherjee et al. 2010; Dalton et al. 2016) and other contexts (Bernard et al., 2011).

The further incorporation of test scores in Column (3) highlight that test scores at 12, especially in mathematics, strongly predict eventual higher-education access. Given that enrolment in higher-education is an advanced stage of having successfully transitioned through multiple previous stages of formal education, and because test scores are also supposed to soak up individual-specific heterogeneity, this is not surprising. The inclusion of test scores does cause declines in coefficients on parental education and wealth, but typically not by very much; this indicates that the effect of these socio-economic characteristics in raising the probability of later higher-education attendance does not result solely through the channel of early childhood investments or aspirations, but that liquidity constraints and current economic circumstances are likely to play a substantial role in determining higher-education access.[28] An interesting pattern that is immediately notable, however, is that coefficients on children's aspirations for higher-education decline significantly once test scores are controlled for. This could either represent a pattern

[28]We are, of course, unable to say this causally, since we do not have experimental, or otherwise plausibly exogenous, variation in either incomes or test scores. However, it is worth noting that these partial correlations do suggest that the differences in the contribution of liquidity constraints vs. early childhood invesments in explaining higher-education access is probably quite different in these contexts than in, for example, the USA, which is the site for most research on this question and where short-term liquidity constraints seem distinctly less effective than increasing invenments earlier in the educational and life stages of children (see e.g. (Cameron and Heckman, 2001; Keane and Wolpin, 2001; Cameron and Taber, 2004)).

wherein individuals revise their aspirations based on current performance, and hence both measures capture similar variation, or it could indicate that lower aspirations result in lower effort and thereby lower realised human capital even at 12.

Finally, including cluster fixed effects increases the proportion of variation that we can explain, thus indicating that the location of households is importantly associated with eventual higher-education attendance in ways not fully captured by our previous controls. But the partial correlations with most variables are left essentially unchanged, indicating that differences in wealth and education are as salient within communities in their association with higher-education attendance as the variation across communities.

As mentioned in Section 3, some individuals in the sample are still attending secondary school, and many of them are likely to go to college. Thus, our results could in part be reflecting the importance of the factors associated with higher-education attendance at the right age, rather than just higher-education attendance. To deal with this possibility, in the Appendix we report results assuming that all individuals currently enrolled in secondary schooling will continue to higher-education (Table 9). In addition, to verify that the choice of a linear probability model is adequate, we report results using a Logit specification. The marginal effects obtained from such model are reported in Table 10 (see Appendix). As it can be seen, in both cases results are largely similar to those reported in Table 5.

5 Heterogeneity in the predictors of achievement across groups

Thus far, we have modelled the relationship of various characteristics to eventual enrolment in higher-education by 19 as being the same for all individuals in the sample. However, it is possible that there is significant heterogeneity in this respect - for example, maternal education may have very different predictive ability for the higher-education of girls vs. boys. In this section, we explore the possibility of such heterogeneity further along three dimensions - gender, rural location, and parental education. We explore this heterogeneity in the possible effects of various fcators by splitting samples along these dimensions and allowing all coefficients to differ across the groups.[29]

[29]This flexibility comes at a cost. For most coefficients, especially in the case of rural/urban splits where the sample is mostly concentrated in one or the other category, we will not be sufficiently

5.1 Heterogeneity by gender

The first dimension of heterogeneity that we investigate is across the sex of the sample individual. Splitting the sample by sex in each country, we present estimates of specifications (1) and (3) i.e. first only regressing higher-education attendance with a vector of background characteristics X_{ij} and then presenting a fuller specification also including aspirations and test scores. Results from this exercise are presented in Table 6.

As is evident, there is interesting and significant heterogeneity across sex in the association of various factors with eventual higher-education attendance. This heterogeneity is most notable in the Indian sample. Looking first at the various background characteristics, it is evident that being in the richest tercile of wealth has a larger positive effect for boys than for girls, as is also true, with a large coefficient, for living in rural areas. This is, however, nuanced by a somewhat larger positive association for maternal secondary education and paternal higher-education for girls' access to higher-education than for boys. Much more interestingly, it appears that parental and child aspirations for higher-education have a significant effect only in the case of girls, making them significantly more likely to transition to post-secondary education; coefficients on these variables in the male sub-sample are very imprecisely estimated and not significantly different from zero. Looking also at the change in coefficients in both male and female subsamples, we note that it is these aspirations that serve as the channel for the disproportionate protective role of maternal education for girls. Splitting the sample by gender in India also reveals interesting heterogeneity in birth-order effects, with third-born girls being less likely than their elder siblings to advance to post-secondary education. However, this is not the case for boys. In fact, second-born boys are more likely to advance to post-secondary education than their elders.

In Peru and Vietnam, there are some indications of gender-based heterogeneity in these partial correlations. The first of these concerns rural location: in Peru, conditional on other characteristics, boys from rural areas are more likely to enrol in higher-education than their urban counterparts, but the same is not true for girls. In both Peru and Vietnam, although in the latter only modestly so, maternal education seems to play a more salient role for boys than for girls; across all three countries, thus, it appears that maternal education acts as an equaliser for the levels of access to higher-education across sex. Finally, as in India, it seems that children's

justified in most cases to reject the equality of coefficients across samples. The results in this case will be suggestive.

self-reported aspirations at 12 in Vietnam are predictive on higher-education access for girls but not for boys.

5.2 Heterogeneity by rural location

Some of the most surprising results in the previous tables relate to a pattern where, conditional on various background characteristics, children in rural areas seemed to be more likely to access post-secondary education than their urban counterparts (see Tables 5 and 6). Further, this pattern seemed to be driven entirely by men in India and Peru. In this section, we split the sample by rural location in each country in order to understand which categories of social or individual characteristics seem to be responding differently in urban vs. rural areas in relation to eventual enrolment in higher-education.

Table 7 presents the results for this particular exercise; as in Table 5, we present results for two specifications for each category (urban and rural) within country. The clearest results are those on gender across urban and rural areas: in both India and Peru, girls in rural areas are less likely than boys to eventually reach higher-education, but there are few indications of such bias in urban areas, where the coefficient on female dummy is usually positive and sometimes significant. In Vietnam, this pattern holds true in the opposite direction; there is a pronounced pro-girl difference in rural areas, but no signs of similar differences in urban areas. We also observe that birth order seems to play a role in rural areas in India and Vietnam, but not in urban areas. There are few other notable differences.

5.3 Heterogeneity by parental education

The final dimension for our investigation of heterogeneity in partial correlations is by parental education (Table 8). Here, we divide the sample into two categories: individuals with and without at least one parent with secondary education. As in Tables 6 and 7, we present specifications with and without aspirations and lagged test scores.

In India and Vietnam, parental education acts as a significant factor for increasing women's education levels. In the Indian sample, there is no discernible gender bias in higher-education enrolment in families with at least one parent educated at the secondary level — in stark contrast to families with lower parental education where a pronounced gender difference favours boys. In Vietnam, on the other hand,

parental education seems to exacerbate gender differences – the pro-girl bias noted in previous tables seems to exist only in households with at least one parent educated at secondary school level or higher. For most other coefficients considered, we do not observe strong evidence of heterogeneity.

6 Discussion

Access to higher-education, despite its likely relevance for middle-income countries, remains understudied. This paper has tried to address that gap partially by presenting the first individual-level longitudinal analysis of factors determining access to higher-education in three middle-income countries for recent cohorts. We have documented that there are important inequalities in access to higher-education arising from parental background, household wealth, location, and gender – as we highlighted in the introduction, these inequalities are important not just in terms of distributional concerns but also because they may have implications for the economic prospects of these countries. In keeping with a vast literature on human-capital formation, we document that much of the inequality in higher-education access appears early, and is evident in the correlation between early measures of learning and later enrollment in higher-education. However, as we also document, accounting for these differences which emerge in school years, or the aspirations that children and parents express for higher-education, reduces only a part of the correlation with household wealth and higher-education access, indicating a potentially important role for liquidity constraints at secondary school ages and afterwards.

While these results are novel and important, they remain limited in important ways. In particular, we do not claim to provide causal treatment effects for any particular determinants of eventual access to higher-education. More generally, the literature on higher-education in developing countries remains constrained in a number of directions. To give a few examples within this broad research area, we are not aware of even careful descriptive analyses of the trends in the 'college premium' on wages or in marriage markets in these countries, or of analyses of the determinants and consequences of the type of higher-education track or chosen by individuals or careful structural analyses of eventual higher-education completion. We also know remarkably little aboout the quality of education imparted in these institutions; as experience with the rapid expansion of schooling in developing countries has demonstrated, actual skill levels imparted in education may lag significantly below

the formal expectations set out in curricula (Pritchett, 2013). These are all areas in which we hope to see our knowledge expand rapidly as concerns about tertiary education rise prominently in academic and policy thinking on education policy in developing countries.

References

Acemoglu, Daron and David Autor (2011) 'Skills, Tasks and Technologies: Implications for Employment and Earnings', chapter in *Handbook of Labor Economics* 4 (edited by David Card and Orley Ashenfelter) pp: 1043–1171.

