

# Understanding higher education access: Inequalities and early learning in low and lower-middle-income countries

Sonia Ilie<sup>\*</sup>, Pauline Rose and Anna Vignoles

*University of Cambridge, UK*

Globally, access to higher education has increased, but inequalities by socio-economic background remain. This article explores the relationship between early schooling opportunities (and learning) and progression into higher education in four low and middle-income countries. We analyse data from the Young Lives longitudinal study, following cohorts of young people from age 5 to 22 in four country settings: Ethiopia, Peru, Vietnam and India. We reveal wide variability in higher education participation between the four countries, with a common pattern of a very strong association between early learning and later higher education participation, even after allowing for a range of demographic characteristics. Whilst early learning is important in predicting later higher education participation, we also find that significant barriers to higher education participation remain for low socio-economic status groups, even if they initially show good levels of learning. We track the trajectories of children who have initial good levels of learning, and hence arguably the potential to progress to higher education, and assess the extent to which socio-economic background plays a mediating role in these trajectories. Pupils with initially good levels of learning at primary school age, but who are from poor backgrounds, fall back in terms of their relative attainment during secondary schooling years. This implies that socio-economic status continues to be a barrier to educational attainment throughout these children's lives. We discuss the implications of these findings for policy initiatives aimed at narrowing inequalities in higher education access in poorer countries.

**Keywords:** access and learning; higher education; inequality; socio-economic status

## Introduction

Access to higher education (HE) in low and middle-income countries is undergoing a period of unprecedented growth (Ilie & Rose, 2016; Marginson, 2016a). This reflects earlier increases in primary school enrolment and completion, and improvements in secondary school access. Primary and to some extent secondary education have captured the majority of global development efforts, bolstered in part by the inclusion of universal primary completion as one of the Millennium Development Goals. Conversely, HE has received less policy attention. The inclusion of a target on equal access to HE in the Sustainable Development Goals (SDGs) reflects a recognition that rapid global growth in HE enrolment has been accompanied by wide disparities

---

<sup>\*</sup>Corresponding author. Faculty of Education, University of Cambridge, 184 Hills Rd, Cambridge CB2 8PQ, UK. Email: isi22@cam.ac.uk

in access. It also highlights an acknowledgement of the contribution of HE to other SDGs as a result of a re-evaluation of the individual and social value of HE (Oketch *et al.*, 2014; Altbach, 2016; Altbach *et al.*, 2019).

From a policy perspective, we locate our work within the specific SDG target on HE, namely: ‘by 2030 ensure equal access for all women and men to affordable quality technical, vocational and tertiary education, including university’ (SDG Target 4.3). This focus signals renewed attention towards commitments made in the 1948 United Nations Universal Declaration of Human Rights, which states: ‘higher education shall be equally accessible to all on the basis of merit’ (Art. 26, para. 1). However, growth in HE access has not occurred evenly across countries; and within countries, it has not occurred in an equitable manner (Marginson, 2016b). Notably, wealth, and to some extent gender, have been identified as major determinants of young people’s chances of accessing HE in many low and lower-middle-income countries (Ilie & Rose, 2016). This links to later life outcomes, with HE not entirely fulfilling its otherwise expected promise of social mobility, precisely because of a pattern of inequitable access, even if the social and financial benefits from HE are uniformly distributed across socio-economic backgrounds (Shafiq *et al.*, 2019).

In high-income country contexts, it has long been recognised that a life-course perspective to understanding education inequalities is needed. Access to HE is highly skewed towards students from higher-income backgrounds, since early socio-economic disadvantage in childhood leads to low levels of academic achievement. Specifically, given the importance of early investments in children’s cognitive and non-cognitive skills, socio-economic disadvantage in the early years has long-term negative effects on academic outcomes (Cunha *et al.*, 2006). Students from lower socio-economic status (SES) backgrounds are therefore far less likely to have the high levels of attainment needed to progress to HE. Indeed, in some countries such as the UK, low prior attainment is the main explanation for the socio-economic gap in HE participation (Anders, 2012; Chowdry *et al.*, 2013; Crawford *et al.*, 2017). In other countries, financial barriers at the point of entry into HE exacerbate the problem (Jerrim *et al.*, 2015). For example, in the USA, the high costs of HE are a major additional barrier to lower-income students (Barr *et al.*, 2017, 2019). However, the evidence on the nature of this problem in low and middle-income countries is far more limited. In these contexts, it is less clear whether socio-economic gaps in access to HE are attributable primarily to inequalities earlier in the education system or to barriers at the point of entry to HE.

This article seeks to address this gap in the literature. Specifically, our research questions are:

- What is the SES gap in HE participation across four low and lower-middle-income countries? How does this interact with gender?
- To what extent can the SES gap in HE participation be explained by differences in early learning attainment? Do these gaps widen or narrow at particular transition points in the education system?
- When do initially high-achieving low-SES students, who show early potential to progress to HE, start to fall behind their equally able but higher-SES peers?

We analyse data from the Young Lives longitudinal study, following cohorts of young people from age 5 to 22 in four country settings: Ethiopia, India, Peru and Vietnam. We reveal wide variability in HE participation between the four countries, but with a common pattern of a very strong association between early learning and later HE participation, even after allowing for a range of demographic characteristics. Whilst early learning is important in predicting later HE participation, we also find that significant barriers to HE participation remain for low-SES groups, even if they have high educational attainment in school. We track the trajectories of children who have high initial attainment in learning, and hence arguably have the potential to progress to HE, and assess the extent to which socio-economic background plays a mediating role in these trajectories. Pupils who are initially high attaining in primary school, but who are from poor backgrounds, fall back in terms of their relative attainment in secondary school. This implies that SES continues to be a barrier to educational attainment throughout these children's schooling. We discuss the implications of these findings for policy initiatives aimed at narrowing inequalities in HE access in poorer countries, and also for the data required for the continued monitoring of progress towards this goal.

### **Country contexts**

While all four countries in the Young Lives study have experienced rapid economic growth in recent years, they are at very different stages of economic development. Ethiopia is a low-income country, both India and Vietnam are lower-middle-income countries, and Peru is an upper-middle-income country. These four nations were intentionally chosen to illustrate the experiences of poverty over the life course in four distinct regions undergoing economic development, and because at the start of the study they shared a poverty-reduction policy agenda (Young Lives, 2017a). Access to schooling has expanded in all four countries and patterns of access largely reflect their stage of economic development. Comparable data on primary and secondary learning levels across all four countries is difficult to obtain, not least because a lot of data (e.g. PISA) only covers those who are enrolled in the school system and ignores those who have already dropped out, or never enrolled. Indeed, this is one reason why the Young Lives longitudinal study is so valuable: it follows all participating children, irrespective of whether they are enrolled in school, and assesses them in literacy and numeracy. Bearing in mind these limitations of existing data, we provide some contextual information on the enrolment rates by level of education and country and on the quality of provision and institutional features.

### **Primary and secondary enrolment**

Many children in Ethiopia continue to struggle to complete primary school and there are stark inequalities even in this early phase of schooling (see Figure 1), with 81% of boys in the richest quintile of households completing primary school in 2016, compared to just 26% of the poorest fifth. This translates into low levels of secondary school completion and HE attendance, with wide inequalities continuing to be apparent throughout the education life course. The three middle-income countries all

exhibit higher primary school completion rates than Ethiopia, reaching near-universal completion for the richest students, and relatively narrow inequality gaps in Peru and Vietnam. India similarly has near-universal primary school completion for the richest fifth of students, but with a wider socio-economic gap compared to the other two middle-income countries.

