Unbundling Returns to Degrees and Skills: Evidence from Postsecondary Education in Colombia*

Matías Busso	Juan Sebastián Muñoz	Sebastián Montaño
Inter-American Development Bank	University of Illinois at Urbana Champaign	University of Maryland at College Park

November 22, 2019

Abstract

This paper shows that returns to education are not enough to capture all the returns to human capital. Using longitudinal data of all college graduates in Colombia, we estimate labor market returns to postsecondary degrees and to various skills –including literacy, numeracy, foreign language, field-specific, and non-cognitive skills. Graduates of longer programs, of private institutions, and of schools with higher reputation earn higher wages. Even after controlling for all the characteristics of the degree, a one standard deviation increase in each skill predicts an average wage increase of two percent. Returns to skills vary along the wage distribution, with tenure, with the field of specialization and the type of job obtained immediately after graduation.

Keywords: Returns to skills, returns to education, numeracy, literacy, foreign language, specific skills, non-cognitive, Colombia.

JEL codes: I20, I24, J24, J31

*Corresponding authors: mbusso@iadb.org, munozmo2@illinois.edu, and montano@umd.edu.

The data used in this research was provided by the Ministry of National Education throughout the Agreement # 1465 (of 2017). We are especially grateful with the team of the Labor Observatory for Higher Education, who collects and manage the administrative records used in this project. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the University of Illinois, the University of Maryland, and the Inter-American Development Bank, its Board of Directors, or the countries they represent.

1 Introduction

There is a long-lasting tradition in economics of estimating returns to human capital by using the years of education with which the individual enters the labor market. This paper shows that returns to education are not enough to capture all the returns to human capital. We jointly estimate the labor market returns to various types of postsecondary degrees as well as to several types of skills (academic, specific and non-cognitive). We find that for individuals that graduated from similar programs, a one standard deviation increase in the measure of skills yields an average wage premium of 2 percent. Similarly, for students of comparable skills, the market valuation of postsecondary degrees are positive and vary widely with the length and quality of the program.

We combine administrative records from a variety of sources to build a longitudinal data set that follows students from high school, to college, and into the labor market. Students in Colombia are evaluated after high school graduation by means of a mandatory standardized test (analogous to the SAT in the United States) on mathematics, language (Spanish), foreign language, and other subjects. We combine those data with information on students' enrollment and graduation in postsecondary programs. For all college graduates we observe a rich set of characteristics including their field of study and the length of the program. Students who are about to graduate from two- and four-year colleges undergo evaluations on mathematics, literacy and foreign language, as well as on an specific tests related to their field of studies (akin to the subject GRE). Furthermore, after graduation, Colombia's Ministry of Education surveys a subgroup of recent graduates in a follow-up survey that contains measures of non-cognitive skills. We combine all these data with records from the Colombian Social Security Administration that contain information on wages and employment characteristics.

We use an expanded Mincer earnings function to jointly estimate the labor market returns to different types of degrees and skills. We assuage concerns of an ability bias by controlling for a broad range of measures of both baseline and contemporaneous skills as well as quality measures of the different programs. A similar approach was taken recently, for instance, by Saiz and Zoido (2005), Hanushek et al. (2015), and Lindqvist and Vestman (2011).

We first estimate returns to finishing different types of higher education degrees. We distinguish between four types: 1) public two-year programs; 2) private two-year programs; 3) public four-year programs; and 4) private four-year programs. We find that the annual returns to graduating from a four-year private school (instead of a two-year public program) is 6 percent. A one standard deviation in our measure of college reputation carries a wage premium of 3 percent. There is also a large heterogeneity in the returns by fields of study. Similarly to Kirkeboen et al. (2016), we find that the returns to certain fields are as large as the returns to completing a four-year college.

We then estimate that, conditional on the degree, the program and its reputation, an increase in one standard deviation in skills yields an wage return of 2 percent. Returns to education are not enough to capture all the returns to human capital. Returns to skills are fairly homogeneous. The returns to purely academic skills, numeracy and literacy, are 2 and 2.5 percent respectively. Returns to foreign language are 1.6 percent. Returns to specific skills are 2.2 percent and the wage premia to non-cognitive skills are 2.3 percent.[1] Because tests scores are imperfect measures of a person's skills, returns could be biased towards zero. We provide instrumental variables estimates that use high school exams as instruments to alleviate the attenuation bias potentially caused by measurement error.

We explore different patterns of heterogeneity in the returns to skills. First, consistent with Farber and Gibbons (1996), we find that the returns to those skills that are less observable to the employer (numeracy and literacy), increase with tenure in the firm while returns to a more easily observable skill (such as foreign language) is constant. We assume that foreign language skills are easier to evaluate in an interview than the skills in numeracy of literacy. Second, individuals realize what we term "returns to specialization." For example, people who work or study in more math-oriented fields and industries have a higher return to numeracy, whereas individuals who graduate from social sciences have a larger return to literacy, and individuals who work in tourism have a higher return to foreign language skills. Third, returns to numeracy are higher for people graduating in most fields of study, and working in most economic sectors. Fourth, we find that women have more returns to foreign language skills than men.