Andrabi, Tahir, Jishnu Das, Asim Khwaja and Tristan Zajonc (2011) 'Do Value-added Estimates Add Value? Accounting for Learning Dynamics', *American Economic Journal of Applied Economics* 3.3: 29–54.

Angrist, Joshua and Victor Lavy (2009) 'The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial', *American Economic Review*: 1384–1414.

Angrist, Joshua, Sarah Cohodes, Susan Dynarski, Parag Pathak and Christopher Walters (2016) 'Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry and Choice', *Journal of Labor Economics* 34.2 (Part 1): 275-318.

Angrist, Joshua, Parag Pathak and Christopher Walters (2013) 'Explaining Charter School Effectiveness', *American Economic Journal: Applied Economics* 5.4:1–27.

Autor, David, Frank Levy and Richard Murnane (2003) 'The Skill Content of Recent Technological Change: An Empirical Exploration', *Quarterly Journal of Economics* 118.4: 1279–1333.

Barro, Robert and Jong-Wha Lee (2013) 'A New Data Set of Educational Attainment in the World, 1950–2010', *Journal of Development Economics* 104:184–198.

Bernard, Tanguy, Stefan Dercon and Alemayehu Seyoum Taffesse (2011) 'Beyond Fatalism: An Empirical Exploration of Self-efficacy and Aspirations Failure in Ethiopia', Technical report, Centre for the Study of African Economies, University of Oxford.

Bigsten, Arne, Anderk Isaksson, Mans Soderbom, Paul Collier, Albert Zeufack, Stefan Dercon, Marcel Fafchamps, Jan Gunning, Francis Teal, Simon Appleton, Bernard Gauthier, Abena Oduro, Remco Oostendorp and Cathy Pattillo (2000) 'Rates of Return on Physical and Human Capital in Africa's Manufacturing Sector, *Economic Development and Cultural Change* 48.4:801–827.

Bloom, David, David Canning and Kevin Chan (2014) '*Higher Education and Economic Development in Africa*', *International Journal of African Higher Education* 1.1: 22-57.

Cameron, Stephen and James Heckman (1998) 'Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males', *Journal of Political Economy* 106.2: 262–333.

Cameron, Stephen and James Heckman (2001) 'The Dynamics of Educational Attainment for Black, Hispanic and White Males', *Journal of Political Economy* 109.3: 455–499.

Cameron, Stephen and Christopher Taber (2004) 'Estimation of Educational Borrowing Constraints using Returns to Schooling', *Journal of Political Economy* 112.1: 132–182.

Chetty, Ray, John Friedman and Jonah Rockoff (2014) 'Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates', *American Economic Review* 104.9:2593–2632.

Chowdry, Haroon, Claire Crawford, Lorraine Dearden, Alisa Goodman and Anna Vignoles (2013) 'Widening Participation in Higher Education: Analysis using Linked Administrative Data', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176.2: 431–457.

Crawford, Claire and Ellen Greaves (2015) *Socio-economic, Ethnic and Gender Differences in HE Participation*. BIS Research Paper 186, London: UK Department for Business Innovation and Skills.

Cutler, David and Adriana Lleras-Muney (2008) *Education and Health: Evaluating Theories and Evidence*, New York: Russell Sage Foundation.

Dalton, Patricio, Sayantan Ghosal and Anandi Mani (2016) Poverty and Aspirations Failure, *The Economic Journal* 126.590: 165–188.

Deming, David (2014) 'Using School Choice Lotteries to Test Measures of School Effectiveness', *American Economic Review* 104.5: 406–11.

Deming, David, Justine Hastings, Thomas Kane and Douglas Staiger (2014) 'School Choice, School Quality, and Post-secondary Attainment', *American Economic Review* 104.3: 991–1013.

Department of Business, Innovation and Skills (2015) Participation Rates in Higher Education: Academic Years 2006/07 - 2013/14 (provisional), Statistical First Release (2 September 2015), London: UK Department of Business, Innovation and Skills.

Dercon, Stefan and Abhijeet Singh (2013) 'From Nutrition to Aspirations and Self-efficacy: Gender Bias over Time among Children in Four Countries', *World Development* 45: 31–50.

Escobal, Javier and Eva Flores (2008) *An Assessment of the Young Lives Sampling Approach in Peru*, Technical Note 3, Oxford: Young Lives.

Goldberg, Pinelopi and Nina Pavcnik (2007) 'Distributional Effects of Globalization in Developing Countries', *Journal of Economic Literature* 45.1: 39–82.

Grant, Monica and Jere Behrman (2010) 'Gender Gaps in Educational Attainment in Less-developed Countries', *Population and Development Review 36.1:* 71–89.

Hsieh, Chang-Tai and Peter Klenow (2009) 'Misallocation and Manufacturing TFP in China and India', *Quarterly Journal of Economics* 124.4: 1403–1448.

Kane, Thomas, Daniel McCaffrey, Trey Miller and Douglas Staiger (2013) *Have We Identified Effective Teachers? Validating Measures of Effective Teaching using Random Assignment,* MET Project Research Paper, Bill & Melinda Gates Foundation.

Kane, Thomas and Douglas Staiger (2008) *Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation,* NBER Working Papers 14607, Cambridge, MA: National Bureau of Economic Research, Inc.

Kapur, Devesh and Megan Crowley (2008) Beyond the ABCs: *Higher Education and Developing Countries*, Working Paper 139, Washington DC: Center for Global Development.

Keane, Michael and Kenneth Wolpin (2001) 'The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment', *International Economic Review* 42.4: 1051–1103.

Kumra, Neha (2008) *An Assessment of the Young Lives Sampling Approach in Andhra Pradesh, India*, Technical Note 22, Oxford: Young Lives.

Mookherjee, Dilip, Debraj Ray and Stefan Napel (2010) 'Aspirations, Segregation and Occupational Choice', *Journal of the European Economic Association* 8.1: 139–168.

Nguyen, Ngoc (2008) *An Assessment of the Young Lives Sampling Approach in Vietnam*, Technical Note 4, Oxford: Young Lives.

OECD (2013) *PISA 2012 Results: What Students Know and Can Do: Student Performance in Mathematics, Reading and Science*, Volume 1, Paris: OECD Publishing.

Oreopoulos, Philip and Kjell Salvanes (2011) 'Priceless: The Non-pecuniary Benefits of Schooling', *Journal of Economic Perspectives* 25.1: 159–184.

Pritchett, Lant (2013) *The Rebirth of Education: Schooling ain't Learning*, Washington, DC: Brookings Institution Press for Center for Global Development.

Rau, Tomas, Eugenio Rojas and Sergio Urzua (2013) 'Loans for Higher Education: Does the Dream Come True?', National Bureau of Economic Research Inc.

Ray, Debraj (2006) 'Aspirations, Poverty and Economic Change', chapter in A.V. Banerjee, R. Benabou and D. Mookherjee (eds) *Understanding Poverty*, Oxford: Oxford University Press.

Restuccia, Diego and Richard Rogerson (2013) 'Misallocation and Productivity', *Review of Economic Dynamics* 16.1:1–10.

Singh, Abhijeet (2014) 'Emergence and Evolution of Learning Gaps across Countries: Linked Panel Evidence from Ethiopia, India, Peru and Vietnam', CSAE Working Paper 2014-28, Oxford: Centre for the Study of African Economies, University of Oxford.

Singh, Abhijeet (2015) 'Private School Effects in Urban and Rural India: Panel Estimates at Primary and Secondary School Ages', *Journal of Development Economics* 113:16–32.

Soderbom, Mans, Francis Teal, Anthony Wambugu and Godius Kahyarara (2006) 'The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing', *Oxford Bulletin of Economics and Statistics* 68.3: 261–288.

Todd, Petra and Kenneth Wolpin (2003) 'On the Specification and Estimation of the Production Function for Cognitive Achievement', *The Economic Journal* 113.485: F3–F33.

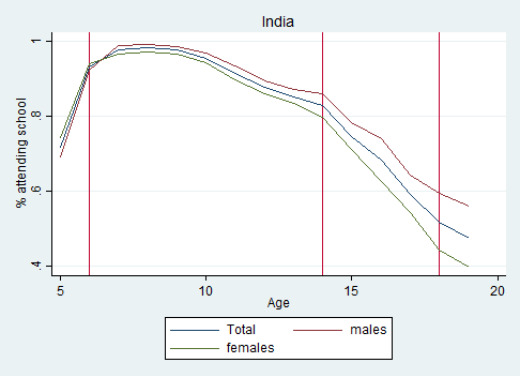
Todd, P.E. and K.I. Wolpin (2007) 'The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps', *Journal of Human Capital* 1.1: 91–136.

UNESCO (2015) *Education For All 2000-2015: Achievements and Challenges*, EFA Global Monitoring Report 2015, Paris: UNESCO.