The majority of the richest students also complete upper secondary school in all three middle-income countries, but the completion rate for poorer students is far lower, resulting in very wide socio-economic inequality gaps at the secondary stage in both India and Vietnam. These gaps are illustrated in Figure 1 with internationally comparable household survey data collated by the UNESCO Institute for Statistics.<sup>1</sup>

### Quality of provision

The four countries all face challenges in terms of the quality of their education systems, but to varying degrees. Peru and Vietnam both participate in PISA.<sup>2</sup> As noted above, one concern with interpreting the PISA results is that the survey only assesses children in school. A considerable proportion of the poorest students are not in secondary school and are, therefore, excluded from the picture. Leaving that issue aside, the PISA data suggests wide socio-economic gaps in learning amongst those who are in school and tested. Socio-economic background explains around one-fifth of the variance in reading at age 15 in Peru in 2018, for example.<sup>3</sup> Peru is also amongst one of the poorer performers in PISA overall. While Vietnam's data for 2018 has so far not been included in the rankings, in previous PISA rounds it has been identified as performing better, with narrower socio-economic gaps compared to what might have been expected given the country's income level. Again, there is substantial selection of those who take the test in Vietnam. As for Peru, it excludes the poorest, who have

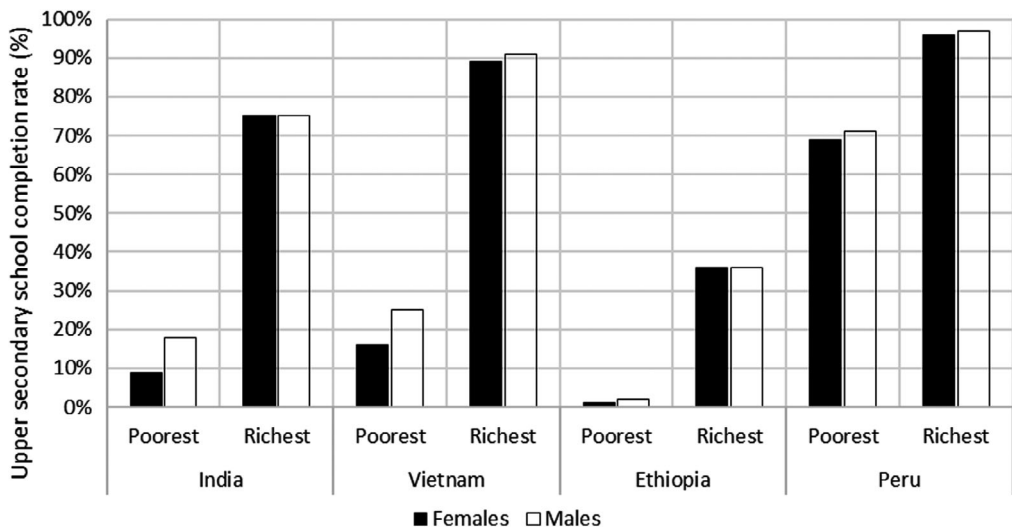


Figure 1. Upper secondary completion rate by gender and family income.

Source: UNESCO Institute for Statistics, drawing on Demographic and Health Survey and UNICEF Multiple Indicator Survey data.

already dropped out of school by that stage. Additionally, in Vietnam, only those who pass the entrance examination required to progress from Grade 9 to Grade 10 are able to proceed (Glewwe *et al.*, 2017).

While Ethiopia and India do not participate in internationally comparable assessments, rich national surveys and data identify that many children in both countries are not able to perform the most basic tasks of reading a sentence or basic numeracy in primary school, with wide socio-economic gaps also apparent (see e.g. Iyer *et al.*, 2020 on Ethiopia; Alcott & Rose, 2017 on India). As such, across the four Young Lives countries, the ability of young people from the poorest households to gain access to HE is likely to be seriously hampered by low levels of secondary school completion, coupled with poor quality of education for those who do make it to this stage.

### Higher education access

There is wide variability in access to HE in all four Young Lives countries. According to 2018 administrative data available from the UNESCO Institute for Statistics, the gross enrolment rate for tertiary education only reaches around 8% in Ethiopia. In India and Vietnam, the rate is 28% and 29%, respectively. In Peru, it is considerably higher, with a gross enrolment rate of around 71%. Growth in enrolment over the last 20 years has also varied. Ethiopia has yet to make a transition to a mass tertiary education system, increasing only marginally from an extremely low level in 2000, when only 1% of the cohort was enrolled in tertiary education. India and Vietnam have seen substantial expansion, both growing from a 10% enrolment rate in 2000 to nearly one-third of the cohort going on to HE in 2018. Peru already had a high HE rate in 2000, at 30%, and has since moved decisively to a mass HE system.

According to household survey data from these countries, access to HE is highly unequal across students from different socio-economic backgrounds. Even in the countries with higher overall enrolment rates, access for the poorest remains extremely limited. In both Peru and Vietnam, around 55% of the richest quintile of the age 18–22 population participate in HE. By contrast, the participation rate is just 5% for young people from the poorest fifth of households. For the same age group in Ethiopia, enrolment overall is considerably lower, with only 16% of the richest quintile attending HE, and a negligible number of the poorest making it to this stage.<sup>4</sup> Across all the countries, there is very little difference in enrolment by gender amongst either the richest or poorest groups. In Ethiopia, the relatively few students who are enrolled in HE are more likely to be male. By contrast, enrolment favours females in Vietnam and Peru, and gaps are very small indeed in India. The modest gaps by gender motivate our main focus on socio-economic gaps, although we also consider their intersection with gender given the focus on this within the SDGs. It is also worth noting that other background demographic characteristics also affect access to education, with all four countries having wide ethnic and language gaps in HE participation, and caste being a factor in India. We are able to account for these demographic factors in our modelling.

The four Young Lives countries vary somewhat in their institutional arrangements and funding of HE. In Ethiopia, the supply of places in public HE institutions has grown in recent years. The Ministry of Education sets minimum grade requirements

and quotas for different programmes each year, with public institutions generally receiving the highest-performing students.<sup>5</sup> Given the relatively small proportion of young people with access to HE, the country spends a significant proportion of its education budget on HE, reaching 48% in 2015 according to the UNESCO Institute for Statistics. Ethiopia HE institutions charge relatively modest fees for tuition and accommodation, although this may still be prohibitive for those from poorer backgrounds.<sup>6</sup> The other three countries spend rather less on HE, around 29% of their total education budget in the case of India (2013), 16% in Peru (2015) and 15% in Vietnam (2015). In Vietnam, India and indeed Peru, there is a mixed economy of public and private HE providers, with substantial growth in the latter and some concerns about quality of provision. Vietnam and India charge fees for HE, although they are low by international standards. India has means testing for fees, with affirmative action for those from disadvantaged social groups and castes. In Peru, public institutions are free but entry is highly selective and many students attend private institutions to prepare for the entry examinations, presenting another potential barrier for poor students.

### Relevant literature

We position our work within the literature that examines the global expansion of HE systems (Schofer & Meyer, 2005; Keeling, 2006; Marginson & Van der Wende, 2007), with many systems moving from elite-only to mass systems (Trow, 1973, 2007). This process of massification has improved access, but has not equalised it in high-income country contexts (e.g. Chowdry *et al.*, 2013 for the UK). In poorer countries, the growth in HE participation has also been considerable, with some improvements in access, albeit against a backdrop of very small numbers of students enrolled in HE overall (Carnoy *et al.*, 2013; Salmi & Bassett, 2014; Chien & Montjourides, 2016). Nonetheless, access to HE in southern contexts remains starkly unequal (Ilie & Rose, 2016; Shafiq *et al.*, 2019).

Our theoretical framing for the analysis is the cumulative skill model proposed by Cunha *et al.*, (2006). That model is derived from empirical evidence from high-income country contexts and has shown that investment in children's education and skills is both cumulative and complementary. Cunha *et al.*, (2006), building on theories about child development from the disciplines of education, economics and psychology, show that children who develop good cognitive skills (e.g. literacy, numeracy) in early childhood tend to learn more as they progress through the education system. The model implies that investment in the learning of young children is necessary to ensure their later success, in terms of educational outcomes. Note that the model does *not* suggest that early investment alone is sufficient to reduce later gaps in learning outcomes. Indeed, evidence from developed countries confirms that sustained investment throughout schooling in children and young adults from disadvantaged backgrounds is necessary. However, what is clear from this model is that *early* socio-economic disadvantage is likely to compound over time and hence modelling the determinants of HE participation needs to be based on a life-course perspective.

