PERUVIAN ECONOMIC ASSOCIATION

Stratification of returns to higher education in Peru: the role of education quality and major choices

Alan Sánchez

Marta Favara

Catherine Porter

Working Paper No. 180, May 2021

Stratification of returns to higher education in Peru: the role of education quality and major choices[1]

March 30, 2021

Alan Sánchez[2]

Marta Favara[3]

Catherine Porter[4]

Abstract

In the last two decades, access to higher education has increased substantially in Latin America. The quantity of new programs available has created concerns about education quality, which has implications for the labor market. We use rich longitudinal data from a Peruvian cohort tracked from ages 8 to 26 (the Young Lives study) to analyze the profile of students enrolled in different 'types' of higher education, and to explore the returns to higher education before and during the COVID-19 crisis. We find evidence of stratification at higher education level: (a) students from the wealthiest households tend to enroll in universities (as opposed to technical institutes), and choose majors and institutions with the highest income rewards; (b) students with higher levels of cognitive skills and socio-emotional competencies tend to attend better quality universities; (c) there are hidden gender gaps: females are more likely to enroll in majors that are the least rewarded in the labor market. In the 2020 labor market, by age 26 we find that: (d) pre-COVID, positive returns to higher education are only observable for those that attended better quality universities; (e) during the pandemic, higher education became a protective factor, with the income premium being higher for everyone that attended this education level; (f) the male income premium doubled during the pandemic.

[1] The authors would like to thank ESRC-GCRF for funding this study. We are grateful to Santos Zhu and Alessandra Hidalgo for excellent research assistance.

[2] Group for the Analysis of Development. Corresponding author: asanchez@grade.org.pe

[3] University of Oxford, Oxford Department of International Development.

[4] Lancaster University.

1. Introduction

Enrolment in higher education has increased substantially in Latin America and the Caribbean (LAC) in the last two decades (Avitabile, 2017). Although this is a pattern observed across the world (Schofer and Meyer, 2005), the speed of the regional expansion has been among the highest in the world.[5] The number of programs as well as its variety has increased to accommodate the increase in demand, especially in countries such as Chile and Peru (Ferreyra, 2017). Although it is well documented that individuals from wealthier families are more likely to enroll in higher education in developing countries (also known as educational stratification),[6] there is less evidence about how students' life course trajectories lead to students sorting into different 'types' of higher education—e.g., by the quality of education provided by the institution and/or by major/degree—, and what are the labor market returns of these choices.[7]

We study these behaviors in the context of Peru, a middle-income country with large inequalities in access to higher education,[8] where evidence suggests that the quality of higher education has reduced over time (Diaz, 2009). We exploit the recently implemented regulatory changes aimed at securing that educational institutions fulfill minimum requirements (accreditation process, "licenciamiento")[9] to identify quality institutions. We seek to contribute in two ways. First, we document the nature of stratification in higher education in Peru. For this we consider two proxies for higher education quality: (a) whether the university is eligible to receive accreditation; (b) according to the average income of graduates by major/degree and name of institution. Measuring the quality of education is a very difficult task.[10] Since the process of accreditation was completed a few years after the enrolment decision was made in our data, eligibility for accreditation captures several characteristics of the institution linked to

[5] Gross enrolment in LAC went up from 17% in 1991 and 21% in 2000 to 40% in 2010 (Avitabile, 2017).
[6] See, for instance, Wu (2010) and Marteleto et al. (2016) with evidence from China and Brazil respectively, and Sánchez and Singh (2018) with evidence from Ethiopia, India, Peru and Vietnam.
[7] Hastings et al. (2013), Yamada and Oviedo (2017), and MacLeod et al (2017), provide evidence about the labour market returns of certain features associated to higher education quality in the context of Chile, Peru, and Colombia (respectively), however these studies lack detailed information about students' characteristics prior to enrolment, i.e., the life-course trajectory of students.
[8] In Peru, gross enrolment in higher education (including universities and technical institutes) has increased steadily between 2001 and 2019, from 38% to 77%—as a proportion of the population aged 17 to 21 years of age, see ESCALE (n.d.). However, enrolment falls to 16% and 27% among the extreme and non-extreme poor, versus 87% for the non-poor.
[9] Evidence suggests that quality has been reducing over time for a long period (Diaz, 2009), however the process might have exacerbated due to the issue of Law 882 ("Ley de Promoción de la Inversión en Educación") in 1996, which allowed the creation of for-profit (private) universities that currently represents 38% of the universities (SUNEDU, n.d,). Between 1996 and 2009 the university admission ratio (proportion of admitted students by number of applicants) increased from 30% to 45%, whereas the proportion of full-time professors reduced from 47% to 25% (Castro and Yamada, 2012). Both changes might be indicative of a worsening of the quality of the education provided.
[10] Even a value-added approach—comparing entry with exit scores—does not by itself allow to measure quality, a point made by Haimovich Paz (2017).

quality—from infrastructure, faculty, and program characteristics to reputation (MacLeod et al., 2017) —as perceived by families, but it does not capture the impact of accreditation per se. A second contribution of this study is to measure the returns of educational choices in the labor market, in two phases: before and during the COVID-19 pandemic. This distinction is important because, as many other countries in the world, the Peruvian government implemented national and regional lockdowns throughout 2020 to deal with the pandemic.

To carry out this analysis, we used longitudinal data from the older cohort of the Young Lives study in Peru, a cohort that has been tracked between 2002 and 2020, from ages 8 to 26 years. The richness of the Young Lives data and its longitudinal nature of the data allows us to establish a profile of enrolment in higher education at age 22 (in 2016) according to individual and household endowments observed prior to the enrolment decision, including the role of household wealth and skills,[11] whereas the sample designs allow to capture the geographical and socio-economic diversity of the country. Similarly, the data allows us to obtain credible estimates of the returns to higher education of this cohort at age 26. In 2020, a phone survey was administered to the participants, through which data on labor market participation and earnings at the age of 26, just before and during the COVID-19 pandemic, was collected. This data provides a unique opportunity to measure the earnings equation controlling for skills accumulated before entering higher education, thus dealing with one of the main sources of endogeneity in the Mincerian equation. It also allows us to measure whether returns to higher education changed during the COVID-19 pandemic.

We find evidence of educational stratification by family wealth and students' skills. This is consistent with the skills formation model proposed by Cunha et al. (2015). Although access to higher education has largely democratized in Peru, students from wealthier families and that have accumulated more cognitive skills and socioemotional competencies over the life course are able to attend better quality institutions. This translates in large income gaps in the labor market, which are already visible by the age of 26.