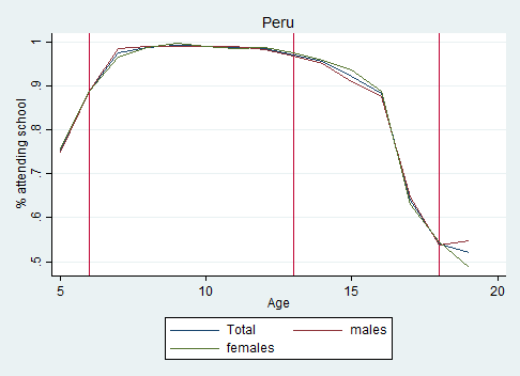
Walker, Maurice (2011) *Pisa 2009 Plus Results: Performance of 15-year-olds in Reading, Mathematics and Science for 10 Additional Participants,* Melbourne: ACER Press.

Figure 1: Enrolment in formal education by age (national data)

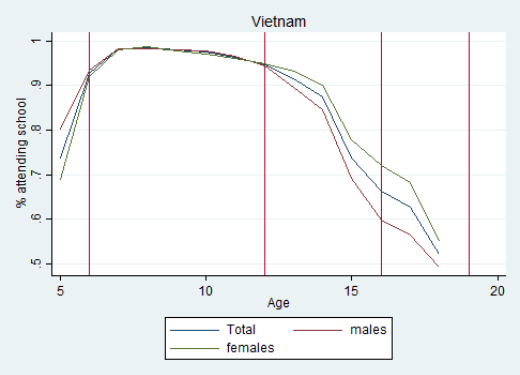
(a) India



(b) Peru



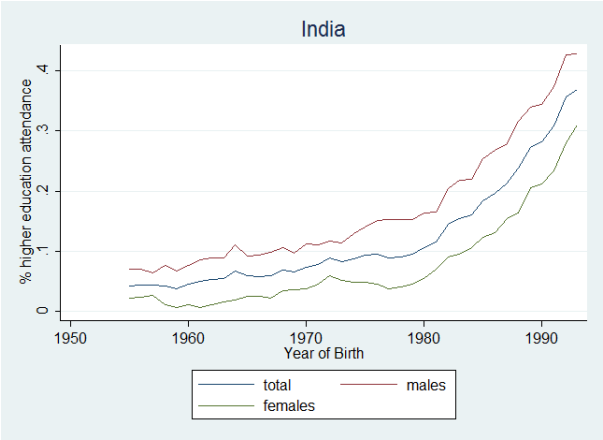
(c) Vietnam



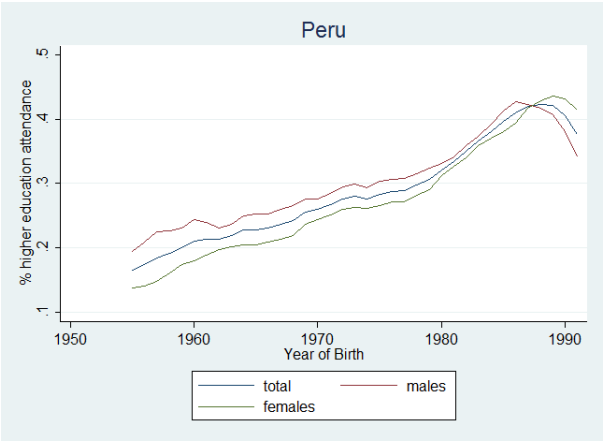
Note: Authors' computations based on India Human Development Survey - II (2012), the National Household Survey (2010) in Peru and the Multiple Indicator Cluster Survey (2010-11) in Vietnam. For each country, the information corresponds to all household members aged between 5 and 18 years at the time of the interview.

Figure 2: Trends in higher-education attendance

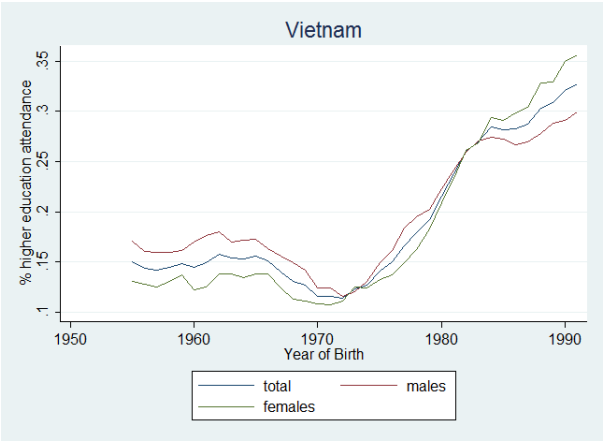
(a) India



(b) Peru



(c) Vietnam



Note: Authors' computations based on India Human Development Survey - II (2012), National Household Survey (2010) in Peru and the Multiple Indicator Cluster Survey (2010-11) in Vietnam. For each country, the information corresponds to all household members born in 1955 or later and aged 19 or more at the time of the interview. In the figures, each point represents a 5-year-of-birth moving average of higher-education attendance.

Table 1: Descriptive statistics

	India			Peru			Vietnam		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Household characteristics									
% Rural	0.76	0.43	950	0.25	0.43	622	0.82	0.39	879
Wealth Index	0.41	0.21	950	0.48	0.23	616	0.44	0.20	878
Mother's education level:									
— None	0.60	0.49	944	0.10	0.30	619	0.10	0.30	872
— Primary School	0.27	0.45	944	0.35	0.48	619	0.27	0.45	872
— Secondary School	0.10	0.30	944	0.40	0.49	619	0.68	0.47	872
— higher-education	0.03	0.16	944	0.16	0.36	619	0.05	0.22	872
Father's education level:									
— None	0.42	0.49	946	0.02	0.14	596	0.07	0.26	846
— Primary School	0.29	0.45	946	0.32	0.47	596	0.23	0.42	846
— Secondary School	0.22	0.41	946	0.47	0.50	596	0.71	0.45	846
— Higher Education	0.06	0.24	946	0.19	0.39	596	0.07	0.25	846
Birth Order:									
— First Child	0.28	0.45	950	0.32	0.47	621	0.37	0.48	879
— Second Child	0.33	0.47	950	0.25	0.43	621	0.32	0.47	879
— Third Child, +	0.39	0.49	950	0.43	0.50	621	0.32	0.47	879
Number of siblings	1.90	1.11	950	1.80	1.36	622	1.80	1.24	879
Individual characteristics									
Age in Round 4	18.72	0.46	950	18.41	0.57	622	18.76	0.47	878
Female	0.51	0.50	950	0.46	0.50	622	0.52	0.50	879
Height-for-age z score, (8 y)	-1.55	1.03	950	-1.41	1.01	618	-1.49	0.97	879
Aspirations at 12									
Caregiver's aspirations for child:									
— Complete Secondary or Less	0.30	0.46	915	0.06	0.23	618	0.22	0.41	871
— Higher Education	0.70	0.46	915	0.94	0.23	618	0.78	0.41	871
Child's aspirations:									
— Complete Secondary or Less	0.36	0.48	940	0.09	0.29	617	0.24	0.42	876
— Higher Education	0.64	0.48	940	0.91	0.29	617	0.76	0.42	876
Test scores at 12:									
PPVT (raw score, standardised)	0.00	1.00	931	-0.00	1.00	611	-0.00	1.00	838
Maths (raw score, standardised)	0.00	1.00	938	0.00	1.00	614	-0.00	1.00	874

Note: Data from the Young Lives surveys. Household characteristics were measured in Round 1 (2002), except for parental education, which comes from the Parental Background module of Round 2 (2006). The primary-school category in parental education includes adult literacy programmes. Aspiration and Test Scores measured at age 12 come from Round 2. Height-for-age at age 8 comes from Round 1.

Figure 3: Age of Young Lives sample individuals in successive survey rounds

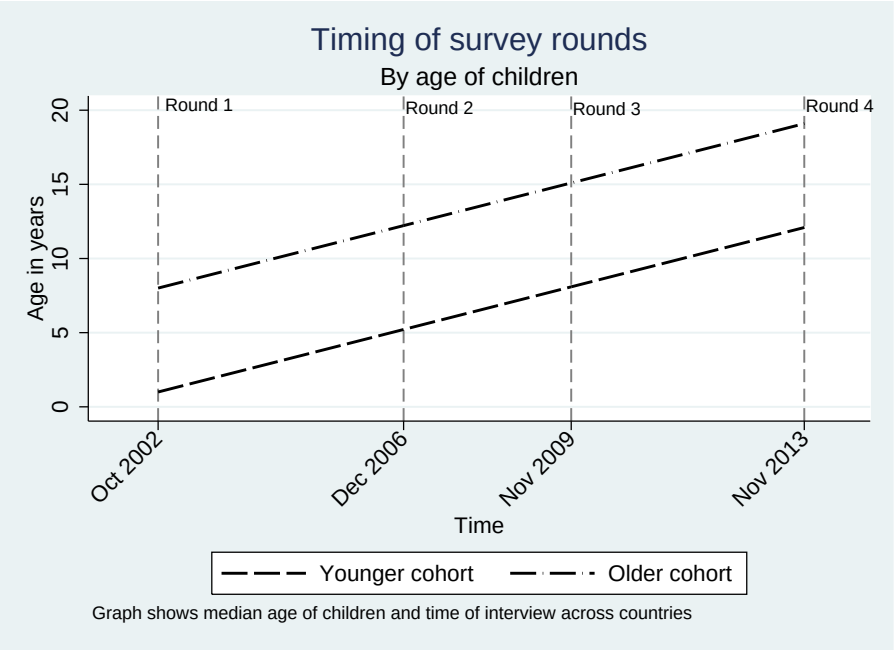


Table 2: Enrolment in higher-education at the age of 18-19

	India				Peru				Vietnam			
			%				%				%	
	N	Total	M	F	N	Total	M	F	N	Total	M	F
Never enrolled in HE	454	47.8	38.9	56.4	283	44.6	42.5	46.9	402	45.9	50.4	41.8
Enrolled in Secondary or lower	87	9.2	12.6	5.8	62	9.8	11.7	7.5	162	18.5	17.2	19.7
Ever enrolled in HE	409	43.1	48.5	37.8	290.0	45.7	45.8	45.6	312.0	35.6	32.5	38.5
— (a) Technical/vocational post secondary college	72	7.6	10.3	5.0	137	21.6	21.4	21.8	142	16.2	15.3	17.1
— (b) University	305	32.1	33.8	30.5	122	19.2	19.4	19.1	165	18.8	16.5	21.0
— (c) No longer enrolled	32	3.4	4.5	2.3	31	4.9	5.0	4.8	5	0.6	0.7	0.4
	950				635				876			