This study contributes to a large literature estimating returns to skills. Most of the previous literature provide estimates for developed countries. Our estimates for Colombia, a developing country with relatively low quality of education, fall within the range of estimates previously found in other studies. Returns to numeracy skills have been found to be in the order of 2 to 20 percent (Levine and Zimmerman, 1995; Murnane et al., 1995; Tyler, 2004; De Coulon et al., 2008; Song et al., 2008; Joensen and Nielsen, 2009; James, 2013; Hanushek et al., 2015). Returns to foreign language have been found to be around 2.5 to 60 percent (Bleakley and Chin, 2004; Saiz and Zoido, 2005; Christofides and Swidinsky, 2010; Azam et al., 2013; Guo and Sun, 2014; Budría and Swedberg, 2015; Di Paolo and Tansel, 2015; Stöhr, 2015). Returns to literacy have been found to be as high as 20 percent (Ishikawa and Ryan, 2002; De Coulon et al., 2008; Fasih et al., 2013; Hanushek et al., 2015; Sanders, 2016). Several papers find that non-cognitive skills are valued in the labor market as much as cognitive skills (Bowles et al., 2001; Heckman et al., 2006; Lindqvist and Vestman, 2011; Dasgupta et al., 2017). Similarly to this literature, we find that field-specific and non-cognitive skills are also as valued in the labor market as academic skills are.[2] To the best of our knowledge we are the first to estimate all these measures jointly.

This paper also contributes to the literature analyzing the heterogeneity in the returns to different types of postsecondary degrees on two dimensions (Hastings et al., 2013; Rodríguez et al., 2015; Kirkeboen et al., 2016; Busso et al., 2017b).[3] First, controlling for a wide range

[1] The estimates of returns to one year of education were obtained using the Integrated Household Survey data (GEIH, according to the Spanish acronym) and are available upon request.

[2] Appendix Table 1 summarizes the results, methodologies and samples used in previous studies.

[3] Several studies review the evidence of returns to education in Latin America. See for instance, Psacharopoulos and Ng (1994), Behrman et al. (2007), Bassi et al. (2012), and Lustig et al. (2012).

of measures of skills and for the quality of the college (as measured by its reputation), we find that the types of degrees as well as the field of study affects dramatically future income. Second, we estimate the returns to college reputation, and find that such measure leads to sizable increases in wages.

The rest of the paper is organized as follows. Section 2 describes the Colombian education system and the data used. Section 3 presents the results for the returns to skills and postsecondary degrees. Section 4 analyzes the robustness and heterogeneity in the estimation of the returns to skills. Section 5 then presents the robustness and heterogeneity in the estimation to different types of postsecondary degrees. Section 6 concludes.

2 Background and Data

Education in Colombia is divided into primary school (first to fifth year), middle school (sixth to ninth year), high school (tenth and eleventh), and postsecondary education. Programs in postsecondary education are divided into college degrees (equivalent to a U.S. bachelor's degree, with a duration of four years) and two-year programs that offer vocational or technical instruction in different fields. We refer to all the institutions in postsecondary education as colleges, regardless of the duration of the program.

During high school, students take classes on mathematics, language (Spanish), and foreign languages as part of the school curriculum. Around 95 percent of schools choose to teach English as a foreign language; therefore, we refer to English and foreign language interchangeably. During postsecondary education, the level and intensity of instruction in these areas depend on the student's major, but most institutions require a minimum of foreign language knowledge as a graduation requirement.

After graduation, students' abilities and qualifications are extremely valuable to find a job. Moreover, a mismatch of occupations and skills is more likely to happen among individuals with lower levels of abilities. A survey implemented in 2013 shows that 67 percent of the firms have employed at least one graduate with less than two years of experience and 73 percent, among these firms, consider that knowledge or specific abilities constitute the main selection criteria to hire a graduate.

Measures of Skills. Since 1980, all high school seniors in Colombia have been evaluated before graduation through a mandatory, high-stakes exit exam (Saber 11). The exam is a requirement for graduation, and its results are sometimes used by colleges to take decisions.[4] The test resembles the SAT in the United States. It evaluates students in several subjects that, for convenience, we divide in two broad areas: (i) general skills, which includes mathematics, reading and language (in Spanish), and foreign language, and (ii) subject skills, which includes biology, philosophy, physics, chemistry, and social sciences. The test is administered in two sessions each of four hours and thirty minutes. It evaluates general knowledge of each

[4] Some colleges establish minimum scores that students must achieve to be considered for admission, while others apply their own admission exams and take Saber 11 only as an enrollment requirement.

subject with roughly 40 questions. The mathematics test measures basic knowledge in algebra, calculus, geometry, probability, and statistics; students must interpret information, design solutions to problems, follow procedures, and justify steps in problem solving. The reading test evaluates the student's ability to understand, interpret, and analyze critically written texts. The language exam tests the ability of the student to communicate in Spanish. The English test evaluates reading, grammar, and vocabulary.

Since 2003, college students who completed at least seventy percent of their coursework have been also required to take a college exit exam (Saber Pro). This exam is a graduation requirement and resembles the GRE, both the general and subject tests. The exam provides a strong signal for students' readiness to make the transition from college into the labor market. It is very common for schools to provide incentives (e.g., public recognition to the best-performing students within each school) to motivate students to aim for high scores because the results are used for calculating school quality indexes that are routinely published by the Ministry of Education.