The Cunha *et al.*, (2006) model is also consistent, for example, with evidence from Jerrim and Vignoles (2013) and Crawford *et al.*, (2017), which has shown that

initially higher achieving but socio-economically disadvantaged children in the UK tend to lose their attainment advantage compared to their richer peers. The implication here is that even if poor children show very early promise academically, they tend to lose out to more advantaged peers over time. Understanding when and why this happens is obviously key to improving outcomes for the poorest children. Yet there is extremely limited evidence from low and middle-income countries about the educational learning attainment trajectories of children and young people. The evidence that is available is also usually only over a short period of 1 or 2 years (see e.g. Carter *et al.*, 2020).

In low and lower-middle-income country contexts there is also more generally very little evidence available on the link between initial learning and HE access. Data limitations have been acute. Whilst commonly available data (such as administrative and household survey datasets) in South Asian and sub-Saharan African countries is extremely valuable for the identification of HE access patterns, it does not provide a longitudinal perspective. This limits the opportunity to link HE access to earlier learning and schooling experiences. The Young Lives data described below allows a unique opportunity to examine this issue in low and middle-income countries, providing as it does a rich longitudinal perspective on cohort members. Some publications have already used the Young Lives data to identify the heterogeneous patterns in schooling across the four Young Lives contexts (Rolleston *et al.*, 2013), and to begin to document disparities in access to HE in each of the countries (Sanchez *et al.*, 2017; Young Lives, 2017b; Araya *et al.*, 2018; Espinoza *et al.*, 2018). Other evidence points to existing gaps in higher education access in specific national contexts (e.g. Jerrim *et al.*, 2015; Brewis, 2019; Chea, 2019). However, these papers have overall not focused specifically on how initial attainment determines later access and how this might vary by socio-economic background and gender.

One exception is Sanchez and Singh (2018), who use Young Lives data between 2002 and 2014 on India, Peru and Vietnam to examine the role of earlier childhood circumstances and HE enrolment.<sup>7</sup> Their findings suggest that academic attainment at age 12, and indeed parental and child aspirations at this age, do impact on HE enrolment by the time these children reach 19 years old. They conclude that these early factors only explain a small proportion of the variation in HE enrolment by SES, suggesting that barriers later in a child's schooling may be important.

We take this work as a starting point but extend it in three important ways. First, we include the most recent round of the Young Lives survey, which allows us to identify whether these patterns hold by the age at which it is likely that more children will have reached the HE stage (namely, age 22). In poor-country contexts, children and young people are often 'over-age' in comparison to the expected age of a particular grade (Lewin & Sabates, 2012). Assessing transitions into HE at age 19 is unlikely to capture all those who eventually make the transition into HE. Extending the possible period of enrolment to age 22 will capture those who have delayed entry. Second, we extend the analysis to include a low-income country, Ethiopia, which exhibits a different pattern of education access—notably with lower school enrolment overall. The example of Ethiopia could therefore be informative for other countries at an earlier stage of growth in HE participation. Third, we also focus on *when* SES gaps open up in children's educational trajectories, thereby identifying critical periods in the child's

schooling and providing evidence that can inform policymakers as to where/when intervention is most needed. Much of the existing literature addresses the question as to whether early learning predicts later HE enrolment, after allowing for differences in child circumstances, such as household SES (e.g. using approaches adopted by Anders, 2012a,b and Chowdry *et al.*, 2013 for England). Here we also take a different approach by analysing the transitions of young children who show initial high academic potential and who we might therefore expect to see enrolling in HE later on. We adopt the method used by Jerrim and Vignoles (2013) to track initially high-achieving children from poor backgrounds through their schooling and identify the trajectory of their initial academic advantage.

## Methodological approach

### *The Young Lives study*

We use data from the Young Lives longitudinal study of young people's experiences (educational and otherwise) in the above four countries: Ethiopia, Peru, Vietnam and two states in India—Andhra Pradesh and Telangana (these were initially one state at the beginning of the Young Lives survey). The study is unique in its longitudinal nature, tracking two cohorts of children in each country. For the purposes of this article, we use data for the older cohort, born in 1994–1995 and tracked from age 8 to 22 (a later round to be available in the future) by means of both quantitative surveys and qualitative interviews, across five survey rounds.

### *Sample and measures*

Young Lives adopts a pro-poor sampling design in principle, focused on 20 sentinel sites in each country, with a rural over-representation. As a result, the sample is not representative of each country's population of young people. Hence, the description of the different education systems across the four countries provided in the previous section by necessity comes from more nationally representative data. However, by over-sampling low-SES children, Young Lives provides valuable longitudinal data with which to study the relationship between early socio-economic disadvantage and later outcomes. Children and young people were surveyed at ages 8, 12, 15, 19 and 22. As with any longitudinal study, the Young Lives survey is liable to attrition issues. However, the survey has been able to retain at least 90% of its initial older-cohort sample by the fifth survey round (age 22), which is very high compared to many other longitudinal studies.<sup>8</sup> We use data from the older cohort sample from the study, which includes 1,000 children from each country who were between 7.5 and 8.5 years of age in the year 2000 (with the exception of Peru, where the older cohort sample starts with 714 participants). Combined with observable data on other variables of interest (see Table S2 in the online Supporting Information), we have an analytical sample of 942 respondents in India, 956 in Vietnam, 936 in Ethiopia and 650 in Peru. This represents an average retention rate in the study of 93.8% for these older cohort participants to age 22, ranging between 91% in Peru and 95.6% in Vietnam.



*Higher education access*

The key outcome of interest for this analysis is HE access. Young Lives provides disaggregated information on the educational level that each respondent has attained, at each respective survey round. We use this information from round 5, supplemented with information from round 4 for those missing in round 5, and compile a binary variable of ‘ever attended higher education’ that takes the value 1 if the respondent has ever indicated that they have attended any of the different categories of HE (see Table S1 in the online Supporting Information), and 0 otherwise.

We adopt a broad measure of HE access, which includes different types of post-secondary institutions, as relevant to each country context (see Table S1 for further information on individual tertiary qualifications included in this definition and sample sizes for each country). In India, we include technical institutes in our definition; in Ethiopia, we include technical and vocational tertiary courses, as well as post-secondary teaching qualifications; in Vietnam, we include (non-vocational) colleges; and in Peru, we include qualifications in technical and pedagogical institutes (but not qualifications in occupational/vocational adult education centres). Ideally, one would want an entirely comparable measure of HE across these countries. For example, one might want to only include individuals who enrolled in undergraduate degrees. However, given the heterogeneity of qualifications in each country, and the small number of people enrolled in this narrower definition of HE in three of the countries, we have opted to use a broad definition of HE.<sup>9</sup>

The addition of the fifth round of data in our analysis includes young people aged around 22. This addresses the issue that round 4 collected information when participants were aged roughly 19, and therefore would have only captured *on-time* access to HE. Being older than the official age-in-grade is a widespread phenomenon in some countries, notably in sub-Saharan Africa (Lewin & Sabates, 2012). This is apparent in the Ethiopian Young Lives sample. In Ethiopia, the official age to start school is 7. Yet in the Young Lives sample, only 51% of children are enrolled in school by age 8. Many start school later. Specifically, by age 12 we see 97.4% enrolled in school in Ethiopia, but with 62.6% of enrolled children over-age for their respective grade. The implication is that, at any age, a large proportion of children are in school grades that are lower than their chronological age would suggest. At the other extreme, in the Peru Young Lives sample, only 2.9% of all age 8 children are observed to be over-age for their grade, with an overall enrolment rate of 97.2%. In India, automatic progression means over-age enrolment is less prevalent initially. Equally importantly, in all countries except India, we observe a strong socio-economic gradient, with children in the wealthiest quartile significantly less likely than those in the poorest quartile to be over-age. Since poorer students are more likely to be above the official school age for their grade, it is important to avoid under-estimating their eventual enrolment in HE by only measuring ‘on-time’ HE enrolment.