The rest of the paper is divided as follows. Section 2 explains the data available from the Young Lives study and how this data is merged with administrative data to characterize higher education institutions. Section 3 provides descriptive statistics that are informative of the profile

[11] Existing evidence from the Young Lives study for Peru shows large disparities in tertiary enrolment according to household wealth and area of residence (Sánchez and Singh, 2018). The wealth gradient persists even after adjusting for skills accumulated during earlier stages of the life cycle. However, these results do not incorporate the role of education quality. Given marked differences in quality, it is not obvious that all higher education institutions across the spectrum have a positive return in the Peruvian labor market.

of young people enrolled in higher education. Sections 4 and 5 explain our empirical strategy and report our main results, respectively. Section 6 discusses our findings, and Section 7 concludes.

2. The Young Lives study

The Young Lives study (YLS) is a multi-country study investigating the causes and consequences of childhood poverty. YLS collects data in four Low-and-Middle-Income countries (LMICs): Ethiopia, India (Andhra Pradesh and Telangana), Peru and Vietnam. In each country, it collects information for a *younger cohort*, born in 2001-2, and an *older cohort*, born in 1994-5. This study uses data from the Peruvian *older cohort* (Jones, 2018; Boyden, 2018a and 2018b; Woldehanna et al., 2018; Sánchez et al., 2018). Data for this cohort was collected from 20 districts in Peru, randomly selected from the universe of districts in the country, excluding the wealthiest 5%. The districts included in the sample cover urban and rural areas in the three climatic regions of the country: coast, highlands, and Amazonian rainforest. In each district, 25 to 50 families that had a child aged approximately 8 years old were selected to be part of the study. More information about the sampling design can be found in Escobal and Flores (2008).

The original sample of the *older cohort* included 714 children. After the first visit in 2002, cohort participants were re-visited at ages 12, 15, 19, and 22 (in 2006, 2009, 2013, and 2016, respectively). After 15 years of data collection, in the last visit the participation rate was 84% (n=597), equivalent to an annual attrition rate of approximately 1%, relatively low compared to other longitudinal studies in developing countries (Sánchez and Escobal, 2020). During each of these visits, information was collected about socio-economic characteristics of the household—including access to basic services, household infrastructure, and holding of durable goods, used to calculate the household wealth index (Briones, 2017)—, parents' level of education and native tongue; at the child level, cognitive skills were measured through the Peabody Picture Vocabulary Test (PPVT) and a mathematics test designed by the Young Lives team (Cueto et al., 2009; Cueto and León, 2012) and socioemotional competencies were measured through scales designed to measure self-efficacy and self-esteem (Yorke and Ogando, 2018).

To measure enrolment in higher education we used information from the education history module administered in 2016 (round 5), which contains self-reported information about enrolment in formal education programs in every year between 2013 and 2016—complemented, when necessary, with the same module from round 4 that covers years 2009 to 2013. We include in the definition of higher education programs provided by technical institutes and

universities, which typically take 2-3 and 5 years for completion.[12] Our definition excludes postsecondary education provided by "productive technical centers" (known as CETPROS), because these programs do not require having completed high school. Based on this information, we find that 57% of the sample had been enrolled in higher education by the age of 22. We did not make a distinction between complete and incomplete degrees because, typically, at age 22 many students are still attending classes—especially at the university level. The education history module also contains information about the name of the higher education institutions attended by the participant, which we used for matching purposes, as discussed in detail below.

In 2020, the YLS implemented the 'Listening to Young Lives at work COVID-19 phone survey', designed to measure aspects related to the short-term impact of the COVID-19 crisis on cohort participants. For this purpose, all participants that had been visited in round 5 and with up-to-date contact information (n=514) were considered to participate.[13,14] The phone survey took place in three phone calls administered in June-July, August-October, and November-December (n=489). From this survey we used information related to labor market participation and income activities of YLS participants in the period prior to the pandemic (January-February 2020) and in the last 7 days prior to the phone call (Tuc et al., 2021).

Matching higher education choices with administrative data

We used information about the name of the last higher education institution attended by each YLS participant to characterize his/her educational choice in three dimensions: whether he/she attended a technical institute or a university, the quality of the education received (for those that attended university), and the average earnings from graduates that either attended the same institution or studied the same major/degree.

To proxy education quality, we used information from *Superintendencia Nacional de Educación Superior Universitaria* (SUNEDU), a government institution created in 2015 with the objective of "protecting the right of students to receive a quality university education to improve their professional competencies". As part of its duties, SUNEDU oversees that all

[12] Although the information is self-reported, the YLS team in Peru verified if the name of the institution mentioned by the participant indeed provided the level of education specified.
[13] YLS had planned a sixth visit in 2020, which was ultimately postponed due to the COVID-19 pandemic. Before this occurred, in the second semester of 2019, all participants visited in round 5 in 2016 (n=597) were tracked to obtain up-to-date contact information. In the case of the *older cohort*, 86% of them were found (n=514).
[14] During the execution of the phone survey, 11 participants that were not found in the tracking were added to the sample, increasing the reference sample from 514 to 525.

universities comply with basic quality conditions, including the existence of study plans, an educational offer consistent with these plans, adequate infrastructure and equipment, areas of research, qualified professors (at least 25% full-time), transparency of information, among others.[15] Those universities that fulfill these conditions are given a license to continue operating. By early 2021, SUNEDU concluded the license process. After evaluating all 145 universities in the country, it granted permission to 94 of them to continue operating and denied accreditation to 51 universities, which must close operations in the next few years. In the YLS data, we classify those universities attended by YLS participants according to whether they were eligible to receive a license in 2020 and used this information as a proxy of the quality of the institution. Technical institutes are not part of this classification and, thus, remain as a separate category.

To measure the income of graduates by institution and major/degree, we used information from the education and labor market modules of the Peruvian National Household Survey (ENAHO). Conventionally, ENAHO collects information about educational attainment and annual income (from formal and informal activities) of all household members. Since 2014, ENAHO also collects information about the name of the educational institution attended. We used 2014-2018 ENAHO data (n=85,296) to obtain average earnings of graduates by higher education institutions and major/degree studied in Peru.[16] We classified all these institutions by income quintiles and terciles, focusing on graduates aged 25 to 40 years old. Similarly, majors were classified by income quintiles and terciles for graduates from the same age range. Institutions and majors were classified separately to have a meaningful number of observations in each data cell. The matching with the ENAHO data was successful for 98% of the YLS participants that attended higher education. For sample size considerations, in the rest of the analysis we classify YLS participants enrolled in higher education according to whether they attend institutions that belong to the top quintile of the average graduates' income distribution, and by terciles for the average graduates' income distribution by majors. The main majors observed in the YLS sample are as follows: in the top income tercile, engineering (civil, electronic, mining), law, economics, and army and police forces; in the middle tercile, management, accounting, psychology, and mechanical engineering; and, in the bottom tercile, nursing, computing and 'secretarial'.

[15] https://www.sunedu.gob.pe/8-condiciones-basicas-de-calidad/

[16] Although ENAHO is not a census, using cross-sections from multiple years allow us to observe a very diverse pool of graduates. In fact, we observe average earnings for 95% of the existing institutions in the country (1,208 out of 1,277 universities and technical institutes registered at the Ministry of Education).