The college exit exam has two main components; one is general and the other is specific to the field of study from which the student is graduating.[5] The general section includes tests in numeracy, writing, reading, English, and citizenship abilities. Students have four hours and forty minutes to complete the test, which includes a total of 161 questions (35 of numeracy, 35 of reading, 35 of citizenship abilities, 55 of English, and one for writing). The numeracy section evaluates basic mathematics knowledge needed to analyze and solve problems using quantitative methods and procedures. The reading section examines the capacity to read analytically by understanding the text and identifying different perspectives and value judgments. The writing section evaluates the ability to communicate ideas of a particular given topic. The English section focuses on testing the ability to communicate effectively in English. The specific section evaluates basic knowledge in the student's field of study. There is a total of 40 specific exams and students have ninety minutes to answer between 30 and 60 questions included in each of them. The questions, designed by experts in each of the different areas, follow specific standards that assure comparability of the exams. Economics majors, for instance, are suggested to take an Economics Analysis test with micro, macro, and econometrics questions.

In addition, since 2005, the Ministry of Education implemented a follow-up survey to recent college graduates that included measures of non-cognitive skills, with the purpose of tracking and evaluating professional performance. The survey was tested from 2005 to 2008 using a sub-sample of recent graduates that voluntarily completed an online form. Starting in 2011, and until 2014, the survey was implemented in person at the moment of graduation, and over the telephone to collect information from older cohorts.[6] Respondents were asked to answer a 10 to 15 minutes interview that demanded information about five modules: (1) family background; (2) level of skills; (3) life plan; (4) job situation; and (5) identity with

[5] Although the general part of the exam started in 2003, major-specific exams were introduced by a staggered roll out. By the second semester of 2011, all students were taking the general tests, while the specific exams have been applied since then to a large share of the students.

[6] Respondents were contacted one, three, and five years after graduation by the telephone.

the institution they graduated from.

The survey includes questions regarding the use of non-cognitive skills acquired in college and applied to the current job. Former students were asked to rank (from one to four) the use of skills related to communication, persuasion, innovation, and readiness to learn, among others. We use this information to compute a factor analysis index that measures non-cognitive skills among the sub-sample of surveyed graduates.[7] This measure quantifies those non-cognitive skills that have been proven to affect job market performance (Heckman et al., 2006). More information and details about the different components and the index are given in Appendix B, on measures of skills.

Postsecondary Degrees. Similar to the experience in other Latin American countries, postsecondary education in Colombia has expanded dramatically in the last 10 years (Busso et al., 2017a). Postsecondary education in Colombia is delivered in the form of two-year and four-year programs. Two-year programs are mostly technically oriented, and resemble the vocational programs in the United States. Four-year programs, by contrast, are academically oriented and are equivalent to a bachelor's degree in the United States. The Colombian postsecondary education also includes public and private institutions; 30 percent of the postsecondary institutions in the country are public, while the other 70 percent are private. We classify academic programs into four categories: 1) two-year public; 2) two-year private; 3) four-year public; and 4) four-year private. In addition, our data also allow us to identify individuals who have pursued graduate studies.

The degrees that students receive vary, not only in terms of the length of study required, but also in terms of quality, reputation, and tuition cost, among others dimensions.[8] Following MacLeod et al. (2017), we compute a measure of college reputation: for each individual the reputation of her college is calculated as the average of the high school exit exam among graduates in the same college.[9]

Data and sample. The database used in this paper corresponds to a unique data that we assembled in collaboration with the Colombian Ministry of Education. We merged five different administrative datasets:

1. We use test score measures from the college exit exam from 2011 to 2015. The general components are compulsory for each student, but the specific tests are not. The decision

[7] Appendix Table 2 presents the loadings, corresponding to the factor with the largest eigenvalue, for each question used to compute the non-cognitive or non-cognitive measure used herein. We use a different set of non-cognitive questions to build a second measure, which is used later as an instrumental variable. Appendix Table 3 presents the results from the factor model applied to compute this latter measure.

[8] We lack information on the cost of tuition for each program in each university that would allow us to net out the cost of attending each program in calculating returns

[9] In Colombia, as in the case of most countries in Latin America, each program (major) in each college takes its own admissions decisions. This decision is decentralized and puts a heavy weight in the high school exit exam. The measure relies on the assumption that better regarded programs-colleges (i.e. with better reputation) can be more selective and will allow to enter a set of applicants with higher high school exit exams.

on which field-specific tests to take is made by the college and not by the student. So, for instance, students graduating from economics may take the economics-specific test but also the business-specific test if their respective college indicates so. Some institutions do not require any specific exam, so a considerable number of students do not take field-specific tests. For this reason, we only observe field-specific test scores for a subgroup of students.

2. We merge data on the high school exit exam from 1996 to 2013. These data include information on the multiple test scores and, in addition, it include information on the student's municipality, school, and some individual characteristics, such as gender and age.

3. We merge records of all students enrolled in college between 1998 and 2016. These records include yearly information on graduation and enrollment of all students in postsecondary education for all universities and programs, and have detailed information on the program of study, as well as socioeconomic information about the student and her family.