*Socio-economic background*

The background variables of interest reflect our focus on wealth, gender and early learning. These include demographic characteristics captured in the first survey

round (age 8); schooling-related characteristics at age 8 and 12; and household characteristics, also measured in the first survey round. Demographics include gender and ethnicity, captured as a binary indicator of belonging to the majority ethnic group in the respective setting in the sample.<sup>10</sup>

Household characteristics include locality (sample site/cluster), whether the household is in an urban or rural environment, parental level of education and household wealth. All these variables are sourced from the first survey round when the child is age 8, and so identify their location and SES at the beginning of their schooling. This is important in contexts where students might have moved to urban areas to study.

We use a categorical variable of parental education, with the following categories: 'No education' (37.7% of the sample across all four countries), each additional year of schooling up to Grade 12, 'Post-secondary, vocational education' (3.2% of the whole sample) and 'University' (1.3% of the whole sample).

The household wealth variable is crucial given our focus on the relationship between household SES and outcomes. We use a country-specific wealth index, thereafter SES (see Briones, 2016 for further details) derived by the Young Lives study from household possession at the first survey wave. We then split the sample of households into quartiles of SES.

#### *School enrolment and attainment measures*

Data on individuals' schooling includes whether or not they are enrolled in school in each of the survey rounds and educational learning attainment in both literacy and numeracy in each of the first three survey rounds. Literacy is captured by a test of basic reading skills at age 8 and 12 (rounds 1 and 2 of the survey). Specifically, we know whether children are able to read a sentence, able to read only words/letters or not able to read at all. We also have a continuous measure of their literacy from a Peabody Picture Vocabulary Test (PPVT) administered in round 3 at age 15. Numeracy is captured at age 8 with a single arithmetic question (identifying if respondents provided a correct answer or not); at age 12 with a 10-item basic mathematics test (scored as the number of correct items out of 10); and at age 15 with a 20-item basic mathematics test (scored as the number of correct responses out of 20). Table S2 details the summary statistics for each of the measures included.

Young Lives also uniquely provides a wealth of information on the aspirations of the participants as well as their parents, in terms of their desire to go on to HE. Indeed, this was a focus of Sanchez and Singh (2018). However, previous research (e.g. Anders, 2012) suggests that such measures are endogenous to HE access. In other words, children and young people who are achieving well at school, and whose parents may envisage being able to support their child through HE, are more likely to be positive about their intentions to go to HE. Hence, high aspirations may in fact be an outcome from education rather than an independent predictor of whether a child is likely to go on to HE. This is consistent with results from Sanchez and Singh (2018) (table 4 of their paper) which suggest that the relationship between socio-economic background (measured by household wealth and parental education) and subsequent HE enrolment is significantly stronger when child and parent aspirations are

not included in the model. This indicates that one pathway where socio-economic circumstances impact upon outcomes is via parental and child aspirations. Given that the ordering of these effects cannot be clearly established, we omit such potentially endogenous variables from our models.

### Analytical approach

We use the above data in linear probability models with site fixed effects to estimate the relationship between SES and higher education background while controlling for a range of demographic and household characteristics; and for school enrolment and literacy levels, respectively, at higher ages (8, 12 and 15 years of age). Linear probability models were selected for ease of interpretation and since they allow for the implementation of site fixed effects, as discussed below. Before estimating the models with site fixed effects, raw models that only account for SES were estimated, to provide an indication of the raw, or unadjusted, gap between the richest and poorest quartile in HE enrolment in each of these countries; then also adding in gender, interacted with SES quartile. While these raw models do not account for the nature of the Young Lives sample, they are informative overall in terms of the across-sites average SES gap in HE enrolment in each sample country.

The models including only site fixed effects provide an unadjusted estimate of the within-site gap in HE enrolment between the richest and the poorest quartiles in each respective Young Lives country. Subsequently, models controlling for demographic and household characteristics, but no individual education-related characteristics, provide evidence around the HE enrolment gaps accounting for the variation of these background characteristics. Finally, models that sequentially add education-related characteristics (enrolment and literacy levels) to the above provide insight into the HE enrolment gaps by SES at the same level of prior educational attainment, the focus of the research questions.

The models account for the clustering of participants in the sample sentinel sites through the inclusion of the site fixed effects and use robust standard errors (with a Huber–White estimator only). The models, estimated separately for each of the four countries, and sequentially as above, take the following general form:

$$y_{is} = \beta_0 + \beta_1 SES_i + \beta_2 C_i + \beta_3 E_i + u_s + \varepsilon_{is}$$

where  $y_{is}$  is a binary variable capturing HE enrolment for individual  $i$  in site  $s$ , taking the value 0 or 1;  $SES_i$  is a vector of SES dummy variables with the reference category being the top richest quartile;  $C_i$  is a vector of control variables that includes demographic characteristics and household characteristics;  $E_i$  is a vector of education-related characteristics, specifically enrolment and literacy levels, captured at consecutive ages (8, 12 and 15, respectively) for each individual;  $u_s$  is the site fixed effect (not included in the initial raw models); and  $\varepsilon_{is}$  is the error term.

## Results

### *Patterns of higher education access*

We start by documenting the wide variability in HE participation across the four countries in the Young Lives samples. Given the nature of the sampling approach, as noted above, these data are not representative of the populations in these countries since a pro-poor sample has been selected. The proportion enrolled in HE will therefore not match the proportion enrolled as recorded in national statistics. It is evident, however, that the enrolment rate increases for HE enrolment up to age 22 (from the round 5 survey) compared with age 19 (from round 4), thereby justifying considering HE enrolment at a rather later age.

Table 1 indicates the extent to which HE enrolment varies by socio-economic background across the four countries. Figure 2 then breaks these enrolment rates down by SES and gender, based on the raw models outlined above. HE enrolment displays a strong socio-economic gradient, with the poorest quartile least likely to have attended HE, and the richest quartile most likely to have attended in all four Young Lives countries.

Gender interacts with socio-economic background differently across the four settings. In India, in this pro-poor sample who (theoretically) could have first entered HE in 2014, women are still consistently less likely to have attended HE than men in each wealth quartile (though national data suggests that by 2018, the overall HE participation rate for males and females was similar). In Vietnam, there is no systematic gender imbalance. In Ethiopia and Peru, women are more likely to have enrolled in HE in most socio-economic groups. The exception is the 'poorest' and 'poorer' groups in Peru, which still have higher participation rates for males. For Ethiopia, the fact that HE appears to favour females in the sample could be because of the wider definition of HE (including post-secondary teaching qualifications) used, in comparison to administrative data reported in an earlier section of the article.

### *Explanatory factors*

*Sentinel site sample structure.* Figure 2 illustrates the entirely unadjusted gaps in HE enrolment by socio-economic background. However, these gaps belie the nature of

Table 1. Rate of access to higher education (ever), by wealth quartile at round 1 and Young Lives country

Country	Higher education enrolment rate (%) wealth quartile				Overall	Gap between richest and poorest (percentage points)
	Poorest	Poorer	Richer	Richest		
India	31.0	42.2	52.2	66.2	47.7	35.2
Vietnam	13.9	28.1	39.9	51.2	32.8	37.3
Ethiopia	9.9	12.4	26.0	37.3	21.4	27.4
Peru	33.1	45.3	55.1	76.4	52.5	43.3