Note: Data from the Young Lives surveys. Enrollment was calculated using information from the Education History module from Round 4 for each country. The education history module contains information about education enrolment in primary, secondary and higher level, for each year between 2010 and 2013. An individual is reported as "ever enrolled in higher-education" if he/she was enroled in higher-education at least one year between 2010 and 2013. Rows (a) and (b) correspond to those enrrolled in 2013, the latest observation. Row (c) correspond to those that are not enrolled in 2013 but that were enrolled in higher-education at least one year between 2010 and 2012.

Table 3: higher-education enrolment by household characteristics and sex

		India		Peru		Vietnam	
		M	F	M	F	M	F
Location:							
Urban	%	0.54	0.59	0.50	0.53	0.43	0.49
	N	113	110	259	218	83	80
Rural	%	0.47	0.31	0.33	0.25	0.30	0.36
	N	355	372	82	76	335	376
Terciles of wealth:							
— Poorest third	%	0.34	0.20	0.30	0.23	0.13	0.21
	N	156	163	105	105	141	151
— Middle third	%	0.48	0.37	0.42	0.44	0.34	0.43
	N	154	162	120	91	127	162
— Richest third	%	0.63	0.57	0.64	0.71	0.50	0.52
	N	158	157	112	96	149	143
Mother's education level:							
— None	%	0.38	0.28	0.31	0.27	0.07	0.08
	N	282	286	35	26	46	40
— Primary	%	0.61	0.42	0.32	0.36	0.11	0.27
	N	129	130	115	104	110	124
— Secondary	%	0.67	0.74	0.48	0.49	0.44	0.47
	N	42	50	135	115	234	268
— higher-education	%	0.92	0.92	0.77	0.75	0.78	0.68
	N	13	12	53	44	23	22
Father's education level:							
— None	%	0.40	0.21	0.29	0.40	0.06	0.09
	N	205	196	7	5	36	23
— Primary	%	0.44	0.41	0.28	0.32	0.13	0.16
	N	131	143	99	92	89	101
— Secondary	%	0.58	0.50	0.48	0.48	0.42	0.47
	N	96	109	152	132	236	285
— higher-education	%	0.86	0.87	0.67	0.69	0.61	0.59
	N	35	31	69	48	38	34
Birth order:							
— First Child	%	0.55	0.45	0.59	0.51	0.38	0.40
	N	129	136	106	91	152	166
— Second Child	%	0.57	0.43	0.41	0.57	0.41	0.47
	N	155	157	85	76	138	139
— Third Child, +	%	0.37	0.29	0.39	0.35	0.17	0.28
	N	184	189	149	127	128	151

Note: Data from the Young Lives surveys. Area of location (urban and rural), wealth terciles and birth order are from Round 1 (2002); parental education is from Round 2. Enrolment is calculated using the definition "ever enrolled in higher-education" reported in Table 2.

Table 4: Aspirations for higher-education of the Young Lives cohorts at the age of 11-12

	India				Peru				Vietnam			
			%				%				%	
	n	T	M	F	n	T	M	F	n	T	M	F
Child aspirations												
— Secondary or lower	337	35.8	28.8	42.7	57	9.2	10.9	7.3	206	23.5	26.3	21.0
— higher-education	604	64.2	71.2	57.3	560	90.8	89.1	92.7	670	76.5	73.7	79.0
	941				617				876			
Caregiver aspirations												
— Secondary or lower	274	29.9	22.0	37.5	36	5.8	5.1	6.6	191	21.9	23.9	20.2
— higher-education	642	70.1	78.0	62.5	582	94.2	94.9	93.4	680	78.1	76.1	79.8
	916				618				871			

Note: Data from the Young Lives surveys. Aspirations were measured using the question "Imagine you had no constraints and could stay at school as long as you liked, what level of formal education would you like to complete?" asked to the child and "Ideally, what level of education would you like NAME to complete?" asked to the caregiver. Information is from Round 2. In the case of India, 104 children (10.9% of the sample) did not answer this question in Round 2 (compared to only 5 children in Peru and 3 in Vietnam). In the case of India only, we used the analagous aspiration question from Round 3.

Table 5: Factors affecting access to higher-education

	India				Peru				Vietnam			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Female	-0.111***	-0.058*	-0.055*	-0.062**	-0.004	-0.010	0.004	-0.002	0.068**	0.063**	0.063**	0.072**
	(0.031)	(0.031)	(0.030)	(0.029)	(0.040)	(0.040)	(0.039)	(0.040)	(0.032)	(0.032)	(0.031)	(0.031)
Rural (2002)	0.054	0.055	0.068		0.041	0.039	0.079		0.084*	0.079*	0.092**	
	(0.052)	(0.050)	(0.048)		(0.058)	(0.057)	(0.058)		(0.048)	(0.047)	(0.047)	
Wealth (2002)												
— Middle tercile	0.101***	0.087**	0.077**	0.054	0.038	0.027	0.002	0.023	0.101**	0.083*	0.068	0.092**
	(0.039)	(0.038)	(0.036)	(0.038)	(0.056)	(0.056)	(0.056)	(0.063)	(0.043)	(0.043)	(0.043)	(0.046)
— Top tercile	0.207***	0.188***	0.160***	0.165***	0.175***	0.161**	0.118*	0.167**	0.158***	0.129***	0.113**	0.124**
	(0.053)	(0.051)	(0.049)	(0.050)	(0.066)	(0.065)	(0.066)	(0.077)	(0.049)	(0.049)	(0.048)	(0.055)
Maternal Education												
— Primary School	0.107***	0.070*	0.039	0.054	-0.006	0.006	-0.001	0.038	0.038	-0.003	-0.094	-0.074
	(0.040)	(0.039)	(0.038)	(0.040)	(0.075)	(0.074)	(0.074)	(0.086)	(0.069)	(0.069)	(0.071)	(0.079)
— Secondary School	0.203***	0.132**	0.075	0.086	0.054	0.058	0.041	0.074	0.185***	0.112	0.016	0.010
	(0.064)	(0.062)	(0.061)	(0.062)	(0.083)	(0.082)	(0.081)	(0.092)	(0.071)	(0.072)	(0.075)	(0.081)
— higher-education	0.272**	0.227**	0.158	0.184*	0.184*	0.191*	0.155	0.207*	0.414***	0.344***	0.233*	0.226*
	(0.111)	(0.107)	(0.103)	(0.106)	(0.101)	(0.100)	(0.100)	(0.109)	(0.129)	(0.128)	(0.129)	(0.132)
Paternal Education												
— Primary School	0.033	0.014	-0.004	0.008	0.074	0.058	0.069	0.197	-0.083	-0.095	-0.105*	-0.089
	(0.040)	(0.038)	(0.037)	(0.036)	(0.162)	(0.161)	(0.159)	(0.163)	(0.064)	(0.063)	(0.062)	(0.065)
— Secondary School	0.028	-0.019	-0.019	0.002	0.087	0.051	0.043	0.157	0.076	0.032	0.005	0.016
	(0.049)	(0.048)	(0.046)	(0.046)	(0.163)	(0.162)	(0.161)	(0.164)	(0.061)	(0.060)	(0.060)	(0.062)
— higher-education	0.255***	0.183**	0.156**	0.149**	0.173	0.137	0.121	0.200	0.030	-0.024	-0.047	-0.070
	(0.080)	(0.077)	(0.074)	(0.074)	(0.171)	(0.169)	(0.168)	(0.171)	(0.109)	(0.108)	(0.107)	(0.107)
Birth Order												
— Second Child	0.024	0.032	0.045	0.024	-0.038	-0.041	-0.033	-0.051	0.030	0.028	0.021	0.019
	(0.040)	(0.039)	(0.038)	(0.038)	(0.053)	(0.053)	(0.052)	(0.053)	(0.039)	(0.038)	(0.038)	(0.038)
— Third Child +	-0.084*	-0.066	-0.047	-0.060	-0.053	-0.050	-0.023	-0.012	-0.103**	-0.104**	-0.113**	-0.104**
	(0.044)	(0.042)	(0.041)	(0.040)	(0.051)	(0.051)	(0.051)	(0.051)	(0.047)	(0.047)	(0.046)	(0.046)
Number of siblings	-0.018	-0.018	-0.014	0.003	-0.029*	-0.024	-0.022	-0.024	-0.000	0.006	0.016	0.018
	(0.017)	(0.016)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.018)	(0.017)	(0.017)	(0.018)
Age in whole years, R4	0.046	0.051	0.055*	0.042	0.012	0.016	0.010	0.008	0.169***	0.162***	0.154***	0.168***
	(0.035)	(0.033)	(0.032)	(0.033)	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)	(0.035)
Height-for-age z-score, R1	0.013	0.012	0.004	-0.002	0.072***	0.069***	0.062***	0.070***	0.029*	0.020	0.008	0.010
	(0.016)	(0.015)	(0.015)	(0.015)	(0.021)	(0.021)	(0.021)	(0.022)	(0.017)	(0.017)	(0.017)	(0.018)
Aspirations for higher-education:												
— Caregiver aspirations		0.162***	0.132***	0.134***		0.090	0.042	0.033		0.100**	0.075	0.083*
		(0.046)	(0.045)	(0.045)		(0.096)	(0.096)	(0.097)		(0.049)	(0.049)	(0.050)
— Child aspirations		0.162***	0.126***	0.102**		0.215***	0.158**	0.133		0.152***	0.107**	0.098**
		(0.044)	(0.042)	(0.042)		(0.079)	(0.080)	(0.081)		(0.048)	(0.049)	(0.049)
Test scores (2006)												
— Receptive vocabulary			0.097***	0.104***			0.043	0.062**			0.041*	0.057**
			(0.019)	(0.020)			(0.027)	(0.029)			(0.024)	(0.026)
— Mathematics			0.058***	0.070***			0.065**	0.070***			0.066***	0.071***
			(0.019)	(0.019)			(0.026)	(0.026)			(0.022)	(0.022)
Cluster Fixed Effects	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Constant	-0.510	-0.834	-0.859	-0.597	0.194	-0.146	0.065	-0.001	-3.095***	-3.060***	-2.786***	-3.010***
	(0.650)	(0.627)	(0.604)	(0.611)	(0.667)	(0.669)	(0.665)	(0.670)	(0.656)	(0.646)	(0.642)	(0.661)
Number of observations	878	878	878	878	559	559	559	559	777	777	777	777
R2	0.154	0.217	0.276	0.324	0.156	0.175	0.195	0.237	0.207	0.233	0.253	0.294