4. We merge longitudinal yearly earnings records for workers who graduated from college. These data include individual earnings from 2011 to 2016 of all the workers who finished any college program after 2001, and who contributed to the Social Security System. Four-digit industry codes, the municipality where the contribution was paid, and establishment identifiers are also included.[10]

5. We merge data on the graduates follow-up survey implemented by the Ministry of Education. We linked the 2013 and 2014 cross sections to obtain a random sample of graduates from cohorts 2011 to 2013. These data corresponds to individuals surveyed in 2013 and 2014, who graduated at or before 2013. Therefore, the sample of individuals we observe with non-cognitive measures represents a subgroup of people with more tenure, compared to the average individual we observed in the full sample.[11]

The final sample includes 363,464 individuals who took the college exit exam between 2011 and 2015, graduated after 2011, and worked formally between 2012 and 2016. The high school and the college exit exams are not fully comparable between their different editions. Hence, all test scores are standardized to have mean zero and standard deviation one within each test's edition. We use three different samples: 1) test takers from 2011 to 2015 that are working formally ($N = 363,330$); 2) test takers with major-specific scores available that are working formally ($N = 155,939$); and 3) test takers that were included in the graduate follow-up survey, for whom we have measures of non-cognitive ability, and who are working formally ($N = 2,401$).[12]

[10] Workers who do not contribute to health or pensions are not included in the data.

[11] In Appendix Table 4 we show that individuals in this survey are older and have more tenure.

[12] Appendix C describes in detail the steps followed to build the data and each sample. Initially, it is worth noting that the larger sample of students contains complete information about wages, socioeconomic background, and high school. The same happens to the sample of students with major-specific scores, which is then a perfect subset of the first sample. However, some observations (134 students) within the survey

Descriptive Statistics. Individuals included in the main sample are on average 27 years old. Sixty percent are female, they live primarily in urban areas (73 percent), and 45 percent belong to low-income households.[13] The majority (82 percent) are graduates of four-year programs, and the main fields of study are Science, technology, engineering, and mathematics (STEM) (26 percent), business and economics (31 percent), social sciences and humanities (16 percent), and health and education (22 percent). Six out of ten work for the service sector, and 28 percent are employed by large firms.[14]

Appendix Table 5 presents a correlation matrix of all the different test score measures in our sample. We present a panel for each sample. Several highlights are in order. First, all correlations are positive and large which, as suggested by Rindermann (2007), indicates that these measures are capturing in part a broader trait such as cognitive ability. There is even a positive correlation between cognitive and non-cognitive (captured by non-cognitive) skills. Second, the largest correlations are found within exams (shown in bold type). This is suggestive of skill complementarities. Third, the correlation within subjects and across time (shown in non-bold type) are large and positive which suggests the existence of self-productivity (Cunha and Heckman, 2007; Cunha et al., 2006). Fourth, the magnitudes of the correlations are very stable across the three panels. Fifth, the largest correlation (0.94) is between the average high school exit exam result and what we call the subject exam result, which corresponds to the average of the high school exit exam excluding the numeracy, language, and English components. This high correlation allows us to replace the average result in the high school exit exam with the "subject" exam in our empirical strategy.

3 Returns to Degrees and Skills

The following equation describes our expanded mincer earnings equation which we use to estimate jointly returns to degrees and skills:

$$\log(W_{ifct}) = P_{ifct}\beta_1 + T_{ifc\tau}\beta_2 + \beta_0\theta_i + \alpha_1 CQ_{ct} + X_i\gamma + \mu_f + \mu_t + \mu_\tau + \mu_s + \varepsilon_{ipct}. \qquad (1)$$

W_{ifct} is the wage of individual i who graduated from field of study f at college c in year t. P_{ifct} is a vector that includes three indicator variables that take the value of one if the individual has a postsecondary degree corresponding to: 1) a two-year private program; 2) a four-year public program; or 3) a four-year private program. The omitted category is two-year public programs. $T_{ipc\tau}$ is a vector of college exit tests scores that includes measures of numeracy, literacy, and English. Individual i's test scores correspond to results observed in

sample do not belong to the larger sample because it was not possible to identify the high school from which they graduated from. We decided to keep these individuals to maintain as many observations as we can in the smaller sample.

[13]We identify low-income households using the Colombian housing stratification system. For purposes of targeting social assistance, all houses in the country are assigned to an economic stratum from one to six, depending on the neighborhood and house. We defined low-income households as those living in the first two strata.

[14]In Appendix Table 4 we present a set of statistics that describe our sample.

the edition τ of the college exit exam. Depending on the sample, vector $T_{ipc\tau}$ also includes measures of major-specific or non-cognitive skills. θ_i is a measure of the initial level of ability. We include the measure for college reputation (CQ_{ct}). X_i is a vector of individual characteristics that include gender, age, age-squared, mother's education, an indicator variable for graduate studies, and socioeconomic stratum.[15] μ_f, μ_t, μ_τ, and μ_s are field of study, cohort, test-edition, and high school fixed effects.