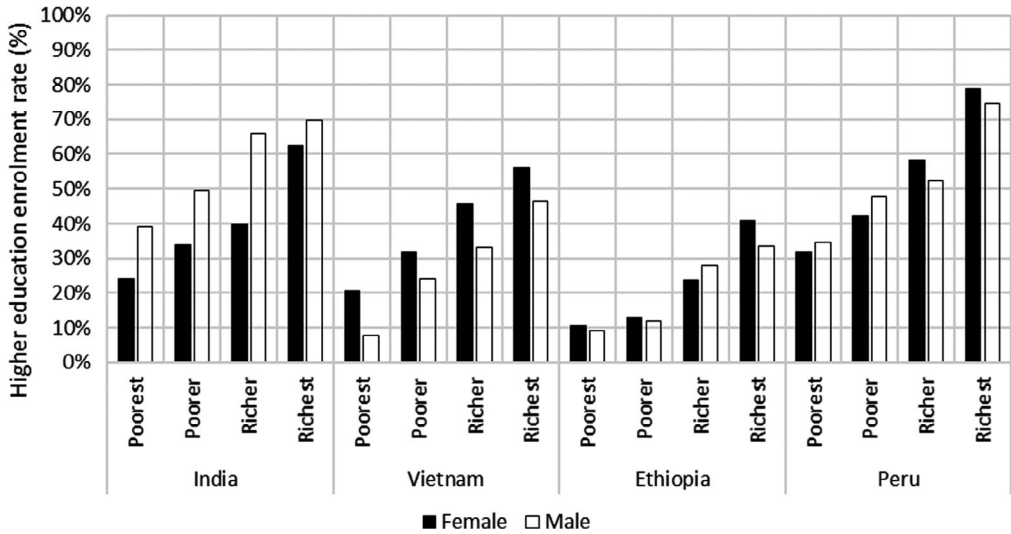


Figure 2. Patterns of higher education access in the Young Lives sample, by wealth and gender.

the Young Lives sample, not only in terms of its pro-poor nature, but also in relation to its sentinel site sampling approach. To account for this, the leftmost bars in Figure 3 illustrate the gap in HE enrolment from a linear probability model with site fixed effects (as specified above in the analytical approach discussion).

In India, Vietnam and Peru, moving from the entirely unadjusted HE enrolment gaps to raw gap that accounts for the sites, the magnitude increases. This is because the average within-site gaps in HE enrolment are higher than when considering all Young Lives respondents independently of their site. This points to substantial between-site variation. This is accounted for by the addition of the site fixed effects in

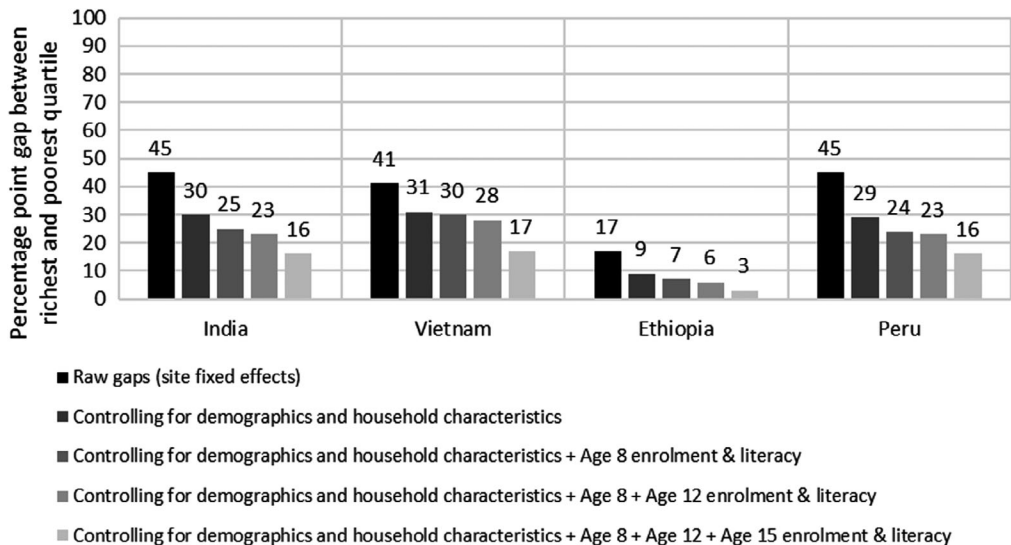


Figure 3. Gap in higher education enrolment rate between the richest and the poorest quartile in each Young Lives country.

the above models, which therefore provide more precise estimates of the HE enrolment gaps. In Ethiopia, the pattern is reversed, indicative of the complex patterns around enrolment in HE and education generally, as we discuss subsequently.

Therefore, in what follows we use the site fixed effects models as the baseline against which to estimate how HE enrolment gaps are explained by enrolment and learning at progressively older ages.

*Demographic and household characteristics.* We find that the raw gaps (allowing for site fixed effects, the leftmost column in Figure 3) are substantially reduced when explanatory variables that capture individual demographics and household characteristics are accounted for (as illustrated in the second column in Figure 3). We include a variable identifying whether, at round 1, individuals were deemed to live in urban or rural environments. While this aligns substantially with the sampling sites in India, Vietnam and Ethiopia, in the Peru sample we observe within-site urban/rural variation, and for the sake of consistency across the four country models, we retain this variable in all countries.

Accounting for gender, ethnicity and urban location reduces the apparent socioeconomic gap in HE enrolment. This implies that in these countries, there are inter-sectional effects whereby these factors interact with SES to predict the likelihood of a person enrolling in HE.

There is remarkable consistency in the results regarding SES gaps in HE enrolment across the country settings. In India, Vietnam and Peru, the SES gap shrinks from around 41–45 percentage points to 29–30 percentage points when moving from the raw gap to the gap after accounting for demographic and household characteristics. Hence, in these countries, demographic and household characteristics explain around a quarter of the total SES gap in HE participation between rich and poor students. In Ethiopia, by contrast, there is a relatively smaller raw gap of 17 percentage points (partly reflecting the lower levels of HE participation overall). This gap is reduced to just 9 percentage points after allowing for children's demographic and household characteristics.

In each country, therefore, a substantial proportion of the apparent SES gap is in fact down to differences in HE enrolment across urban and rural environments, ethnicity and gender. Since living in an urban location and ethnicity are correlated with SES, including these variables in the model is important to identify the independent effect of SES, which is empirically observed to be considerable in all four country settings.

### *Early school enrolment and learning*

Our key hypothesis in this article is that gaps in HE enrolment may be partially explained by early school enrolment and learning, first captured in this Young Lives cohort at age 8. We assume that even if children were not enrolled at age 8, they would have had the opportunity to enter schooling later, beyond the official school starting age. Hence, individuals might start school later but still potentially progress to HE and they are therefore included in the sample. Indeed, this is what we observe in Ethiopia, which exhibits a lower enrolment rate at age 8 than the other Young Lives countries (Table S1).

Once school enrolment and literacy levels at age 8 are included, the SES gaps above are reduced further. In India, for instance, the large raw gap in HE enrolment between the richest and the poorest students is reduced by a substantial amount to 25 percentage points when comparing young people with the same background characteristics and with similar early educational experiences and attainment. In Peru, the raw gap is reduced by almost half (from 45 to 24 percentage points), indicating the important role of early school enrolment, and of learning, even in a country with relatively high levels of primary school access. In Vietnam, after controlling for age 8 enrolment and learning, the gap is also reduced (to 30 percentage points), but early learning makes only a marginal difference and it is demographic factors that matter more. In Ethiopia, the remaining SES gap becomes quite small (7 percentage points) after allowing for differences in early schooling enrolment and attainment.

Overall, the results are somewhat consistent with evidence in high-income settings. In England, Chowdry *et al.*, (2013) also consistently find a large reduction in raw HE access gaps when prior attainment measures are added to the model. It should be noted, however, that the SES gap in HE enrolment remains much larger in these four Young Lives countries, even after accounting for early attainment. Specifically, in India, Vietnam and Peru, the likelihood of enrolling in HE is far greater for students from richer backgrounds, even when comparing students who have had a similar educational start in life (i.e. enrolled in school and achieving similar literacy at age 8). This finding suggests that there are additional barriers to HE participation rates in these country contexts, an issue which is highly relevant to policy, as discussed below.