Notes: Each column report results from a linear probability model. The dependent variable takes the value of 1 if the individual is/was enrolled in higher-education, 0 otherwise. Clustered standard errors reported in brackets. ***p<0.01, **p<0.05, *p<0.1.

Table 6: Main results split by gender

	India				Peru				Vietnam			
	Male		Female		Male		Female		Male		Female	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Rural (2002)	0.237***	0.202***	-0.085	-0.041	0.153*	0.157*	-0.062	0.004	0.090	0.101	0.078	0.086
	(0.080)	(0.076)	(0.067)	(0.062)	(0.078)	(0.082)	(0.087)	(0.087)	(0.068)	(0.067)	(0.069)	(0.068)
Wealth (2002)												
— Middle tercile	0.106*	0.082	0.104*	0.070	0.004	-0.026	0.072	0.043	0.084	0.052	0.111*	0.081
	(0.056)	(0.053)	(0.054)	(0.050)	(0.074)	(0.074)	(0.088)	(0.086)	(0.063)	(0.063)	(0.061)	(0.059)
— Top tercile	0.313***	0.234***	0.137*	0.101	0.151*	0.103	0.217**	0.184*	0.179**	0.135*	0.135*	0.096
	(0.080)	(0.076)	(0.070)	(0.065)	(0.087)	(0.089)	(0.105)	(0.103)	(0.071)	(0.071)	(0.069)	(0.068)
Maternal Education												
— Primary School	0.219***	0.158***	-0.004	-0.083	0.008	0.023	0.007	-0.036	-0.038	-0.173	0.105	-0.039
	(0.059)	(0.056)	(0.056)	(0.053)	(0.103)	(0.102)	(0.112)	(0.110)	(0.103)	(0.107)	(0.096)	(0.098)
— Secondary School	0.169*	0.087	0.245***	0.063	0.103	0.101	0.013	-0.061	0.193*	0.027	0.174*	-0.014
	(0.095)	(0.090)	(0.087)	(0.082)	(0.115)	(0.114)	(0.122)	(0.120)	(0.106)	(0.112)	(0.098)	(0.103)
— higher-education	0.321**	0.204	0.280*	0.143	0.243*	0.252*	0.114	-0.012	0.453**	0.278	0.391**	0.187
	(0.162)	(0.153)	(0.152)	(0.140)	(0.138)	(0.136)	(0.153)	(0.152)	(0.176)	(0.178)	(0.194)	(0.193)
Paternal Education												
— Primary School	-0.068	-0.085	0.120**	0.076	0.111	0.132	0.027	-0.016	-0.069	-0.065	-0.114	-0.168*
	(0.057)	(0.054)	(0.055)	(0.051)	(0.202)	(0.201)	(0.280)	(0.275)	(0.086)	(0.085)	(0.098)	(0.096)
— Secondary School	-0.063	-0.085	0.092	0.045	0.135	0.123	0.032	-0.051	0.077	0.035	0.063	-0.045
	(0.073)	(0.070)	(0.066)	(0.061)	(0.205)	(0.204)	(0.279)	(0.276)	(0.083)	(0.082)	(0.091)	(0.090)
— higher-education	0.128	0.058	0.375***	0.276***	0.242	0.222	0.102	-0.001	0.080	0.024	-0.035	-0.135
	(0.118)	(0.112)	(0.107)	(0.099)	(0.215)	(0.213)	(0.292)	(0.289)	(0.145)	(0.143)	(0.166)	(0.162)
Birth Order												
— Second Child	0.079	0.103*	-0.025	-0.016	-0.158**	-0.150**	0.092	0.089	0.028	0.024	0.036	0.028
	(0.060)	(0.057)	(0.055)	(0.050)	(0.073)	(0.072)	(0.079)	(0.077)	(0.055)	(0.054)	(0.057)	(0.056)
— Third Child +	-0.034	0.023	-0.117**	-0.095*	-0.043	-0.037	-0.057	-0.001	-0.132*	-0.146**	-0.069	-0.069
	(0.068)	(0.064)	(0.057)	(0.053)	(0.072)	(0.072)	(0.074)	(0.074)	(0.069)	(0.069)	(0.066)	(0.064)
Number of siblings	-0.041*	-0.040*	-0.000	0.003	-0.037*	-0.027	-0.015	-0.015	0.000	0.022	-0.005	0.006
	(0.025)	(0.024)	(0.022)	(0.020)	(0.021)	(0.022)	(0.025)	(0.024)	(0.025)	(0.026)	(0.025)	(0.024)
Age in whole years, R4	0.065	0.091**	-0.015	-0.026	0.036	0.044	-0.026	-0.050	0.126**	0.113**	0.201***	0.186***
	(0.048)	(0.045)	(0.050)	(0.046)	(0.046)	(0.045)	(0.052)	(0.051)	(0.050)	(0.049)	(0.049)	(0.047)
Height-for-age z-score, R1	0.043*	0.031	-0.008	-0.015	0.085***	0.080***	0.065*	0.041	0.010	-0.010	0.044*	0.025
	(0.022)	(0.021)	(0.022)	(0.021)	(0.028)	(0.028)	(0.035)	(0.034)	(0.025)	(0.025)	(0.025)	(0.024)
Aspirations for higher-education:												
— Caregiver aspirations		0.128*		0.171***		0.147		-0.068		0.097		0.040
		(0.074)		(0.055)		(0.140)		(0.138)		(0.067)		(0.075)
— Child aspirations		0.068		0.133**		0.152		0.139		0.032		0.188**
		(0.070)		(0.053)		(0.104)		(0.132)		(0.064)		(0.076)
Test scores (2006)												
— Receptive vocabulary		0.076***		0.122***		0.004		0.084**		0.045		0.045
		(0.028)		(0.025)		(0.039)		(0.040)		(0.037)		(0.031)
— Mathematics		0.091***		0.016		0.053		0.084**		0.067**		0.060**
		(0.029)		(0.025)		(0.037)		(0.036)		(0.033)		(0.030)
Constant	-0.965	-1.574*	0.583	0.643	-0.269	-0.716	0.908	1.368	-2.315**	-2.031**	-3.587***	-3.276***
	(0.896)	(0.850)	(0.942)	(0.863)	(0.889)	(0.898)	(1.018)	(1.008)	(0.946)	(0.932)	(0.920)	(0.897)
Number of observations	430	430	448	448	299	299	260	260	362	362	415	415
R2	0.166	0.271	0.187	0.325	0.184	0.215	0.175	0.235	0.255	0.291	0.179	0.239

Notes: Each column report results from a linear probability model. The dependent variable takes the value of 1 if the individual is/was enrolled in higher-education, 0 otherwise. Clustered standard errors reported in brackets. ***p<0.01, **p<0.05, *p<0.1.