To reduce ability bias concerns, we include a measure of initial ability (θ_i) which is measured by the subject skills components of the high school exit exam.[16] This measure is highly correlated (0.94) with the average of all the tests in the high school exit exam.[17]

Equation 1 resembles an estimation of a *value-added* measure in which the effect on students' knowledge is conditioned on initial knowledge.[18] In our case, however, we estimate the economics return to skills enhanced during college and to types degrees, conditioning on the abilities that each student had at the moment of starting. An advantage of such specification is that it eliminates observed and unobserved confounding elements about the history of parental and school inputs, and, therefore, reduces remarkably the likelihood of suffering of omitted variable bias (Rivkin et al., 2005). The inclusion of θ_i in equation (1) together with the full set of skills measures and fixed effects intends to address possible concerns regarding ability bias of $\hat{\beta}$. We interpret $\hat{\beta}_1$ as the return to skills acquired or enhanced during college –conditional on skills prior to college. $\hat{\beta}_2$ should be interpreted as the economic return to type of degrees conditioning on the skills a student has before and after graduation. Note that estimating these jointly isolates the returns of attending a specific program independently of college quality (which is controlled by the measures of skills and the measure of college reputation).

If our set of control variables is not rich enough to ensure our estimators are unbiased, our resulting estimates may still be informative about the relative returns if the bias is the same for all coefficients in the vectors $\hat{\beta}_1$ and $\hat{\beta}_2$. In other words, similar to methodology underpinning some of the previous related literature (Hanushek et al., 2015; Lindqvist and Vestman, 2011), the difference between coefficients will eliminate the bias, and we will still be able to correctly rank the returns to skills and types of degrees.[19]

[15] Households in Colombia are categorized by a number from one to six depending on the type of neighborhood where they reside. This number is used for tax and redistribution purposes, and it is a very strong predictor of wealth. Stratum one is the poorest whereas six is the wealthiest. Households residing in strata five and six subsidize basic public utilities of households living in strata one and two.

[16] We reserve the general high school skills components of the exam (mathematics, language, and foreign language) to be used as instruments later in the analysis. Results, however, are similar when including these general high school skills in the measure of initial ability.

[17] See Appendix Table 5 for more information about the correlation among the measures of skills.

[18] A considerable amount of papers have used value-added measures to estimate the return to better teachers. A discussion about this model is presented in: Hanushek and Rivkin (2010).

[19] Assume a true model of the form: $W = \alpha + \beta_1 T_1 + \beta_2 T_2 + \theta + u$. If the estimated model is $W = \alpha + \beta_1 T_1 + \beta_2 T_2 + \varepsilon$, where $\varepsilon = \theta + u$, then the probability limits of the difference between the OLS estimators will be: $\text{Plim}(\hat{\beta}_1 - \hat{\beta}_2) = (\beta_1 - \beta_2) + (\hat{\beta}_{\theta T_1} - \hat{\beta}_{\theta T_2})$, where $\hat{\beta}_{\theta T_1}$ is the coefficient of a regression of θ and T_1. Our assumption states that $\hat{\beta}_{\theta T_1} \sim \hat{\beta}_{\theta T_2}$.

Table 1 presents ordinary least squares (OLS) estimates of equation (1). Given that we observe wages for multiple periods, we use the current observed wage as dependent variable.[20] Column (1) presents the estimation in the full sample. Columns (2) and (3) restrict the sample to those individuals with field-specific and non-cognitive skills, respectively.[21] Columns (4) to (6) follow the same format, but additionally include the measure of college reputation; these are our preferred estimates.

Returns to different types of degrees are fairly heterogeneous across programs. Using the point estimates in column (5) and adjusting by the years it takes to graduate from each program, we find that relative to two-year public degrees, an additional year of education increases wages in 4.2 percent for two-year private, 3.9 percent for four-year public, and 5.9 percent for four-year private degrees. The reputation of the program also carries a wage premium of three percent. However, we observe that education does not capture the full extent of human capital. The returns to skills conditional on education are meaningful and comparable to the returns to education. Among the returns to skills, numeracy seems to have the largest return (up to 3.7 percent), except in the survey sample. Literacy, foreign language, non-cognitive, and field-specific skills have a similar return of, roughly, two percent. These results are fairly stable across columns which varies both the specification and the sample used in the estimation.

4 Robustness and Measurement Error

Sensitivity of the Mincer Earnings Equation. We analyze the sensitivity of the point estimates of the expanded mincer equation by estimating alternative specifications, and present the results in Table 2. Columns (1) presents estimates of the returns to degrees unconditionally of the returns to skills and field of study, whereas column (2) includes back field of study fixed effects aiming to control for potential pre-graduation sorting into fields. Columns (3) to (5) present the returns to skills across different samples, but unconditionally of the types of degrees and field of study. In columns (8) to (10) we include field of study fixed effects again aiming to control for potential pre-graduation sorting. Comparing these estimates with those in Table 1 give informative evidence on the degree to which the returns to degrees complement with the returns to skills.

The results in Column (1) are quite different from the returns to types of degrees in column (4) of Table 1. This difference can be attributed to the exclusion of the field of study fixed effects or the measures of skills. Including back the fixed effects in column (4), we recover point estimates very close to those in Table 1. This implies that including the

[20] In Appendix Table 6 we present results with alternative measures such as the average wage observed for each individual, the first observed wage after graduation, and total annual earnings. The results are essentially the same.