### Later education enrolment and outcomes

The results above show that demographics, household characteristics and early educational experiences and learning attainment in primary school are important in explaining subsequent HE enrolment for poor students. So, trying to ensure higher levels of school enrolment and early literacy levels in primary school for poorer children is a priority if the goal is higher HE enrolment for all. However, large SES gaps remain even for those who do succeed earlier in the system. We need to understand better *when* the SES gaps narrow, if at all. To address this, we add information about school enrolment and literacy at ages 12 and 15, respectively, to the above models.

These results follow the pattern identified previously, whereby the addition of each subsequent point of data on school enrolment and literacy level provides further explanation of the HE enrolment gap between the richest and the poorest students. This is the case across all four Young Lives country settings.

In India, for instance, the raw gap is more than halved once age 15 enrolment and attainment are accounted for. Note that in Figure 3, allowing for literacy attainment at each age point reduces the SES gap in a relatively linear manner for most countries. This implies that SES continues to play a role in determining attainment throughout the education system, not just in the earlier years. This is true for Vietnam too, but including age 15 enrollment and attainment in the model reduces the SES gap in Vietnam to a far greater extent than in the other countries. This suggests that in Vietnam, attainment in secondary school is both highly correlated with SES and important in predicting HE participation, more so than attainment at an earlier age. The

Vietnamese system is highly selective with very rigorous examinations, including national examinations that are the entry requirement for HE. This might explain why academic success in secondary school is both a stronger determinant of HE participation than in some other countries and more highly correlated with SES.

Overall, even when we allow for differences in literacy at age 15, we still see a sizeable SES gap remaining in India, Vietnam and Peru. This differs from many developed countries, such as the UK. This implies that children's learning (in school) does not explain whether or not they progress to HE to the same extent in these low and middle-income countries. Instead, family background continues to play a strong role in predicting enrolment. This may be attributable to financial barriers, for example tuition fees and other costs associated with access to HE in each country, as well as the opportunity costs of not working.

### *The protective effect of wealth*

Figure 3 shows a steady reduction in the SES gap between the poorest and the richest students in terms of their likelihood of HE enrolment, after allowing for differences in their early and later school enrolment and learning. Yet, as has been said, SES gaps remain even after allowing for differences in children's academic attainment. This implies that the impact of socio-economic background on academic attainment extends throughout children's educational trajectories.

To investigate this still further, we focus on a particular subgroup, to determine how household wealth protects them (or otherwise) from low attainment and facilitates HE participation. We identify children who, on the basis of early potential measured at age 8 (i.e. a good level of learning, whether they attend school or not), we might expect would be most likely to proceed to HE. We term this the 'high-promise' group. We then model the extent to which the trajectories of this high-promise group vary by socio-economic background. This helps us to understand whether poor students who have a good initial level of learning manage to sustain their academic advantage, or alternatively the extent to which their socio-economic circumstances over-ride their initial early promise.

We do this by assigning Young Lives respondents to the high-promise group in each respective country, based on their age 8 literacy, with all children able to read sentences included in this group. Then, to avoid the issue of regression to the mean (Jerrim & Vignoles, 2013), we track their educational progress in terms of a separate measure of attainment, namely numeracy (measured as described above) at ages 8, 12 and 15. We also consider their likelihood of then going on to HE. The data are far from ideal. At age 8, the numeracy measure available is a binary indicator of whether or not the child can perform a simple mathematical calculation. The tests at age 12 and 15 are more informative. Figure 4 therefore shows, for children identified as belonging to the high-promise group at age 8, the percentage who at age 8 can undertake the mathematical calculation. At age 12 and 15, Figure 4 shows this group's percentile rank on the numeracy tests. The sample is divided into the high-promise students from the poorest and the richest wealth quartiles, for each country.

Figure 4 shows that for the high-promise group, even at age 8 there is a socio-economic gap in the likelihood of the child being able to answer the mathematical



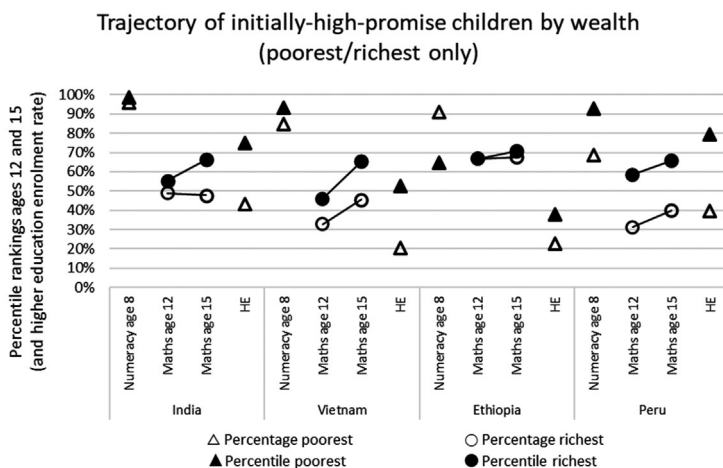


Figure 4. Trajectory of initially high-promise children, by wealth background in each Young Lives country.

question correctly. In India and Vietnam, the gap is very small. In Peru, there is a larger SES gap. Interestingly, in Ethiopia, the high-promise poor group are more likely to be able to complete the mathematical calculation than the richer group at age 8, though they are similarly likely to be enrolled in school (above 90% of children in these respective groups are enrolled). We are mindful of the data limitations here. First, high performance on reading is correlated, but not synonymous, with performance in mathematics. Second, the measures of both literacy and numeracy at age 8 are limited. Nonetheless, it is still informative to investigate what happens to this group of children who make a relatively strong start to their learning.

Our main focus is on what happens between age 12 and 15. In India and Vietnam, it is clear that the gap between high-promise children from rich backgrounds and high-promise children from poor backgrounds widens. Hence, despite a good start, SES continues to impact on academic achievement and poorer students do not fare as well as the high-SES group. This implies that in these countries, high-promise poor students experience barriers during their respective secondary school years that mean their learning is not as strong as that of their high-SES counterparts. In Ethiopia, progress for high and low-SES pupils in this high-promise group is very similar, suggesting no additional impact from SES during secondary school years, with secondary school enrolment relatively lower in Ethiopia compared to the other Young Lives countries. In Peru, although the high-promise, high-SES group does perform better than the high-promise, low-SES group, the gap does not widen. For Peru, it seems that barriers earlier in the sample are more likely to explain SES gaps for this high-promise group than issues arising during secondary school years.

Figure 4 also shows the proportion in each group that goes on to be enrolled in HE. In every country there is a wide gap in HE participation between rich and poor within this high-promise group. So initially good levels of learning, and even relatively good progress in secondary school, are not enough to ensure parity in the likelihood of participating in HE. This implies that there are barriers in place at that critical transition from secondary to HE.

## Conclusions

Our findings suggest a very strong association between socio-economic background and educational attainment, throughout children's schooling and in all four Young Lives country contexts. There is considerable variation in the proportion of students who enrol in HE across the four countries but, despite this, there is also clearly a common pattern of a strong association between early learning and later HE participation. Demographic and household characteristics, particularly gender, ethnicity and urban location, also predict HE participation. Some of these demographic factors, such as ethnicity and location, are themselves correlated with SES. This is evidence of the intersectionality of demographic characteristics, SES and educational success.

Our analysis therefore clearly suggests that early learning is an important predictor of later educational attainment, particularly HE participation. Therefore, whether a child is enrolled in school and achieving well at age 8 is an important predictor of later HE participation. However, we also observe that enrolment and attainment at age 12 and 15 are also strong predictors of HE participation (and in the case of Vietnam, even stronger than age 8 attainment) and explain a significant amount of the SES gap. Even when we allow for prior attainment up to age 15 and a range of other background characteristics, we find large residual SES gaps, suggesting significant barriers for poor students at the point of entry into HE. To explore this issue further, we focused on a group of students who showed strong early promise in reading at the age of 8. Even for this group, the students from poor backgrounds fall away in India and Vietnam, in terms of their relative attainment as they progress through the system. In Peru and Ethiopia, this high-promise group fares better in secondary school. However, in all four countries, poor students in this high-promise group are far less likely to go on to HE. Together, these findings imply that SES continues to be a barrier to attainment throughout the system. We now discuss the implications of these findings for policy initiatives aimed at narrowing inequalities in HE access in poorer countries, and also for the data required for the continued monitoring of progress towards this goal.