Table 7: Results split by area of location

	India				Peru				Vietnam			
	Urban		Rural		Urban		Rural		Urban		Rural	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Female	0.069	0.094	-0.161***	-0.098***	0.041	0.043	-0.141*	-0.130	0.108	0.092	0.076**	0.072**
	(0.068)	(0.061)	(0.036)	(0.034)	(0.046)	(0.046)	(0.080)	(0.081)	(0.087)	(0.083)	(0.035)	(0.034)
Wealth (2002)												
— Middle tercile	0.380	0.300	0.097**	0.077**	0.087	0.026	-0.088	-0.088	0.022	-0.039	0.104**	0.077*
	(0.278)	(0.246)	(0.040)	(0.037)	(0.069)	(0.069)	(0.095)	(0.094)	(0.273)	(0.269)	(0.044)	(0.043)
— Top tercile	0.482**	0.376*	0.195***	0.154***	0.214***	0.130*	0.518	0.446	0.300	0.142	0.134***	0.101**
	(0.242)	(0.218)	(0.057)	(0.054)	(0.074)	(0.075)	(0.472)	(0.470)	(0.252)	(0.252)	(0.051)	(0.050)
Maternal Education												
— Primary School	0.138	0.018	0.099**	0.039	0.085	0.111	-0.007	-0.008	0.590**	0.453*	-0.022	-0.145*
	(0.089)	(0.081)	(0.046)	(0.043)	(0.118)	(0.116)	(0.098)	(0.099)	(0.251)	(0.249)	(0.073)	(0.075)
— Secondary School	0.263**	-0.015	0.136	0.028	0.082	0.096	0.351**	0.354**	0.676***	0.441*	0.134*	-0.026
	(0.106)	(0.101)	(0.087)	(0.082)	(0.120)	(0.118)	(0.144)	(0.145)	(0.245)	(0.247)	(0.076)	(0.080)
— higher-education	0.292*	0.016	0.413*	0.324	0.206	0.199	(dropped)	(dropped)	0.935***	0.659**	0.326**	0.150
	(0.150)	(0.138)	(0.216)	(0.202)	(0.133)	(0.131)			(0.305)	(0.305)	(0.160)	(0.160)
Paternal Education												
— Primary School	-0.089	-0.060	0.049	0.006	0.096	0.146	0.097	0.052	-0.110	-0.068	-0.062	-0.087
	(0.103)	(0.091)	(0.043)	(0.040)	(0.216)	(0.213)	(0.239)	(0.243)	(0.216)	(0.209)	(0.067)	(0.066)
— Secondary School	-0.055	-0.018	0.026	-0.033	0.083	0.085	0.101	0.033	-0.059	-0.073	0.105	0.036
	(0.106)	(0.094)	(0.058)	(0.054)	(0.214)	(0.210)	(0.250)	(0.254)	(0.174)	(0.167)	(0.065)	(0.065)
— higher-education	0.093	0.039	0.348***	0.231**	0.198	0.194	-0.121	-0.250	-0.201	-0.248	0.150	0.073
	(0.136)	(0.122)	(0.106)	(0.100)	(0.219)	(0.215)	(0.338)	(0.342)	(0.223)	(0.217)	(0.141)	(0.139)
Birth Order												
— Second Child	0.053	0.035	0.012	0.042	-0.062	-0.053	0.158	0.167	-0.121	-0.091	0.061	0.052
	(0.082)	(0.074)	(0.047)	(0.044)	(0.059)	(0.058)	(0.122)	(0.122)	(0.103)	(0.101)	(0.043)	(0.042)
— Third Child +	-0.088	-0.088	-0.083*	-0.038	-0.067	-0.030	0.097	0.122	-0.130	-0.017	-0.096*	-0.113**
	(0.092)	(0.083)	(0.050)	(0.047)	(0.059)	(0.059)	(0.110)	(0.110)	(0.141)	(0.139)	(0.050)	(0.049)
Number of siblings	0.004	0.043	-0.026	-0.026	-0.015	-0.010	-0.055**	-0.050*	0.009	0.006	-0.005	0.012
	(0.037)	(0.034)	(0.019)	(0.018)	(0.020)	(0.019)	(0.028)	(0.028)	(0.054)	(0.052)	(0.019)	(0.018)
Age in whole years, R4	0.024	-0.002	0.045	0.059	-0.032	-0.026	0.141*	0.114	0.213*	0.199*	0.173***	0.158***
	(0.070)	(0.062)	(0.040)	(0.038)	(0.039)	(0.039)	(0.072)	(0.074)	(0.112)	(0.107)	(0.036)	(0.035)
Height-for-age z-score, R1	0.044	0.078**	0.010	-0.005	0.088***	0.082***	-0.007	-0.028	0.013	-0.003	0.032*	0.011
	(0.037)	(0.033)	(0.018)	(0.016)	(0.024)	(0.024)	(0.046)	(0.047)	(0.048)	(0.046)	(0.019)	(0.019)
Aspirations for higher-education:												
— Caregiver aspirations		0.175*		0.136***		0.115		-0.049		0.190		0.077
		(0.100)		(0.050)		(0.140)		(0.132)		(0.189)		(0.051)
— Child aspirations		0.223**		0.107**		0.156		0.146		0.191		0.087*
		(0.087)		(0.048)		(0.098)		(0.139)		(0.159)		(0.051)
Test scores (2006)												
— Receptive vocabulary		0.207***		0.080***		0.055*		-0.037		0.135		0.046*
		(0.045)		(0.021)		(0.032)		(0.058)		(0.135)		(0.024)
— Mathematics		0.083*		0.056***		0.062*		0.086*		0.143**		0.051**
		(0.047)		(0.021)		(0.032)		(0.045)		(0.071)		(0.023)
Constant	-0.420	-0.176	-0.398	-0.843	0.915	0.518	-2.317*	-1.899	-4.363*	-4.229**	-3.060***	-2.756***
	(1.321)	(1.175)	(0.757)	(0.708)	(0.768)	(0.778)	(1.368)	(1.419)	(2.240)	(2.137)	(0.678)	(0.667)
Number of observations	205	205	673	673	426	426	133	133	146	146	631	631
R2	0.163	0.360	0.141	0.256	0.164	0.204	0.195	0.237	0.152	0.255	0.232	0.273

Notes: Each column report results from a linear probability model. The dependent variable takes the value of 1 if the individual is/was enrolled in higher-education, 0 otherwise. Clustered standard errors reported in brackets. ***p<0.01, **p<0.05, *p<0.1.

Table 8: Results split by education

	India				Peru				Vietnam			
	Less than sec		Sec or more		Less than sec		Sec or more		Less than sec		Sec or more	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Female	-0.136***	-0.075**	-0.053	-0.008	0.010	-0.003	-0.010	0.008	0.086**	0.090**	0.072*	0.060
	(0.038)	(0.035)	(0.060)	(0.057)	(0.074)	(0.075)	(0.046)	(0.045)	(0.044)	(0.042)	(0.038)	(0.037)
Rural (2002)	0.086	0.116*	-0.013	0.006	-0.015	0.019	0.057	0.103	-0.099	-0.080	0.093*	0.099*
	(0.072)	(0.066)	(0.078)	(0.071)	(0.079)	(0.080)	(0.077)	(0.076)	(0.096)	(0.092)	(0.052)	(0.051)
Wealth (2002)												
— Middle tercile	0.115***	0.079**	0.083	0.094	-0.012	-0.045	0.156**	0.079	0.147***	0.075	0.155***	0.103**
	(0.043)	(0.039)	(0.108)	(0.098)	(0.088)	(0.087)	(0.070)	(0.070)	(0.056)	(0.056)	(0.051)	(0.050)
— Top tercile	0.273***	0.212***	0.211*	0.138	0.202	0.160	0.338***	0.216***	-0.001	-0.060	0.238***	0.152***
	(0.064)	(0.059)	(0.112)	(0.101)	(0.151)	(0.150)	(0.072)	(0.076)	(0.085)	(0.082)	(0.055)	(0.055)
Birth Order												
— Second Child	-0.014	0.009	0.049	0.087	0.150	0.165	-0.076	-0.060	0.000	-0.036	0.041	0.047
	(0.051)	(0.047)	(0.070)	(0.065)	(0.123)	(0.121)	(0.058)	(0.057)	(0.063)	(0.060)	(0.045)	(0.044)
— Third Child +	-0.130**	-0.083*	-0.045	-0.013	0.094	0.140	-0.118**	-0.072	-0.068	-0.082	-0.099	-0.089
	(0.051)	(0.048)	(0.089)	(0.083)	(0.102)	(0.102)	(0.056)	(0.056)	(0.057)	(0.055)	(0.062)	(0.060)
Number of siblings	-0.025	-0.023	-0.018	0.010	-0.074***	-0.067***	0.003	0.005	-0.004	0.014	-0.007	0.007
	(0.019)	(0.017)	(0.036)	(0.033)	(0.023)	(0.023)	(0.021)	(0.020)	(0.016)	(0.016)	(0.025)	(0.024)
Age in whole years, R4	0.043	0.061	0.033	0.040	0.016	0.011	-0.000	0.005	0.024	0.021	0.219***	0.192***
	(0.042)	(0.039)	(0.064)	(0.058)	(0.063)	(0.065)	(0.040)	(0.039)	(0.044)	(0.042)	(0.043)	(0.042)
Height-for-age z-score, R1	0.018	0.006	0.026	0.010	0.013	0.003	0.084***	0.071***	0.018	-0.008	0.049**	0.022
	(0.019)	(0.017)	(0.030)	(0.028)	(0.041)	(0.040)	(0.024)	(0.024)	(0.021)	(0.021)	(0.021)	(0.021)
Aspirations for higher-education:												
— Caregiver aspirations		0.140***		0.142		-0.043		0.110		0.119**		0.103*
		(0.048)		(0.120)		(0.121)		(0.156)		(0.052)		(0.062)
— Child aspirations		0.124***		0.211**		0.199*		0.073		0.041		0.146**
		(0.047)		(0.101)		(0.110)		(0.101)		(0.052)		(0.062)
Test scores (2006)												
— Receptive vocabulary		0.076***		0.172***		-0.000		0.067**		0.038*		0.076**
		(0.021)		(0.040)		(0.051)		(0.032)		(0.023)		(0.033)
— Mathematics		0.064***		0.057		0.072*		0.087***		0.008		0.105***
		(0.021)		(0.048)		(0.041)		(0.032)		(0.023)		(0.029)
_cons	-0.393	-0.962	-0.056	-0.670	0.110	0.043	0.477	0.195	-0.260	-0.318	-3.839***	-3.616***
	(0.782)	(0.722)	(1.206)	(1.101)	(1.174)	(1.226)	(0.741)	(0.738)	(0.827)	(0.792)	(0.813)	(0.789)
Number of observations	616	616	269	269	159	159	428	428	178	178	635	635
R2	0.089	0.236	0.049	0.237	0.106	0.162	0.121	0.170	0.099	0.201	0.106	0.178