[21] Columns (3) and (6) use the information of the follow up survey to college graduates. For more details about this dataset visit the webpage: https://ole.mineducacion.gov.co/portal/A-quienes-aportamos/Investigadores/

Table 1: Returns to Skills and Types of Degrees

	Dependent Variable: log(Current Wage)					
	(1)	(2)	(3)	(4)	(5)	(6)
Two-Years Private	0.086	0.088	0.109	0.080	0.084	0.128
	[0.018]	[0.015]	[0.039]	[0.014]	[0.012]	[0.049]
Four-Years Public	0.210	0.194	0.236	0.174	0.156	0.197
	[0.016]	[0.020]	[0.044]	[0.011]	[0.013]	[0.042]
Four-Years Private	0.268	0.257	0.229	0.249	0.236	0.212
	[0.025]	[0.026]	[0.037]	[0.020]	[0.021]	[0.036]
Graduate Studies	0.158	0.183	0.048	0.156	0.181	0.045
	[0.017]	[0.019]	[0.063]	[0.016]	[0.019]	[0.060]
College Reputation				0.032	0.031	0.061
				[0.004]	[0.005]	[0.023]
Literacy	0.028	0.022	0.018	0.027	0.020	0.017
	[0.001]	[0.002]	[0.007]	[0.001]	[0.002]	[0.007]
Numeracy	0.037	0.028	0.024	0.035	0.025	0.019
	[0.002]	[0.003]	[0.009]	[0.002]	[0.003]	[0.010]
Foreign language	0.020	0.020	0.036	0.016	0.016	0.026
	[0.006]	[0.006]	[0.007]	[0.006]	[0.006]	[0.006]
Field Specific		0.024			0.022	
		[0.004]			[0.004]	
Non-Congnitive			0.023			0.023
			[0.005]			[0.006]
Sample:	Full	Specific	Survey	Full	Specific	Survey
Observations	363,330	155,939	2,401	363,330	155,939	2,401
R-squared	0.192	0.218	0.239	0.194	0.219	0.247
Controls:						
Individual	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Ability	Yes	Yes	Yes	Yes	Yes	Yes
Field of study	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Ordinary least squares estimations. The dependent variable is the natural log of the last observed wage from 2011 to 2016 for each individual. Literacy, numeracy, and foreign language scores are computed from the college exit exam and standardized within each sample using the average and standard deviation of the corresponding test edition. The field specific scores are standardized with respect to the average not only of the test edition but also considering the groups of related majors. The omitted category of postsecondary degrees is 2-3 Years Public, which identifies individuals graduated from technical or vocational tertiary programs taught by public universities. College reputation variable is computed following MacLeod et al. (2017) and then standardized to have mean zero and standard deviation one within each sample. Individual control variables include: gender, age, age squared, socioeconomic stratum, mother's education level, a dummy for graduate studies, a proxy of initial abilities, high school fixed effects, cohort fixed effects, and test edition fixed effects. Initial abilities are proxied using the standardized average of biology, physics, chemistry, philosophy, and social sciences scores computed from the high school exit exam. Field of Study refers to fixed effects of related majors: accounting, agricultural sciences, architecture, arts, business and related, economics, education, engineering, health, humanities, journalism, medicine, natural and exact sciences, nursing, law, psychology, social sciences, and sports. Standard errors are clustered by municipality and in brackets.

measures of skills do not affect the coefficients on the types of degrees.

Similar results are obtained in columns (3) to (5), when we estimate returns to skills unconditionally from the returns to degrees and fields of study. We do see that excluding these controls affects slightly the magnitudes, especially in the return to numeracy skills. However, controlling for potential pre-graduate sorting into fields recovers point estimates similar to those obtained in Table 1. These results as a whole imply that the measures of skills and the measures of types of degrees capture different information when estimating economic returns. In fact, it poses strong evidence suggesting that the returns to human capital cannot be exclusively attributed to returns to education nor to returns to skills, but that a combination of both are jointly important.

Table 2: Returns to Different Types of Postsecondary Degrees

	Dependent Variable: log(Current Wage)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Two-Years Private	0.046	0.078						
	[0.012]	[0.014]						
Four-Years Public	0.127	0.175						
	[0.017]	[0.011]						
Four-Years Private	0.235	0.243						
	[0.011]	[0.020]						
Graduate Studies	0.163	0.160						
	[0.015]	[0.017]						
College Reputation	0.042	0.042						
	[0.005]	[0.006]						
Literacy			0.023	0.022	0.012	0.027	0.022	0.024
			[0.002]	[0.002]	[0.007]	[0.001]	[0.002]	[0.007]
Numeracy			0.055	0.058	0.044	0.034	0.029	0.025
			[0.005]	[0.004]	[0.012]	[0.002]	[0.003]	[0.010]
Foreign language			0.017	0.017	0.024	0.028	0.024	0.042
			[0.003]	[0.003]	[0.008]	[0.006]	[0.006]	[0.007]
Field Specific				-0.003		0.021		
				[0.003]		[0.004]		
Non-Congnitive					0.030			0.023
					[0.006]			[0.005]
Sample:	Full	Full	Full	Specific	Survey	Full	Specific	Survey
Observations	363,330	363,330	363,330	155,939	2,401	363,330	155,939	2,401
R-squared	0.130	0.188	0.118	0.143	0.137	0.177	0.207	0.224
Controls:								
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Ability	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field of Study		Yes				Yes	Yes	Yes