3. Enrolment in higher education in the Young Lives sample

We use information from the balanced sample of participants (with complete information in all the relevant variables in rounds 1 to 5, n=499). Close to 6 out of 10 individuals in the sample have been enrolled in higher education by age 22 (57%). In **Table 1** we report descriptive statistics for YLS participants enrolled in higher education. The majority were enrolled in universities (58% compared to 42% in technical institutes) and in private institutions (64%). Most participants that were enrolled in universities went to institutions eligible for accreditation (70%)—virtually all universities that were denied accreditation by SUNEDU at the national level are private, and this is reflected in the sample.

INSERT TABLE 1

On average graduates from universities attended by YLS participants earn more than those from technical institutes (5,593 USD versus 4,849 USD, annualized earnings), however graduates from technical institutes earn more than graduates from universities that are non-eligible for accreditation (4,346 USD), with graduates from universities eligible for accreditation earning the most on average (6,130 USD) (column (3)). These patterns resemble results observed in graduates' earnings distribution at the national level.

There is a strong association between accreditation eligibility and earnings of graduates. Consequently, there are very few YLS participants attending universities non-eligible for accreditation that belong to the top income quintile. Although graduates from technical institutes earn less than those from universities, it is not the case that the distribution of earnings of university graduates dominate that of technical institute graduates. In fact, of those YLS participants enrolled in institutions from the top income quintile, 33% of them attend technical institutes.

In **Table 2** we report how enrolment differs by individual and household characteristics (Column 1). Enrolment is higher for those from urban areas (61%, compared to 47% from rural areas) and whose mother have a higher level of education. Those in the middle and, especially, top tercile of household expenditure report higher levels of enrolment (52% and 74%) compared to the bottom tercile (46%)—the wealth gradient. No differences by gender are observed.

INSERT TABLE 2

In **Table 2**, Panel A, we report enrolment rates distinguishing among those enrolled in technical institutes (25%) and, for universities, according to whether these are eligible or not to receive accreditation (23% and 10%, respectively). Results show that to a large extent differences in enrolment rates by maternal education and household expenditure are explained by enrolment in universities, and, specifically, in universities eligible to receive accreditation—i.e., that fulfill minimum quality criteria. In the case of technical institutes, if anything, enrolment is higher for those from poorer backgrounds, but the difference is not statistically significant.

Some of these patterns are also observed when those enrolled are classified according to the income quintile of the institution and income tercile of the major/degree (**Table 2**, panels B and C, respectively). The wealth gradient is mainly explained by those attending higher education institutions in the top quintile and majors in the top tercile. In terms of the maternal education gap, there is a direct relationship with the quintile of the institution, but not necessarily with the tercile of the major, which suggests that other aspects drive the choice of major. Furthermore, in this case there is evidence of a gender gap, such that females are less likely to attend institutions that belong to the top quintile, less likely to study majors that belong to the top tercile, and more likely to study majors from the bottom tercile. We revisit many of these results in the econometric analysis in Section 5.

4. Empirical methodology

To formally explore the profile of higher education students we report results from three multinomial logit specifications with categories denoted by $0,1..,J$. The baseline category ($j = 0$) correspond to the state in which the individual has not been enrolled in higher education. The remaining categories vary according to the specification:

- **Profile by institutional accreditation**: consider three alternative states: enrolled in a technical institute ($ac_i = 1$), in a university not eligible for accreditation ($ac_i = 2$) and in a university eligible for accreditation ($ac_i = 3$). Eligibility for accreditation is used as a *proxy* for good quality education.

- **Profile by average income of institution graduates**: there are two states besides the baseline: enrolled in an institution (technical institute or university) with an average income of graduates in the first (bottom), second, third or fourth quintile of the income distribution ($iq_i = 1$); and, in the fifth (top) quintile ($iq_i = 2$).

- **Profile by average income of major graduates**: three states (besides the baseline): enrolled in a major with an average income of graduates in the bottom tercile ($it_i = 1$), in the middle tercile ($it_i = 2$), and in the top tercile ($it_i = 2$). In practice, there is a group of YLS participants enrolled in higher education for which we are unable to observe the name of the major (n=55), this group of participants is classified in a fifth group.

It is important to note that individuals are not literally faced with these choices. By the time the data was collected (in 2016) the accreditation model had just started, and no decisions had been made by SUNEDU. This means that our results do not capture the impact of accreditation on educational choices. Instead, we treat eligibility to accreditation as an ex-post revealed indicator of quality. Similarly, individuals might not necessarily be aware whether their institutions/major of choice belong to the top quintile/tercile in the graduates' income distribution. However, from a demand perspective, individuals are likely to apply to certain institutions taking into consideration the information available and their own resources.

Individuals are expected to have priors about the quality and monetary returns of the institutions/majors they choose to apply to. Priors are based on information such as infrastructure, equipment, research reputation, media marketing, the performance of graduates in the labor market, information obtained through a person's own network, among others. These factors are partially observable, especially for well-established institutions (those created before the passing of Law 882 in 1996, which allowed the creation of for-profit universities) and can be used to inform priors. Although information is noisy, those universities perceived to have been of better quality prior to the creation of SUNEDU are likely to eventually have received accreditation. A similar argument can be made about universities, educational institutions, and majors/degrees with the largest income premiums in the labor market: this information should be partially observable, but individuals from wealthier households are likely to be better informed and more able individuals are likely to make a more efficient use of the information available.

Tertiary education is also costly, especially in a system where public and private institutions coexist. Individuals from wealthier households can pay more for education, including tuition fees, food, transportation as well as opportunity costs. In addition, more able individuals are more likely to enroll in higher education, and in more demanding institutions and programs—i.e., they are more likely to be successful at admission exams and/or to fulfill admission requirements. For the two reasons mentioned above (the role of information and the cost of education), enrolment is likely to be driven by socio-economic status and skills.

Based on the availability of information, the probability that an individual belongs to category $j > 0$ is modeled as a function of $x_i = (x_{1,i}, x_{2,i})$, which includes individual and household socio-demographic characteristics ($x_{1,i}$) and measurements of skills ($x_{2,i}$). Specifically, vector $x_{1,i}$ includes sex, age in years (at the time of the interview), current area of residence (urban or rural), mother´s education level, and household location in the expenditure distribution by tercile (measured at age 12); in turn, $x_{2,i}$ includes results in cognitive test scores (raw scores in math and vocabulary, expressed as Z-scores),[17] socio-emotional competencies (self-esteem and self-efficacy, expressed as Z-scores),[18] and educational aspirations, all measured at the age of 12. We simultaneously include measurements of household expenditure and individual skills to incorporate the direct and indirect role that household wealth might have on determining tertiary enrolment, the former through short-term liquidity constraints and the latter through the impact of household wealth on the acquisition of skills over the life course (Carneiro and Heckman 2002; Cunha et al., 2005). The model also controls for a binary variable that takes the value of 1 if the individual has a scholarship and 0 otherwise, however we refrain from interpreting this coefficient because the proportion of the sample with a scholarship is very small (3%).