Notes: Each column report results from a linear probability model. The dependent variable takes the value of 1 if the individual is/was enrolled in higher-education, 0 otherwise. Clustered standard errors reported in brackets. ***p<0.01, **p<0.05, *p<0.1.

APPENDIX

Table 9: Factors affecting access to higher-education - Robustness check (upper bound of higher-education)

	India				Peru				Vietnam			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Female	-0.111***	-0.058*	-0.055*	-0.062**	-0.004	-0.010	0.004	-0.002	0.068**	0.063**	0.063**	0.072**
	(0.031)	(0.031)	(0.030)	(0.029)	(0.040)	(0.040)	(0.039)	(0.040)	(0.032)	(0.032)	(0.031)	(0.031)
Rural (2002)	0.054	0.055	0.068		0.041	0.039	0.079		0.084*	0.079*	0.092**	
	(0.052)	(0.050)	(0.048)		(0.058)	(0.057)	(0.058)		(0.048)	(0.047)	(0.047)	
Wealth (2002)												
— Middle tercile	0.101***	0.087**	0.077**	0.054	0.038	0.027	0.002	0.023	0.101**	0.083*	0.068	0.092**
	(0.039)	(0.038)	(0.036)	(0.038)	(0.056)	(0.056)	(0.056)	(0.063)	(0.043)	(0.043)	(0.043)	(0.046)
— Top tercile	0.207***	0.188***	0.160***	0.165***	0.175***	0.161**	0.118*	0.167**	0.158***	0.129***	0.113**	0.124**
	(0.053)	(0.051)	(0.049)	(0.050)	(0.066)	(0.065)	(0.066)	(0.077)	(0.049)	(0.049)	(0.048)	(0.055)
Maternal Education												
— Primary School	0.107***	0.070*	0.039	0.054	-0.006	0.006	-0.001	0.038	0.038	-0.003	-0.094	-0.074
	(0.040)	(0.039)	(0.038)	(0.040)	(0.075)	(0.074)	(0.074)	(0.086)	(0.069)	(0.069)	(0.071)	(0.079)
— Secondary School	0.203***	0.132**	0.075	0.086	0.054	0.058	0.041	0.074	0.185***	0.112	0.016	0.010
	(0.064)	(0.062)	(0.061)	(0.062)	(0.083)	(0.082)	(0.081)	(0.092)	(0.071)	(0.072)	(0.075)	(0.081)
— higher-education	0.272**	0.227**	0.158	0.184*	0.184*	0.191*	0.155	0.207*	0.414***	0.344***	0.233*	0.226*
	(0.111)	(0.107)	(0.103)	(0.106)	(0.101)	(0.100)	(0.100)	(0.109)	(0.129)	(0.128)	(0.129)	(0.132)
Paternal Education												
— Primary School	0.033	0.014	-0.004	0.008	0.074	0.058	0.069	0.197	-0.083	-0.095	-0.105*	-0.089
	(0.040)	(0.038)	(0.037)	(0.036)	(0.162)	(0.161)	(0.159)	(0.163)	(0.064)	(0.063)	(0.062)	(0.065)
— Secondary School	0.028	-0.019	-0.019	0.002	0.087	0.051	0.043	0.157	0.076	0.032	0.005	0.016
	(0.049)	(0.048)	(0.046)	(0.046)	(0.163)	(0.162)	(0.161)	(0.164)	(0.061)	(0.060)	(0.060)	(0.062)
— higher-education	0.255***	0.183**	0.156**	0.149**	0.173	0.137	0.121	0.200	0.030	-0.024	-0.047	-0.070
	(0.080)	(0.077)	(0.074)	(0.074)	(0.171)	(0.169)	(0.168)	(0.171)	(0.109)	(0.108)	(0.107)	(0.107)
Birth Order												
— Second Child	0.024	0.032	0.045	0.024	-0.038	-0.041	-0.033	-0.051	0.030	0.028	0.021	0.019
	(0.040)	(0.039)	(0.038)	(0.038)	(0.053)	(0.053)	(0.052)	(0.053)	(0.039)	(0.038)	(0.038)	(0.038)
— Third Child +	-0.084*	-0.066	-0.047	-0.060	-0.053	-0.050	-0.023	-0.012	-0.103**	-0.104**	-0.113**	-0.104**
	(0.044)	(0.042)	(0.041)	(0.040)	(0.051)	(0.051)	(0.051)	(0.051)	(0.047)	(0.047)	(0.046)	(0.046)
Number of siblings	-0.018	-0.018	-0.014	0.003	-0.029*	-0.024	-0.022	-0.024	-0.000	0.006	0.016	0.018
	(0.017)	(0.016)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.018)	(0.017)	(0.017)	(0.018)
Age in whole years, R4	0.046	0.051	0.055*	0.042	0.012	0.016	0.010	0.008	0.169***	0.162***	0.154***	0.168***
	(0.035)	(0.033)	(0.032)	(0.033)	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)	(0.035)
Height-for-age z-score, R1	0.013	0.012	0.004	-0.002	0.072***	0.069***	0.062***	0.070***	0.029*	0.020	0.008	0.010
	(0.016)	(0.015)	(0.015)	(0.015)	(0.021)	(0.021)	(0.021)	(0.022)	(0.017)	(0.017)	(0.017)	(0.018)
Aspirations for higher-education:												
— Caregiver aspirations		0.162***	0.132***	0.134***		0.090	0.042	0.033		0.100**	0.075	0.083*
		(0.046)	(0.045)	(0.045)		(0.096)	(0.096)	(0.097)		(0.049)	(0.049)	(0.050)
— Child aspirations		0.162***	0.126***	0.102**		0.215***	0.158**	0.133		0.152***	0.107**	0.098**
		(0.044)	(0.042)	(0.042)		(0.079)	(0.080)	(0.081)		(0.048)	(0.049)	(0.049)
Test scores (2006)												
— Receptive vocabulary			0.097***	0.104***			0.043	0.062**			0.041*	0.057**
			(0.019)	(0.020)			(0.027)	(0.029)			(0.024)	(0.026)
— Mathematics			0.058***	0.070***			0.065**	0.070***			0.066***	0.071***
			(0.019)	(0.019)			(0.026)	(0.026)			(0.022)	(0.022)
Cluster Fixed Effects	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Constant	-0.510	-0.834	-0.859	-0.597	0.194	-0.146	0.065	-0.001	-3.095***	-3.060***	-2.786***	-3.010***
	(0.650)	(0.627)	(0.604)	(0.611)	(0.667)	(0.669)	(0.665)	(0.670)	(0.656)	(0.646)	(0.642)	(0.661)
Number of observations	878	878	878	878	559	559	559	559	777	777	777	777
R2	0.154	0.217	0.276	0.324	0.156	0.175	0.195	0.237	0.207	0.233	0.253	0.294

Notes: Each column report results from a linear probability model. The dependent variable takes the value of 1 if the individual is/was enrolled in higher-education or is enrolled in secondary education, 0 otherwise. Clustered standard errors reported in brackets. ***p<0.01, **p<0.05, *p<0.1.