Notes. Ordinary least squares estimations. The dependent variable is the log of the last observed wage for each individual from 2011 to 2016. 2-3 Years Private is a dummy variable for individuals graduated from technical or vocational tertiary programs taught by private universities, 4 > Years Public includes graduates from bachelor programs taught by public universities, and 4 > Years Private includes graduates from bachelor programs taught by private universities. The omitted category of postsecondary degrees is 2-3 Years Public, which identifies individuals graduated from technical or vocational tertiary programs taught by public universities. Columns (4) and (8) include measure of skills –literacy, numeracy and foreign language– computed from the college exit exam. Individual and field of study controls as described in the notes of Table 1. Standard errors clustered at the municipality level and in brackets.

Skills Collinearity. The results in Table 1 include simultaneously multiple test score measures that are highly correlated. As previously discussed in Section 2, there is a high degree of skill complementarity: the level of skills of any given individual in a given subject is positively correlated to the skills of that individual on other subjects. This could be for instance, because having better mathematics skills can allow a student to acquire science skills faster; or having higher literacy skills, by better understanding the underlying structure of language, can affect the capacity of an individual to learn a foreign language. The simultaneous estimation of equation (1) will therefore yield a lower bound of the effect of skills on wages because the coefficient of one test score will bias the coefficient of the other test score. The separate inclusion of the test score measures, however, will estimate the upper bound of such effect (see (Lindqvist and Vestman, 2011)).

Table 3 shows OLS estimates of equation (1) including every measure of skills simultaneously and separately (to save space, these point estimates are presented stacked).[22] In columns (1) to (3) we present the same results as in Table 1, for comparison. Columns (4) to (6) show the point estimates of three and four separate regressions that use equation (1), but include each measure of skills separately.

Using the full sample of workers, we find again that numeracy has the highest return (up to 4.4 percent), whereas literacy and foreign language have returns of 3.8 and three percent. Once again we see that the return to numeracy decreases the most when considering the specific sample, although it remains being the largest. Specific skills show again a point estimate among the highest (3.8 percent) when estimated in the same restricted sample. The coefficients on numeracy and literacy decrease, and suggest again that individuals who score high in generic skills also score high in specific skills. Regarding the estimations using the survey sample, we observe that the estimated returns to cognitive skills increase (compared to column (3)), while the return to non-cognitive skills remains stable (2.3 percent).

Measurement Error Correction. Test scores are just a proxy variable of the true level of an individual's skills. In other words, even though the exit exams are official and well established it is possible that tests scores measure skills with error, biasing the estimated coefficients towards zero (attenuation bias). To account for this we assume a classical measurement error setting and instrument the college exit test scores with the high school exit test scores. We additionally control for initial ability. The identification of this model relies on the assumption that the level of a given skill in high school only affects wages through the skills level in college.

Columns (7) to (11) of Table 3 show the IV estimates. In column (7) we include all the three measures of generic skills simultaneously and instrument them with the mathematics, language, and English result in the high school exit exam. The point estimates in the IV estimation increase considerably for literacy and numeracy, and they are similar in magnitude to those found by Lindqvist and Vestman (2011). The coefficient on literacy increases more than the others, and becomes larger than the coefficient on numeracy. Notice as well that the returns to foreign language decrease and are negative possibly because it estimates a lower bound when estimated jointly with the other measures of skills. In column (8) we

[22]Appendix Table 8 presents the same results in Table 3, but including the point estimates for each type of degrees.

Table 3: Extensions: Estimates' Ranges and Measurement Error

	Dependent Variable: log(Current Wage)										
	OLS						IV: Measurement Error Correction				
	Simultaneous			Stacked			Simultaneous		Stacked		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Literacy	0.027	0.020	0.017	0.038	0.036	0.026	0.075	-0.004	0.105	0.106	0.112
	[0.001]	[0.002]	[0.007]	[0.002]	[0.002]	[0.007]	[0.011]	[0.057]	[0.012]	[0.013]	[0.046]
Numeracy	0.035	0.025	0.019	0.044	0.041	0.027	0.045	0.154	0.087	0.083	0.140
	[0.002]	[0.003]	[0.010]	[0.003]	[0.003]	[0.010]	[0.006]	[0.046]	[0.008]	[0.010]	[0.028]
Foreign Language	0.016	0.016	0.026	0.030	0.031	0.035	-0.002	0.002	0.040	0.041	0.053
	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.007]	[0.004]	[0.027]	[0.006]	[0.005]	[0.022]
Field Specific		0.022			0.038					0.123	
		[0.004]			[0.003]					[0.017]	
Socioemotional			0.023			0.023		0.026			0.022
			[0.006]			[0.006]		[0.007]			[0.006]
Sample:	Full	Specific	Survey	Full	Specific	Survey	Full	Survey	Full	Specific	Survey
Observations	363,330	155,939	2,401	363,330	155,939	2,401	363,330	2,401	363,330	155,939	2,401
Controls:											
Individual & Field	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Types of Degrees	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The dependent variable and measures of skills are as described in the notes of Table 1. Estimations within the full sample use the same specification of column (1) of Table **??**. Likewise, estimations within the specific and survey samples use respectively the specifications of columns (3) and (5) of Table **??**. Stacked results refer to separe regressions for each ability measure. In columns (7) to (9), cognitive skills were instrumented using each time numeracy, literacy and foreign language standardized scores computed from the high school exit exam. Individual and types of degrees area as described in the notes of Table 1. Standard errors clustered by municipality and in brackets.