In the accreditation status model, the probability that an individual belongs to category $j = 1,2,3$ can be expressed as follows:

$$\Pr\left(ac_i = j | X_i, j > 0\right) = \frac{\exp(X_i\beta_j)}{1+\exp(X_i\beta_1)+\exp(X_i\beta_2)+\exp\left(X_i\beta_3\right)} \qquad (1)$$

In the other two models, the probability takes an analogous form. All models are estimated by maximum likelihood.

To explore the returns to higher education, we used information collected in the phone survey administered in 2020, when YLS participants aged 26 years. First, we estimate a linear probability model and focus our attention on the role of accreditation status on the probability of having an adequate job, proxied by having a job and earning at least the minimum wage ($F_i = 1$):

[17] The mathematics test was designed by the Young Lives team, whereas the vocabulary test is the Peabody Picture Vocabulary Test, a test designed to measure receptive vocabulary (Cueto et al., 2009; Cueto and León, 2012). We used the raw scores, normalized with the mean equal to zero and variance equal to one.

[18] We used the generalized self-efficacy and generalized self-esteem scales, which have been validated in the psychological literature (Yorke and Ogando, 2018). The two scales use a Likert response scale. To calculate each scale, the following procedure was followed: (i) the negative statements were put in reverse order; (ii) all the statements were standardized, with mean and variance equal to zero and one; (iii) the average was calculated for each individual.

$$\Pr(F_i = 1|X_i, ac_i) = (ac_i = 1)\gamma_1 + (ac_i = 2)\gamma_2 + (ac_i = 3)\gamma_3 + X_i\beta \quad + \mu_i \quad (3a)$$

Similarly, for choice of major,

$$\Pr(F_i = 1|X_i, it_i) = (it_i = 1)\gamma_1 + (it_i = 2)\gamma_2 + (it_i = 3)\gamma_3 + X_i\beta \quad + \mu_i \qquad (3b)$$

Second, to calculate the earnings that these individuals receive in the labor market as a function of their education choices, we estimate (separately) earnings equations for accreditation status and choice of major:

$$\ln y_i = (ac_i = 1)\delta_1 + (ac_i = 2)\delta_2 + (ac_i = 3)\delta_3 + X_i\beta \quad + \mu_i \qquad (4a)$$

$$\ln y_i = (it_i = 1)\delta_1 + (it_i = 2)\delta_2 + (it_i = 3)\delta_3 + X_i\beta \quad + \mu_i \qquad (4b)$$

where $\ln y_i$ is the log of self-reported monthly earnings. In all these models, X_i is a vector that allows to adjust for sex, age in years, area of residence (urban or rural) and, importantly, for cognitive skills and socio-emotional competencies observed at the age of 12. Controlling for lagged skills is important as it allows to control for the main source of omitted variable bias in the estimation of the earnings equation. To test the protective role of higher education, we estimate two versions of all these models, one considering labor market outcomes just before the beginning of the COVID-19 pandemic (in January-February 2020, asked retrospectively), and another set of results using information about the situation of the individual at the time of the interview (in November-December 2020).

5. Results

a) Profile of students in higher education

Table 3 reports the marginal effects from the accreditation status model, the multinomial logit, which considers the following categories: enrolled in a technical institute, in a university not eligible for accreditation, and in a university eligible for accreditation (columns 1, 2, and 3, respectively). In all cases, the baseline includes those individuals never enrolled in higher education. Results show that those enrolled in technical institutes mothers who did not complete higher education, and are less likely to belong to the upper terciles of household expenditure, compared to those not enrolled. In contrast, those enrolled in university are more likely to belong to the top tercile of household expenditure by 11 pp—the result is the same regardless of eligibility for accreditation.

INSERT TABLE 3

University enrolment is more likely for those with higher cognitive skills and socio-emotional competencies (relatively to those not enrolled in higher education): a one standard deviation increase in math and vocabulary test scores is associated with an increase in enrolment in universities eligible for accreditation by 10 pp and 6 pp, respectively. Similarly, a one standard deviation increase in self-efficacy increases the probability of attending a university eligible for accreditation by 11 pp. These relationships are not observed for enrolment in technical institutes, nor for enrolment in universities ineligible for accreditation, which suggests educational stratification. In addition, having aspired for university education is associated with an increase in university enrolment (by 7 or 8 pp depending on eligibility to accreditation) but is only statistically significant for ineligible universities.

In **Table 4**, we report results of the enrolment profile according to whether the institution in which the participant is enrolled (be that a university or a technical institute) belongs to the top quintile. Individuals from the top tercile of household expenditure are 9 pp more likely to attend institutions from the top quintile (relatively to those not enrolled in higher education). Similarly, improvements in cognitive skills (math and vocabulary) and socio-emotional competencies (self-efficacy) by one standard deviation, and aspiring for university education, are associated with increases in the probability to attend top income institutions by 9 pp, 7 pp, 15 pp and 13 pp, respectively. None of these patterns are observed for enrolment in institutions below the top income quintile, which is again suggestive of substantial educational stratification. One key difference with respect to previous results is that, in this case, we find evidence of a gender gap: females are more likely to enroll in institutions that are not in the top quintile (by 10 pp).

INSERT TABLE 4

In **Table 5** we report results of the enrolment profile according to whether the major belongs to the bottom/middle/top income tercile. As before, the baseline corresponds to those not enrolled. In this case, we find a more diverse profile. Males are more likely to enroll in majors that belong to the top income tercile while females in majors that belong to the bottom income tercile—in both cases relatively to those not enrolled in higher education. Those from the top tercile of household expenditure are 20 pp more likely to enroll in majors from the top tercile, and those from the middle tercile of household expenditure are 18 pp more likely. Similarly, improvements in math test scores predict an increase in the probability to attend majors from the middle and top tercile. In this profiling we do not detect differences in socio-emotional competencies. We discuss the implications of all these results in the next section.

INSERT TABLE 5

b) The returns to higher education

Previous results show evidence of stratification in the higher education market. To complement these findings, we investigate the actual returns to higher education at the age of 26 years in terms of having an adequate job (defined as having a job and earning at least the minimum wage), and, for those with a job, measuring difference in the income premium (see **Table 6** and panels A and B, respectively). For higher education, we consider the information previously reported, which does not distinguish between complete and incomplete degrees.