## Policy implications

Policymakers urgently need robust evidence on the nature and causes of the socio-economic gap in HE participation globally. With the expansion of HE in many countries, and increasing costs, policy decisions about how to expand and fund HE efficiently and equitably have become pressing. In countries with very limited access to HE for poorer students, additional taxpayer funding of HE is highly regressive. It essentially taxes those who are very unlikely to benefit from HE to subsidise a socio-economic elite who themselves gain a great deal of personal economic benefit from their HE (in the form of higher wages and better job opportunities). One solution that many countries have followed is to charge tuition fees for HE. However, charging fees without giving students access to income-contingent state loans is likely to be a major barrier to access to HE for low-income students. If a poor student did make it through the education system with sufficiently high attainment to progress to HE, high tuition fees would then most likely prevent them from doing so. These tensions imply some

difficult choices about how to organise and fund HE, especially when the evidence base around what is effective at meaningfully increasing equitable access to higher education, particularly for the most disadvantaged, rarely offers simple policy solutions, across both low and high-income contexts (Younger *et al.*, 2019). Evidence on when and how inequalities in access and attainment emerge in education systems is certainly vital to having an informed debate on this issue. Much of the evidence to date has been from developed countries and for this reason, this article seeks to provide similar evidence for a selection of low and middle-income countries.

Our findings indicate that poor students are less likely to enrol in HE in these countries partly because they lack the levels of prior attainment needed for entry, and indeed a significant proportion of the poorest students have dropped out of education altogether by the age of 15. The first policy implication from this work is therefore that increasing investment in primary schooling for the poorest students remains a priority. Only if we can increase the learning of the poorest students in the earlier years of their schooling is it likely that they will have sufficient attainment to progress on to HE.

Our second key finding is that attainment at every age is highly correlated with socio-economic background in all four countries. In all countries bar Ethiopia, being stronger academically in primary school is not enough; if you are from a poor background, you are still likely to have lower attainment by the age of 15. Certainly for poorer students, lower attainment at age 15 is still acting as a barrier to progression into HE. The implication is that targeted and sustained interventions and funding need to be directed at the poorest students throughout their schooling, if we are to narrow the SES gap in HE participation.

Lastly, we also found clear evidence of residual large SES gaps in HE participation, even when comparing children with similar levels of prior attainment at age 8, 12 and 15. So poor students are not progressing to HE as much as one would expect given their prior attainment, suggestive of additional barriers at the point of entry into HE. These might be financial barriers, such as tuition fees or the need to earn income, or psychological barriers, such as lower expectations for progression or feeling university is not for them. From a policy perspective, we certainly need better evidence on the effectiveness of different HE financial support systems for poor students. In particular, robust evidence is needed on the effectiveness of targeted interventions that reduce the costs of HE for poor students whilst not subsidising children from wealthier backgrounds who disproportionately participate in HE. So, for example, we might not want to increase the proportion of the education budget spent on HE, given the evidence that investment is needed earlier in the system to improve the learning of poor students in primary and secondary school. We might, however, want to evaluate the effectiveness of means-tested grants in encouraging HE participation among the poorest students.

## **Data implications**

The Young Lives study has yielded incredibly rich and valuable data. It has enabled researchers to really understand the lives of children and young people in an inclusive manner, including as it does children who are often outside the formal education

system altogether. The data does, however, have limitations. The richness of the data means that collecting it is expensive and sample sizes are relatively small. It is therefore vitally important that such survey data is supplemented with better administrative data that can enable researchers to track students through the education system, providing a clearer view of when they drop out, their attainment along the way and their eventual participation in HE. As countries move to mass HE systems, this kind of administrative data will become increasingly valuable in terms of monitoring the effect of massification and ensuring that it does not reproduce inequalities. Further, in several of the countries studied in this article, quality of HE provision has been raised as an issue. Enrolling more students in HE is only valuable if they obtain the skills they need to thrive in society and economically. Data on the outcomes for graduates, in terms of job quality, earnings and wider social outcomes, continues to be vitally important as HE systems expand and concerns about quality grow.

#### NOTES

- <sup>1</sup> Data for Ethiopia, India and Peru are from the Demographic and Health Survey datasets, and those for Vietnam are from the UNICEF Multiple Indicator Cluster Survey. Poorest and richest refers to the top and bottom quintile of households using an index of assets compiled by these surveys as a proxy for wealth.
- <sup>2</sup> Vietnam's results were excluded from the 2018 PISA rankings due to concerns over data quality.
- <sup>3</sup> <https://gpseducation.oecd.org/CountryProfile?primaryCountry=PER&treshold=10&topic=PI>
- <sup>4</sup> Data for Peru are from 2014, Vietnam from 2012 and Ethiopia from 2016. The latest comparable data from India, from 2006, show a similar pattern to Ethiopia at that time, namely, very low access overall, with negligible enrolment for the poorest. All data are from the World Inequality Database on Education (WIDE): [www.education-inequalities.org/](http://www.education-inequalities.org/)
- <sup>5</sup> <https://wenr.wes.org/2018/11/education-in-ethiopia>
- <sup>6</sup> <https://wenr.wes.org/2018/11/education-in-ethiopia>
- <sup>7</sup> They exclude Ethiopia from their analysis because most students at the time of the survey had still not completed secondary school. They therefore focus only on middle-income countries.
- <sup>8</sup> Further information on the sample design and attrition can be found at [www.younglives.org.uk/content/sampling-and-attrition](http://www.younglives.org.uk/content/sampling-and-attrition)
- <sup>9</sup> This approach is consistent with Sanchez and Singh's (2018) analysis of HE using Young Lives data.
- <sup>10</sup> These groups include Telugu in Andhra Pradesh in India; Kinh in Vietnam; Oromo in Ethiopia; Mestizo including Andean Indian in Peru.

#### References

- Alcott, B. & Rose, P. (2017) Learning in India's primary schools: How do disparities widen across the grades?, *International Journal of Educational Development*, 56, 42–51.
- Altbach, P.G. (2016) *Global perspectives on higher education* (Baltimore, MD, Johns Hopkins University Press).
- Altbach, P.G., Reisberg, L. & Rumbley, L.E. (2019) *Trends in global higher education: Tracking an academic revolution* (Leiden, Brill).
- Anders, J. (2012a) The link between household income, university applications and university attendance, *Fiscal Studies*, 33(2), 185–210.
- Anders, J. (2012b) *Using the Longitudinal Study of Young People in England for research into higher education access* (No. 12–13) (London, Institute of Education).
- Araya, M., Woldehanna, T. & Pankhurst, A. (2018) *Youth transitions—skills, work and family formation: Preliminary findings from the round 5 survey in Ethiopia*. Available online at: <https://ora.ox.ac.uk/objects/uuid:3f66f7e4-62e4-4350-ab60-7a7f3efd1e6e> (accessed 12 February 2020).
- Barr, N., Chapman, B., Dearden, L. & Dynarski, S. (2017) *Getting student financing right in the US: lessons from Australia and England*. Centre for Global Higher Education. Working paper no. 16. Centre for Global Higher Education working paper series. Available online at: [https://edpolicy.umich.edu/files/01-2017\\_student-financing.pdf](https://edpolicy.umich.edu/files/01-2017_student-financing.pdf) (accessed 23 August 2019).