APPENDIX

Table 10: Factors affecting access to higher-education - Robustness check (logit estimation)

	India				Peru				Vietnam			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Female	-0.111***	-0.059*	-0.054*	-0.063**	-0.002	-0.009	0.009	0.003	0.078**	0.072*	0.074**	0.086**
	(0.031)	(0.030)	(0.029)	(0.028)	(0.040)	(0.039)	(0.039)	(0.038)	(0.038)	(0.037)	(0.035)	(0.036)
Rural (2002)	0.045	0.046	0.056		0.041	0.038	0.077	0.009	0.080*	0.075	0.095**	
	(0.049)	(0.047)	(0.044)		(0.056)	(0.055)	(0.054)	(0.081)	(0.049)	(0.047)	(0.042)	
Wealth (2002)												
— Middle tercile	0.100***	0.087**	0.077**	0.052	0.036	0.027	0.000	0.021	0.128**	0.112**	0.098*	0.129**
	(0.038)	(0.036)	(0.035)	(0.036)	(0.054)	(0.054)	(0.055)	(0.060)	(0.055)	(0.550)	(0.053)	(0.060)
— Top tercile	0.200***	0.183***	0.152***	0.152***	0.171***	0.154**	0.106	0.164**	0.178***	0.139**	0.119**	0.132*
	(0.054)	(0.052)	(0.049)	(0.050)	(0.069)	(0.068)	(0.068)	(0.077)	(0.060)	(0.060)	(0.059)	(0.069)
Maternal Education												
— Primary School	0.097***	0.062*	0.032	0.047	-0.005	-0.005	-0.016	0.017	0.163	0.080	-0.054	-0.055
	(0.039)	(0.038)	(0.036)	(0.037)	(0.077)	(0.076)	(0.076)	(0.084)	(0.123)	(0.122)	(0.108)	(0.114)
— Secondary School	0.194***	0.109*	0.047	0.053	0.049	0.040	0.019	0.047	0.299***	0.191*	0.046	0.018
	(0.065)	(0.062)	(0.059)	(0.059)	(0.083)	(0.082)	(0.082)	(0.091)	(0.096)	(0.103)	(0.109)	(0.117)
— higher-education	0.357***	0.302**	0.231*	0.251*	0.187	0.178*	0.134	0.184*	0.532***	0.455***	0.259	0.257
	(0.120)	(0.125)	(0.132)	(0.122)	(0.106)	(0.103)	(0.104)	(0.109)	(0.107)	(0.148)	(0.197)	(0.210)
Paternal Education												
— Primary School	0.032	0.012	0.001	0.018	0.078	0.064	0.078	0.189	-0.116	-0.115	-0.101	-0.086
	(0.039)	(0.037)	(0.035)	(0.034)	(0.166)	(0.168)	(0.164)	(0.131)	(0.078)	(0.077)	(0.074)	(0.076)
— Secondary School	0.024	-0.022	-0.018	0.004	0.093	0.056	0.049	0.156	0.085	0.047	0.033	0.380
	(0.048)	(0.045)	(0.043)	(0.043)	(0.175)	(0.175)	(0.174)	(0.159)	(0.074)	(0.075)	(0.072)	(0.074)
— higher-education	0.311***	0.216**	0.179*	0.162**	0.187	0.146	0.128	0.213	0.035	-0.022	-0.039	-0.075
	(0.083)	(0.087)	(0.083)	(0.080)	(0.189)	(0.187)	(0.186)	(0.174)	(0.128)	(0.115)	(0.106)	(0.096)
Birth Order												
— Second Child	0.022	0.030	0.047	0.024	-0.036	-0.037	-0.031	-0.049	0.028	0.029	0.033	0.038
	(0.040)	(0.038)	(0.037)	(0.036)	(0.051)	(0.050)	(0.050)	(0.049)	(0.046)	(0.045)	(0.072)	(0.044)
— Third Child +	-0.083*	-0.068	-0.048	-0.063	-0.053	-0.052	-0.023	-0.011	-0.116**	-0.123**	-0.114**	-0.123**
	(0.044)	(0.042)	(0.040)	(0.039)	(0.050)	(0.050)	(0.049)	(0.049)	(0.055)	(0.053)	(0.050)	(0.050)
Number of siblings	-0.019	-0.017	-0.012	0.006	-0.029*	-0.026	-0.022	-0.025	-0.006	0.005	0.013	0.021
	(0.017)	(0.016)	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.024)	(0.023)	(0.023)	(0.023)
Age in whole years, R4	0.046	0.051	0.05	0.040	0.012	0.017	0.013	0.009	0.214***	0.206***	0.190***	0.217***
	(0.034)	(0.032)	(0.031)	(0.030)	(0.034)	(0.034)	(0.034)	(0.033)	(0.045)	(0.044)	(0.042)	(0.044)
Height-for-age z-score, R1	0.013	0.011	0.002	-0.005	0.072***	0.068***	0.063***	0.069***	0.034	0.022	-0.004	-0.002
	(0.015)	(0.015)	(0.014)	(0.014)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Aspirations for higher-education:												
— Caregiver aspirations		0.176***	0.143***	0.146***		0.144	0.042	0.104		0.160***	0.127**	0.140***
		(0.047)	(0.046)	(0.045)		(0.111)	(0.096)	(0.111)		(0.055)	(0.053)	(0.052)
— Child aspirations		0.159***	0.124***	0.099**		0.262***	0.158**	0.213***		0.218***	0.169***	0.160***
		(0.044)	(0.042)	(0.041)		(0.075)	(0.080)	(0.083)		(0.049)	(0.050)	(0.050)
Test scores (2006)												
— Receptive vocabulary			0.095***	0.102***			0.043	0.065**			0.119***	0.155***
			(0.019)	(0.020)			(0.027)	(0.028)			(0.035)	(0.040)
— Mathematics			0.074***	0.091***			0.074***	0.077***			0.112***	0.132***
			(0.021)	(0.021)			(0.026)	(0.026)			(0.031)	(0.031)
Cluster Fixed Effects	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Number of observations	878	878	878	878	559	559	559	559	777	777	777	777
Pseudo R2	0.122	0.177	0.239	0.292	0.121	0.141	0.161	0.201	0.178	0.212	0.252	0.297

Notes: In each column, we report the marginal effects for the average individual obtained from a Logit estimation. The dependent variable takes the value of 1 if the individual is/was enrolled in higher-education, 0 otherwise. Clustered standard errors reported in brackets. ***p<0.01, **p<0.05, *p<0.1.

Accessing Higher Education in Developing Countries: Panel Data Analysis from India, Peru, and Vietnam

We use unique individual-level panel data from India (Andhra Pradesh), Peru and Vietnam on a cohort of individuals surveyed from the age of 8 years to 19 years to study factors affecting enrolment in higher education in these middle-income countries. We document (a) that similar to nationally representative data, the proportion having accessed higher education at this age is high (~35-45 per cent); (b) that there are steep gradients in higher education access across wealth and parental education; (c) that a substantial part of the gradient with regard to parental education is explained by parental and child aspirations for education at 12 years of age and previous measures of learning; (d) that in contrast, wealth gradients decline much less with the inclusion of these variables, indicating that the correlation between household economic circumstances and higher education access is only partly due to differences in early-childhood human-capital formation; and (e) that there are important differences in terms of gender in access to levels of higher education (favouring boys in India and girls in Vietnam) and in the association of various household and individual characteristics and parental and child aspirations with enrolment in higher education by 19 years of age. To the best of our knowledge, this is the first such comparative longitudinal analysis of access to higher education in developing-country settings.

Young Lives

An International Study of Childhood Poverty

About Young Lives

Young Lives is an international study of childhood poverty, involving 12,000 children in 4 countries over 15 years. It is led by a team in the Department of International Development at the University of Oxford in association with research and policy partners in the 4 study countries: Ethiopia, India, Peru and Vietnam.

Through researching different aspects of children's lives, we seek to improve policies and programmes for children.

Young Lives Partners

Young Lives is coordinated by a small team based at the University of Oxford, led by Professor Jo Boyden.

- *Ethiopian Development Research Institute, Ethiopia*
- *Pankhurst Development Research and Consulting plc, Ethiopia*
- *Centre for Economic and Social Studies, Hyderabad, India*
- *Save the Children India*
- *Sri Padmavathi Mahila Visvavidyalayam (Women's University), Andhra Pradesh, India*
- *Grupo de Análisis para el Desarollo (GRADE), Peru*
- *Instituto de Investigación Nutricional, Peru*
- *Centre for Analysis and Forecasting, Vietnamese Academy of Social Sciences, Vietnam*
- *General Statistics Office, Vietnam*
- *Oxford Department of International Development, University of Oxford, UK*

Contact:
Young Lives
Oxford Department of International Development,
University of Oxford,
3 Mansfield Road,
Oxford OX1 3TB, UK
Tel: +44 (0)1865 281751
Email: younglives@younglives.org.uk
Website: www.younglives.org.uk

Young Lives