Looking in first instance at the labor market observed prior to COVID-19, we find that having attended a technical institute increases the probability of having an adequate job, by 10 pp (column 2). The point estimate for attending a university eligible for accreditation is almost identical to that observed for technical institutes but is not statistically significant, whereas the point estimate for universities not eligible for accreditation is close to zero. Having studied a major from the top tercile increases the probability of having an adequate job by 15 pp (column 1). Conditional on having a job, we observe an income premium of about 23% for those individuals that attended universities eligible for accreditation, and no income premium for those not eligible universities, in fact the point estimate is negative in this case—but statistically insignificant (column 6). For those that attended technical institutes the point estimate of the income premium is 9% but is statistically insignificant. Those that studied a major from the top tercile are rewarded with a 27% income premium (column 5).

When the same models are estimated during the period of the COVID-19 pandemic, the point estimates of all categories of higher education in the earnings equation go up (column 8) and even point estimates that were negative turn positive, suggesting that young people with higher education degrees were in a more secure position during the economic crisis, regardless of the accreditation status of the institution attended and the major choice. There is even suggestive evidence that having attended a university without accreditation arose as a protective factor during the crisis—the coefficient goes from -18% to 19% but remains statistically insignificant. Enrolment in accredited universities is associated with an income premium of 46%, substantially larger compared to the pre-pandemic period; a similar pattern is observed in the point estimate of the income premium for technical institutes, but the result is not statistically significant. Similarly, the income premium of majors from the top tercile almost double, going from 27% to 52%. Finally, another distinctive feature of our results is that the gender gap

(against females) substantially increased during the pandemic, both in terms of having an adequate job and income premium.

INSERT TABLE 6

6. Discussion

Our results provide evidence of substantial stratification in the higher education market in Peru, measured in two dimensions: family wealth, and skills accumulated during childhood and adolescence. Both dimensions are linked through the skills formation model (Cunha et al., 2005). In this setup, family wealth determines access to higher education through two channels: short-term liquidity constraints, which affects resources available to pay for tuition fees and other opportunity costs; and its long-term impact on skills formation, required to access good quality programs—i.e., to be successful at entry exams. Indeed, our results are aligned with the main prediction of this model: while close to 6 out of 10 respondents are enrolled in higher education at age 22, which suggest a large democratization in access, those that attend better quality universities arrive at this stage having accumulated more cognitive skills and socio-emotional competencies during childhood and adolescence. In addition, the wealth gradient is explained by the choice between technical institutes and universities, while there are few differences in the socio-economic status of those that attend technical institutes versus those that are not enrolled. Similar evidence is also found when profiling students according to the average income of their institution graduates—as reported also in Sánchez (2019).

The choice of major also provides evidence of stratification. Those students that accumulated more math skills are more likely to be enrolled in majors from the middle and top income terciles, which include engineering (both middle and top) as well as economics, law, management, accounting, etc. In addition, those students from the middle and top terciles of household wealth are more likely to study majors from the top income tercile. Given that household expenditure was measured when the individual aged 12 years and that the model already control for skills, this provides further evidence that access to the better rewarded careers is constrained for individuals from poorer backgrounds.

Our results also provide evidence of gender gaps against women that are not obvious at first sight, when looking at tertiary enrolment at the aggregate level. Although females are as likely as males to be enrolled in tertiary education and to be enrolled in accredited universities, they are more likely to be enrolled in institutions with a relatively lower reward for their graduates in

the labor market, and, similarly, are less likely to be studying the most rewarded majors (in particular, STEM degrees).

Moving on the returns to higher education choices in the labor market, in normal times (before the beginning of the pandemic) we found virtually no income premium of attending a non-accredited university, while for those attending a better-quality university there is a sizable premium. The pre-pandemic results resemble findings from Yamada and Oviedo (2017) using data from ENAHO, the national household survey, however our estimates adjust for cognitive and non-cognitive skills, which substantially alleviate concerns related to the potential existence of omitted variable bias. Nevertheless, our results must also be interpreted with cautious. As the value-added literature in higher education suggests, the fact that graduates from better quality universities receive a premium in the labor market can occur for many reasons (MacLeod and Urquiola, 2015; MacLeod at al., 2017). While this group of universities might be adding value to their students, the results we observe are most likely a combination of other factors, including signaling (i.e., the student reveals his/her ability by completing a degree in a prestigious institution) and networking effects. Although accreditation might act as a signal for potential employers, it is important to stress that our results are unlikely to be informative of the impact of 'accreditation' per se, because the YLS labout market data is informative of a moment in time when the accreditation process had not yet concluded.

Although we do not detect an income premium for those YLS participants that attended a technical institute, they are more likely to have an adequate job (compared to those that only completed high school). Furthermore, there are technical institutes that provide education that leads to well remunerated major/degrees (see **Table 1**). It is possible that the reason why we are not able to detect an income premium for technical institutes is because we are not able to discriminate by quality in this case—as we do with universities.

The COVID-19 pandemic hit the Peruvian economy very hard, due in part to the institution of a lengthy national lockdown (from mid-March to June 2020) followed by localized lockdowns from July to September. According to the Peru Central Bank, the real GDP decreased by 11,5% in 2020, one of the largest recessions in the Latin American region. During the time of the pandemic, we find that the point estimates of the income premium of higher education increased for all sub-groups (including technical institutes and universities without accreditation). This suggests that being a higher education graduate acted as a protective factor during the crisis. Unfortunately, results also show that females were severely affected by the crisis, as the gender gap in earnings increased.

Overall, our results suggests that, to reduce the reproduction of inequalities, a combination of policies is required at different levels: (a) improve the quality of basic education (to enhance skills); (b) increase the scope of scholarships and credit for higher education for those less able to afford it (to deal with short-term liquidity constraints); (c) include technical institutes in the accreditation process (to improve the information available for students); (d) provide more information to prospective students about the returns to education by institution and degree. Indeed, the Peruvian government has taken some steps in the right direction in recent years, including the implementation of a extended school day reform at the secondary level ("Jornada Escolar Completa"), a scholarship program for higher education—based on merit, "Beca 18"—, the implementation of an information program about the returns to education by institution and degree ("Ponte en Carrera"), among others. However, given the inequalities observed, there is a need to scale-up these initiatives and to put special emphasis on the incentives and information provided to females and to adolescents from poorer backgrounds.

7. Conclusions

Access to higher education has substantially increased in the LAC region. To satisfy the increase in demand, there has been a surge in the creation of new institutions and programs, but this has occurred in the absence of adequate regulation. We focus on Peru, a middle-income country that has undergone exactly this experience, and that only recently has concluded an ex-post process to provide accreditation to those universities that fulfill minimum quality requirements. Using longitudinal data from a cohort tracked from ages 8 to 26, we find substantial evidence of educational stratification, such that students from wealthier households and that accumulated more skills over the life course are more likely to enroll in better quality institutions. Our data allow us to verify that this leads to inequalities in the labor market that have amplified since the COVID-19 pandemic began. Finally, we uncover gender gaps against females in both educational choices and labour market returns—the latter also amplified during the pandemic. To reduce the reproduction of inequalities, a combination of policies is required to improve the quality of basic education, increase the scope of scholarships and credit for higher education, include technical institutes in the accreditation process, and provide more information to prospective students, especially females.