- Barr, N., Chapman, B., Dearden, L. & Dynarski, S. (2019) The US college loans system: Lessons from Australia and England. *Economics of Education Review*, 71, 32–48.
- Brewis, E. (2019) Fair access to higher education and discourses of development: A policy analysis from Indonesia, *Compare: A Journal of Comparative and International Education*, 49(3), 453–470.
- Briones, K. (2016) 'How many rooms are there in your house?' *Constructing the Young Lives Wealth Index*. Available online at: [http://doc.ukdataservice.ac.uk/doc/8357/mrdoc/pdf/8357\\_yl-tn43\\_0.pdf](http://doc.ukdataservice.ac.uk/doc/8357/mrdoc/pdf/8357_yl-tn43_0.pdf) (accessed 04 March 2020).
- Carnoy, M., Loyalka, P., Dobryakova, M., Dossani, R., Froumin, I., Kuhns, K. et al (2013) *University expansion in a changing global economy: Triumph of the BRICs?* (Stanford, CA, Stanford University Press).
- Carter, E., Rose, P., Sabates, R. & Akyeampong, K. (2020) Trapped in low performance? Tracking the learning trajectory of disadvantaged girls and boys in the Complementary Basic Education programme in Ghana, *International Journal of Educational Research*, 100, 101541.
- Chea, P. (2019) Does higher education expansion in Cambodia make access to education more equal?, *International Journal of Educational Development*, 70, 102075.
- Chien, C. & Montjourides, P. (2016) Global trends in access to post-secondary education, in: N. Harrison (Ed.) *Access to higher education: Theoretical perspectives and contemporary challenges* (Abingdon, Routledge).
- Chowdry, H., Crawford, C., Dearden, L., Goodman, A. & Vignoles, A. (2013) Widening participation in higher education: Analysis using linked administrative data, *Journal of the Royal Statistical Society, Series A*, 176(2), 431–457.
- Crawford, C., Macmillan, L. & Vignoles, A. (2017) When and why do initially high-achieving poor children fall behind?, *Oxford Review of Education*, 43(1), 88–108.
- Cunha, F., Heckman, J.J., Lochner, L. & Masterov, D.V. (2006) Interpreting the evidence on life cycle skill formation. In E. Hanushek & F. Welch (Eds) *Handbook of the economics of education* (vol. 1) (Amsterdam, Elsevier), 697–812.
- Espinoza, P., Benny, L., Duc, L.T. & Hang, N.T.T. (2018) *Youth transitions—skills, work and family formation: Preliminary findings from the round 5 survey in Vietnam*. Available online at: <https://ora.ox.ac.uk/objects/uuid:833ceadf-59e9-45df-8ddf-17e6dba561a2> (accessed 04 March 2020).
- Glewwe, P., Lee, J., Vu, K. & Dang, H.A. (2017) What explains Vietnam's exceptional performance in education relative to other countries? Analysis of the 2012 PISA data, paper presented at the *Research on Improving Systems of Education (RISE) Annual Conference, Washington, D.C.*
- Ilie, S. & Rose, P. (2016) Is equal access to higher education in South Asia and sub-Saharan Africa achievable by 2030?, *Higher Education*, 72(4), 435–455.
- Iyer, P., Rolleston, C., Rose, P. & Woldehanna, T. (2020) A rising tide of access: what consequences for equitable learning in Ethiopia? *Oxford Review of Education*, 46(5), 601–618.
- Jerrim, J., Chmielewski, A.K. & Parker, P. (2015) Socioeconomic inequality in access to high-status colleges: A cross-country comparison, *Research in Social Stratification and Mobility*, 42, 20–32.
- Jerrim, J. & Vignoles, A. (2013) Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes, *Journal of the Royal Statistical Society, Series A*, 176(4), 887–906.
- Keeling, R. (2006) The Bologna Process and the Lisbon Research Agenda: The European Commission's expanding role in higher education, *European Journal of Education*, 41(2), 203–223.
- Lewin, K.M. & Sabates, R. (2012) Who gets what? Is improved access to basic education pro-poor in Sub-Saharan Africa?, *International Journal of Educational Development*, 32(4), 517–528.
- Marginson, S. (2016a) High participation systems of higher education, *The Journal of Higher Education*, 87(2), 243–271.
- Marginson, S. (2016b) The worldwide trend to high participation higher education: Dynamics of social stratification in inclusive systems, *Higher Education*, 72(4), 413–434.
- Marginson, S. & Van Der Wende, M. (2007) *Globalisation and higher education*. OECD Education Working Paper No. 8.

- Oketch, M., McCowan, T. & Schendel, R. (2014) *The impact of tertiary education on development. Education rigorous literature review*. Available online at: [https://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=PORTfd\\_qWfo%3D&tabid=3437](https://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=PORTfd_qWfo%3D&tabid=3437) (accessed 25 June 2019).
- Rolleston, C., James, Z. & Aurino, E. (2013) *Exploring the effect of educational opportunity and inequality on learning outcomes in Ethiopia, Peru, India, and Vietnam*. Available online at: [www.younglives.org.uk/sites/www.younglives.org.uk/files/Rolleston%20et%20al\\_UNESCO-GMR-background-paper\\_Jan2014.pdf](http://www.younglives.org.uk/sites/www.younglives.org.uk/files/Rolleston%20et%20al_UNESCO-GMR-background-paper_Jan2014.pdf) (accessed 14 February 2020).
- Salmi, J. & Bassett, R.M. (2014) The equity imperative in tertiary education: Promoting fairness and efficiency, *International Review of Education*, 60(3), 361–377.
- Sánchez, A., Pazos, N., Manrique, G. & Ferrer, J. (2017) *Youth transitions—skills, work and family formation: Preliminary findings from the 2016 Young Lives Survey (round 5): Peru*. Available online at: <https://ora.ox.ac.uk/objects/uuid:480817a1-e28f-4940-8ed1-b37ba8bc4968> (accessed 12 February 2020).
- Sánchez, A. & Singh, A. (2018) Accessing higher education in developing countries: Panel data analysis from India, Peru, and Vietnam, *World Development*, 109, 261–278.
- Schofer, E. & Meyer, J.W. (2005) The worldwide expansion of higher education in the twentieth century, *American Sociological Review*, 70(6), 898–920.
- Shafiq, M.N., Toutkoushian, R.K. & Valerio, A. (2019) Who benefits from higher education in low- and middle-income countries?, *The Journal of Development Studies*, 55(11), 2403–2423.
- Tran, L., Marginson, S., Do, H., Le, T., Nguyen, N., Vu, T. et al (2016) *Higher education in Vietnam: Flexibility, mobility and practicality in the global knowledge economy* (New York, Springer).
- Trow, M. (1973) *Problems in the transition from elite to mass higher education*. Available online at: <http://files.eric.ed.gov/fulltext/ED091983.pdf> (accessed 25 June 2019).
- Trow, M. (2007) Reflections on the transition from elite to mass to universal access: Forms and phases of higher education in modern societies since WWII, in: J.F. Forest & P.G. Altbach (Eds) *International handbook of higher education* (Amsterdam, Springer), 243–280.
- Young Lives (2017a) *A guide to Young Lives research*. Available online at: [www.younglives.org.uk/sites/www.younglives.org.uk/files/GuidetoYLResearch\\_0.pdf](http://www.younglives.org.uk/sites/www.younglives.org.uk/files/GuidetoYLResearch_0.pdf) (accessed 25 September 2019).
- Young Lives (2017b) *Youth transitions—skills, work and family formation: Preliminary findings from the round 5 survey in India*. Available online at: <https://ora.ox.ac.uk/objects/uuid:11b32c3b-1d1e-4dbb-9f31-a56fbc224e3e> (accessed 25 September 2019).
- Younger, K., Gascoine, L., Menzies, V. & Torgerson, C. (2019) A systematic review of evidence on the effectiveness of interventions and strategies for widening participation in higher education, *Journal of Further and Higher Education*, 43(6), 742–773.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Qualifications classified as Higher Education, by country.

Table S2. Variables of interest, by country (analytical sample).

Table S3. Model with site-fixed effects only.

Table S4. Model with site-fixed effects and demographic controls.

Table S5. Model with site-fixed effects, demographic controls, and age 8 school enrolment and learning.

Table S6. Model with site-fixed effects, demographic controls, age 8 school enrolment and learning, and age 12 school enrolment and learning

Table S7. Model with site-fixed effects, demographic controls, age 8 school enrolment and learning, age 12 school enrolment and learning, and age 15 school enrolment and learning.