References

Avitabile, C. (2017). The Rapid Expansion of Higher Education in the New Century. In Ferreyra, M.M., Avitabile, C., Botero, J., Haimovich, F., & Urzúa, S. (2017). *At a Crossroads: Higher Education in Latin America and the Caribbean*. Directions in development—Human Development. World Bank, Washington, D.C.

Boyden, J. (2018a). Young Lives: an International Study of Childhood Poverty: Round 2, 2006. [data collection]. 3rd Edition. *UK Data Service*. SN: 6852, http://doi.org/10.5255/UKDA-SN-6852-3

Boyden, J. (2018b). Young Lives: an International Study of Childhood Poverty: Round 3, 2009. [data collection]. 3rd Edition. *UK Data Service*. SN: 6853, http://doi.org/10.5255/UKDA-SN-6853-3

Briones, K. (2017). 'How Many Rooms Are There in Your House?' Constructing the Young Lives Wealth Index. *Young Lives Technical Note* N° 43 Oxford: Young Lives.

Carneiro, P., Heckman, J. (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *The Economic Journal*, 112(482), p. 705-734.

Castro, J. F., & Yamada, G. (2013). Declining Quality Affects Choice: The Peruvian Case. *International Higher Education*, 70(1), 26-27. https://doi.org/10.6017/ihe.2013.70.8715

Cueto, S., León, J., Guerrero, G. & Muñoz, I. (2009). Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives. *Young Lives Technical Note* N° 15. Oxford: Young Lives.

Cueto, S., & León, J. (2012). Psychometric characteristics of cognitive development and achievement instruments in Round 3 of Young Lives. *Young Lives Technical Note* N° 25. Oxford: Young Lives.

Cunha, F., Heckman, J., Lochner, L., Masterov, D. (2005). Interpreting the Evidence on Life Cycle Skill Formation. *IZA Discussion Paper* N° 1675.

Diaz, J.J. (2008). Educación superior en el Perú: tendencias de la demanda y la oferta. In M. Benavides (Ed.). *Análisis de programas, procesos y resultados educativos en el Perú: contribuciones empíricas para el debate*. (p. 83-129). Lima: GRADE

Escobal, J., & Flores, E. (2008). An Assessment of the Young Lives Sampling Approach in Peru. *Young Lives Technical Note* 3, 37p. Oxford: Young Lives.

Ferreyra, M.M. (2017). The supply side of the higher education expansion. In Ferreyra, M.M., Avitabile, C., Botero, J., Haimovich, F., & Urzúa, S. (2017). *At a Crossroads: Higher Education in Latin America and the Caribbean*. Directions in development—Human Development. World Bank, Washington, D.C.

Haimovich Paz, F. (2017). Equity, quality and variety of higher education. In Ferreyra, M.M., Avitabile, C., Botero, J., Haimovich, F., & Urzúa, S. (2017). *At a Crossroads: Higher Education in Latin America and the Caribbean*. Directions in development—Human Development. World Bank, Washington, D.C.

Hastings, J., Neilson, C., Zimmerman, S. (2013). Are Some Degrees Worth More than Others? Evidence from college admission cutoffs in Chile. *NBER Working Paper* N° 19242.

Jones, N., Huttly, S. (2018). Young Lives: an International Study of Childhood Poverty: Round 1, 2002. [data collection]. 6th Edition. *UK Data Service*. SN: 5307, http://doi.org/10.5255/UKDA-SN-5307-3

Lavado, P., Martínez, J., & Yamada, G. (2014). ¿Una promesa incumplida? La calidad de la educación universitaria y el subempleo profesional en el Perú. *BCRP Working Paper* No 21. Lima: Central Bank of Peru.

MacLeod, Urquiola, M. (2015). The Big Sort: College Reputation and Labor Market Outcomes. *American Economic Review*, 105(11): p. 3471–3488.

MacLeod, W.B., Riehl, E., Saavedra, J.E., Urquiola, M. (2017). The Big Sort: College Reputation and Labor Market Outcomes. *American Economic Journal: Applied Economics*, 9(3): p. 223–261.

Marteleto, L., Marschner, M., Carvalhaes, F. (2016). Educational stratification after a decade of reforms on higher education access in Brazil. *Research in Social Stratification and Mobility*, 46, 99-111.

Sánchez, A. (2019). Habilidades, formación para el trabajo y subempleo juvenil: un enfoque de ciclo de vida. *Informe Final CIES*. Lima: Consorcio de Investigación Económica y Social.

Sánchez, A., Escobal, J. (2020). Survey attrition after 15 years of tracking children in four developing countries: The Young Lives study. *Review of Development Economics*, 109, 1-21.

Sánchez, A., & Singh, A. (2018). Accessing higher education in developing countries: Panel data analysis from India, Peru, and Vietnam. *World Development*, 109, 261-278.

Sánchez, A., Penny, M., Woldehanna, T., Galab, S., Boyden, J., Duc, L. Thuc. (2018). Young Lives: an International Study of Childhood Poverty: Round 5, 2016. [data collection]. *UK Data Service*. SN: 8357

Schofer, E., & Meyer, J. (2005). The Worldwide Expansion of Higher Education in the Twentieth Century. American Sociological Review, 70 (6), 898-920.

SUNEDU (n.d.). *Informe Bienal sobre la Realidad Universitaria Peruana*. Lima: SUNEDU.

Tuc, L., Revathi, E., Sanchez, A., Favara, M., Penny, M., Porter, C., Woldehanna, T. (2021). Listening to Young Lives at Work: COVID-19 Phone Survey, First Call, Second Call and Third Call, 2020. [data collection]. 3rd Edition. *UK Data Service*. SN: 8678, http://doi.org/10.5255/UKDA-SN-8678-3

Urzúa, S. (2017). The Economic Impact of Higher Education. In Ferreyra, M.M., Avitabile, C., Botero, J., Haimovich, F., & Urzúa, S. (2017). *At a Crossroads: Higher Education in Latin America and the Caribbean*. Directions in development—Human Development. World Bank, Washington, D.C.

Woldehanna, T., Galab, S., Sanchez, A., Penny, M., Duc, L. Thuc, Boyden, J. (2018). Young Lives: an International Study of Childhood Poverty: Round 4, 2013-2014. [data collection]. 2nd Edition. *UK Data Service*. SN: 7931, http://doi.org/10.5255/UKDA-SN-7931-2

Wu, X. (2010). Economic transition, school expansion and educational inequality in China, 1990-2000. *Research in Social Stratification and Mobility*, 28, 91-108.

Yamada, G., Oviedo, N. (2017). Premio a la calidad universitaria en el mercado laboral peruano. In Yamada, G., Lavado, P. (Eds.). *Educación superior y empleo en el Perú: una brecha persistente* (p. 71-89). Lima: Universidad del Pacifico.

Yorke, L., & Ogando, M. J. (2018). Psychosocial scales in the Young Lives Round 4 survey: selection, adaptation and validation. *Young Lives Technical Note* N° 45. Oxford: Young Lives.

Table 1: Characteristics of higher education institutions observed in the Young Lives sample

	n	%	Annual earnings of graduates (in USD)
	(1)	(2)	(3)
Balanced sample	499	100	
Enrolled in higher education	286	57	5,281
Not enrolled	213	43	n.d.
Conditional on enrolment:			
Type of management			
Private	184	64	5,082
Public	102	36	5,639
Type of higher education			
Technical institute	120	42	4,849
University	166	58	5,593
Eligibility for accreditation			
Universities non-eligible	50	30	4,346
Universities eligible	116	70	6,130
Income quintile by institution			
Top quintile	147	51	6,384
Below top quintile	139	49	4,114
Income tercile by major			
No information about major	54	19	4,678
Bottom tercile	57	20	4,689
Middle tercile	110	38	5,366
Top tercile	65	23	6,156
Income quintile by institution and type of higher education			
In top income quintile by institution			
Technical institute	49	33	6,237
Universities non-eligible	2	1	5,683
Universities eligible	96	65	6,473
Below top income quintile by institution			
Technical institute	71	51	3,891
Universities non-eligible	48	35	4,290
Universities eligible	20	14	4,482
Income tercile by major and type of higher education			
Technical institute			

No information about major	33	28	4,485
Bottom tercile	38	32	4,266
Middle tercile	35	29	5,004
Top tercile	14	12	6,900
University			
No information about major	21	13	4,980
Bottom tercile	19	11	5,535
Middle tercile	75	45	5,535
Top tercile	51	31	5,952

Note: information obtained from matching the name of the higher education institution attended by the YLS participants with information from SUNEDU and ENAHO.

Table 2: Enrolment in higher education in the Young Lives sample: total and by characteristics of the institution

	n	Enrolled in higher education (%)	Panel A. Enrolment by eligibility to accreditation			Panel B. Enrolment by average income of graduates (institution)		Panel C. Enrolment by average income of graduates (major/degree)			
			Technical institute (%)	University not eligible (%)	University eligible (%)	Below top quintile (%)	Top quintile (%)	Bottom tertile (%)	Middle tertile (%)	Upper tertile (%)	Not known (%)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total	499	57	25	10	23	28	30	11	22	13	11
Gender											
Male	260	57	24	10	24	24	34	7	21	17	12
Female	239	58	25	10	23	33	25	16	23	8	10
Gap		*1*	*1*	*0*	*-1*	*9***	*-9***	*9****	*2*	*-9****	*-2*
p-value		*0.85*	*0.75*	*0.99*	*0.91*	*0.02*	*0.04*	*0.00*	*0.62*	*0.00*	*0.59*
Area of residence											
Rural	116	47	25	5	16	23	23	8	21	6	12
Urban	383	61	24	12	25	29	31	13	23	15	10
Gap		*14****	*-1*	*7***	*9***	*6*	*8**	*5*	*2*	*9***	*-2*
p-value		*0.01*	*0.78*	*0.05*	*0.05*	*0.21*	*0.10*	*0.16*	*0.69*	*0.01*	*0.62*
Maternal level of schooling											

(i) Mother has no formal education	178	43	20	7	15	22	21	8	14	8	13
(ii) Mother has primary or secondary	270	61	28	10	23	31	30	13	24	14	10
(iii) Mother has higher education	51	88	18	20	51	33	55	18	39	26	6
Gap (iii) - (i)		*45****	*-2*	*13****	*36****	*11*	*34****	*10***	*25****	*18****	*-7*
p-value		*0.00*	*0.68*	*0.01*	*0.00*	*0.10*	*0.00*	*0.04*	*0.00*	*0.00*	*0.16*
Household expenditure											
(i) Bottom tercile	164	46	27	6	14	26	21	14	22	3	7
(ii) Middle tercile	172	52	24	7	20	26	26	8	15	16	13
(iii) Top tercile	163	74	21	18	36	32	42	12	30	20	12
Gap (iii) - (i)		*28****	*-6*	*12****	*22****	*6*	*21****	*-2*	*8*	*17****	*5*
p-value		*0.00*	*0.21*	*0.00*	*0.00*	*0.21*	*0.00*	*0.64*	*0.12*	*0.00*	*0.13*
<u>Conditional on enrolment:</u>											
Cognitive test scores											
Math (mean)	286	0.40	0.16	0.34	0.68	0.26	0.54	0.08	0.57	0.66	0.09
PPVT (mean)	286	0.37	0.14	0.38	0.59	0.21	0.51	0.15	0.43	0.58	0.20
Socioemotional competencies											
Aspirations for higher education	286	87	81	94	91	83	92	86	89	91	82
Self-esteem (mean)	286	0.07	0.06	0.09	0.07	0.02	0.12	0.12	0.11	0.03	0.00
Self-efficacy (mean)	286	0.14	0.08	0.19	0.18	0.06	0.22	0.14	0.12	0.18	0.13

Note: information obtained from matching the name of the higher education institution attended by the YLS participants with information from SUNEDU and ENAHO. * p < 0.10, ** p <0.05, *** p < 0.01

Table 3: Profile of higher education students by accreditation status of institution (multinomial logit specification)

	Technical institute	University not eligible for accreditation	University eligible for accreditation
	(1)	**(2)**	**(3)**
	coef/se	**coef/se**	**coef/se**
Has a scholarship	0.243**	-0.055	0.072
	(0.107)	(0.100)	(0.060)
Female	0.029	0.006	0.032
	(0.031)	(0.024)	(0.044)
Age (in years)	-0.002	0.076***	-0.034
	(0.030)	(0.028)	(0.037)
From urban area	0.069*	-0.023	-0.108
	(0.037)	(0.052)	(0.068)
Mother has primary or secondary	0.047	-0.008	0.003
	(0.058)	(0.028)	(0.048)
Mother has higher education	-0.004	0.054	0.100*
	(0.065)	(0.057)	(0.053)
HH expenditure middle tercile	-0.043	0.010	0.017
	(0.048)	(0.054)	(0.050)
HH expenditure top tercile	-0.075	0.108**	0.107**
	(0.061)	(0.042)	(0.042)
Math test (z-score)	0.010	0.010	0.104***
	(0.031)	(0.013)	(0.025)
PPVT (z-score)	0.009	0.002	0.062**
	(0.029)	(0.018)	(0.026)
Self-esteem (Z-score)	0.012	-0.003	-0.046
	(0.056)	(0.019)	(0.042)
Self-efficacy (Z-score)	0.047	0.039	0.112*
	(0.067)	(0.042)	(0.062)
Aspirations for university	-0.032	0.072**	0.080
	(0.045)	(0.029)	(0.055)
Number of observations	499	499	499

Note: marginal effects reported. Baseline category: not enrolled in higher education. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Profile of higher education students by average income of institution graduates (multinomial logit specification)

	Institution below top income quintile	**Institution in top income quintile**
	(1)	**(2)**
	coef/se	**coef/se**
Has a scholarship	0.038	0.230**
	(0.135)	(0.093)
Female	0.101**	-0.036
	(0.049)	(0.056)
Age (in years)	0.123***	-0.086***
	(0.038)	(0.033)
From urban area	-0.013	-0.039
	(0.062)	(0.051)
Mother has primary or secondary	0.036	0.007
	(0.053)	(0.041)
Mother has higher education	0.095	0.068
	(0.089)	(0.083)
HH expenditure middle tercile	-0.022	0.003
	(0.048)	(0.049)
HH expenditure top tercile	0.045	0.090*
	(0.064)	(0.047)
Math test (z-score)	0.037	0.087***
	(0.024)	(0.024)
PPVT (z-score)	0.008	0.066***
	(0.034)	(0.023)
Self-esteem (Z-score)	-0.020	-0.012
	(0.052)	(0.051)
Self-efficacy (Z-score)	0.040	0.153**
	(0.065)	(0.062)
Aspirations for university	-0.013	0.129**
	(0.052)	(0.055)
Number of observations	499	499

Note: marginal effects reported. Baseline category: not enrolled in higher education. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Profile of higher education students by average income of major graduates (multinomial logit specification)

	Major in bottom income tercile	**Major in middle income tercile**	**Major in top income tercile**
	(1)	**(2)**	**(3)**
	coef/se	**coef/se**	**coef/se**
Has a scholarship	0.149***	0.099	0.058
	(0.055)	(0.077)	(0.076)
Female	0.109***	0.033	-0.078*
	(0.034)	(0.047)	(0.041)
Age (in years)	0.031	0.008	0.010
	(0.029)	(0.032)	(0.023)
From urban area	0.019	0.035	-0.020
	(0.041)	(0.075)	(0.039)
Mother has primary or secondary	0.040	0.037	0.008
	(0.033)	(0.052)	(0.038)
Mother has higher education	0.110	0.070	0.056
	(0.080)	(0.098)	(0.048)
HH expenditure middle tercile	-0.073***	-0.143***	0.181***
	(0.025)	(0.044)	(0.058)
HH expenditure top tercile	-0.046	-0.036	0.195***
	(0.035)	(0.030)	(0.059)
Math test (z-score)	-0.014	0.096***	0.050***
	(0.020)	(0.021)	(0.015)
PPVT (z-score)	0.004	0.020	0.025
	(0.022)	(0.020)	(0.017)
Self-esteem (Z-score)	-0.005	0.038	-0.035
	(0.036)	(0.059)	(0.041)
Self-efficacy (Z-score)	0.049	0.022	0.061
	(0.033)	(0.047)	(0.041)
Aspirations for university	0.033	0.058	0.027
	(0.026)	(0.039)	(0.037)
Number of observations	499	499	499

Note: marginal effects reported. Baseline category: not enrolled in higher education. Marginal effects from the following group is omitted: "major not known". * p < 0.10, ** p <0.05, *** p < 0.01

Table 6: Labour market outcomes

	PANEL A: OLS specification. Dependent variable: 1 if has an adequate job (earns at least minimum wage), 0 otherwise				**PANEL B: OLS specification.** Dependent variable: log- monthly earnings			
	Before the pandemic		During the pandemic		Before the pandemic		During the pandemic	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	coef/se	**coef/se**	**coef/se**	**coef/se**	**coef/se**	**coef/se**	**coef/se**	**coef/se**
Female	-0.191***	-0.209***	-0.292***	-0.302***	-0.091*	-0.121**	-0.191*	-0.218**
	(0.039)	(0.039)	(0.056)	(0.059)	(0.046)	(0.047)	(0.091)	(0.095)
Age (in years)	-0.042	-0.041	-0.067	-0.069	-0.052	-0.024	-0.026	-0.009
	(0.036)	(0.039)	(0.043)	(0.043)	(0.058)	(0.062)	(0.095)	(0.097)
From urban area	0.025	0.018	0.038	0.041	-0.077	-0.057	-0.106	-0.063
	(0.076)	(0.074)	(0.064)	(0.066)	(0.137)	(0.139)	(0.088)	(0.090)
Studied major/degree from bottom tercile of income dist.	-0.061		0.048		-0.086		0.293*	
	(0.071)		(0.076)		(0.109)		(0.140)	
Studied major/degree from middle tercile of income dist.	0.085		0.050		0.093		0.169	
	(0.075)		(0.085)		(0.092)		(0.149)	
Studied major/degree from top tercile of income dist.	0.147*		0.171**		0.266**		0.521***	
	(0.071)		(0.062)		(0.120)		(0.087)	
Attended technical institution		0.101**		0.080		0.086		0.161
		(0.041)		(0.068)		(0.057)		(0.109)
Attended university not eligible for accreditation		0.016		0.083		-0.171		0.193
		(0.093)		(0.067)		(0.148)		(0.119)
Attended university eligible for accreditation		0.104		0.067		0.232*		0.464***
		(0.070)		(0.079)		(0.116)		(0.089)
Math test (Z-score)	-0.012	-0.008	0.043	0.047	-0.025	-0.024	0.053	0.052
	(0.025)	(0.025)	(0.029)	(0.030)	(0.031)	(0.038)	(0.059)	(0.059)

PPVT (Z-score)	0.059*	0.062**	0.007	0.011	0.095*	0.094*	-0.000	-0.014
	(0.028)	(0.028)	(0.033)	(0.035)	(0.050)	(0.053)	(0.064)	(0.059)
Self-esteem (Z-score)	0.042	0.034	-0.003	-0.007	-0.050	-0.043	-0.064	-0.065
	(0.053)	(0.054)	(0.060)	(0.063)	(0.070)	(0.069)	(0.107)	(0.113)
Self-efficacy (Z-score)	0.037	0.044	0.023	0.026	0.111	0.112	0.060	0.069
	(0.056)	(0.053)	(0.072)	(0.074)	(0.077)	(0.072)	(0.083)	(0.094)
Constant	1.467*	1.449	1.969*	2.011*	8.220***	7.598***	7.570***	7.163***
	(0.791)	(0.857)	(0.956)	(0.961)	(1.222)	(1.297)	(2.018)	(2.045)
Number of observations	402	402	402	402	316	316	275	275
Adjusted R2	0.080	0.070	0.106	0.103	0.036	0.053	0.063	0.061

Note: * p < 0.10, ** p <0.05, *** p < 0.01