restrict the sample to the individuals from the follow-up survey of graduates. This column includes our three generic cognitive abilities and the measure of non-cognitive skills, which we instrument with the mentioned scores from the high school exit exam and a score computed from another set of non-cognitive related questions contained in the survey sample.[23] We observe that the returns to non-cognitive skills are similar in magnitude to those found using OLS, while there is a remarkable change for the returns to cognitive skills. Observe that literacy and foreign language returns are close to zero, and the returns to numeracy are up to 15 percent.

Column (9) includes each test score measure estimated separately but over-identified using the scores in mathematics, language, and English in the high school exam jointly as instruments.[24] We find that all coefficients increase and are much larger than those estimated separately by OLS (in column (4) of Table 3) or estimated jointly (in column (7)). The point estimate on foreign language (4 percent) is now positive. Column (10) deals with the measurement error for workers who took the major specific test. Each point estimate corresponds to a separate regression. The results show that, as in the OLS case, the returns to literacy and foreign language do not change largely in this sample (compared to columns

[23] Appendix Table 3 shows the results from the factor model used to compute the instrument for the measure of non-cognitive skills.

[24] We also performed estimations using one instrument at a time. The results do not change.

(9)), whereas the returns to numeracy decrease. All the coefficients increase, compared to column (5), but the increase in specific skills and literacy is remarkable changing from 3.8 to 12.3 percent and from 3.6 to 10.6 percent, respectively.

The last column in Table 3 shows the results, adjusting for measurement error, for workers contained in the follow-up survey. Estimations are obtained from separate regressions for each measure of skills. We over-identified each measure using mathematics, language, and English from the high school exam, as well as a non-cognitive score, computed from the questions contained in the survey of graduates. Literacy and foreign language returns increase (in comparison with column (8)) and are up to 11.2 percent and 5.3 percent, respectively. Returns to numeracy are around 14 percent, and returns to non-cognitive skills stay close to 2 percent.[25]

Taken together, these estimates indicates that numeracy skills have a return that ranges between 1.7 percent and 14 percent, literacy a return between 1.4 percent and 11.2 percent, foreign language skills between 1 percent and 12.6 percent, specific skills from 2.1 percent to 12.3 percent, and non-cognitive skills return around 1.2 percent and 2.6 percent. These ranges include upper (regressions with each test score alone) and lower bounds (regressions with all the measures simultaneously) accounting and not accounting for measurement error.

5 Beyond Average Returns

We now focus on the heterogeneity of returns to degrees and skills. First, we explore heterogeneity across the wage distribution, then we present returns by tenure and, finally, we estimate the returns among fields of study, economic sectors, gender, and firm size.

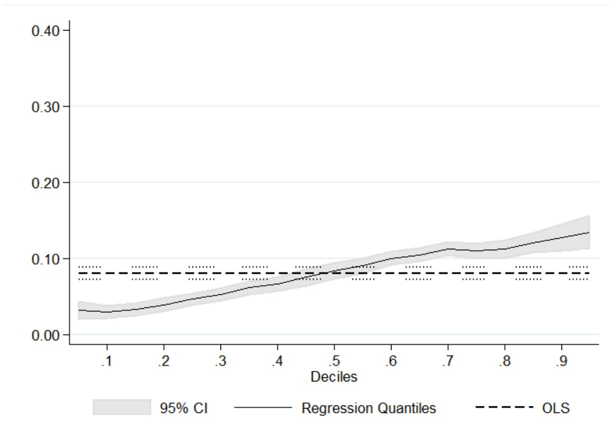
Returns Across Wage Distribution. We estimate conditional regression quantiles for each type of postsecondary degree and present the results in Figure 1. OLS point estimates are plotted to allow comparisons. We use the specification in column (4) of Table 1. Figure 1a shows the results for two-year private, figure 1b for four-year public, and 1c for four-year private.

The results suggest an important degree of heterogeneity in the returns to postsecondary degrees. The three graphs show strictly increasing returns with wages, indicating that postsecondary degrees matter more among higher wage quantiles. Note that the lowest point estimate among four-year private degrees is around 0.8, which is comparable to the point estimates on returns to skills. It means that taking individuals from a two-year public program and placing them in a four-year private, keeping abilities constant, will increase the wage in at least eight percent even if they get a job that pays in the first decile of the wage distribution. A similar increase in wage would be achieved if those same persons stayed in the two-year public program but increased their numeracy skills in one standard deviation.

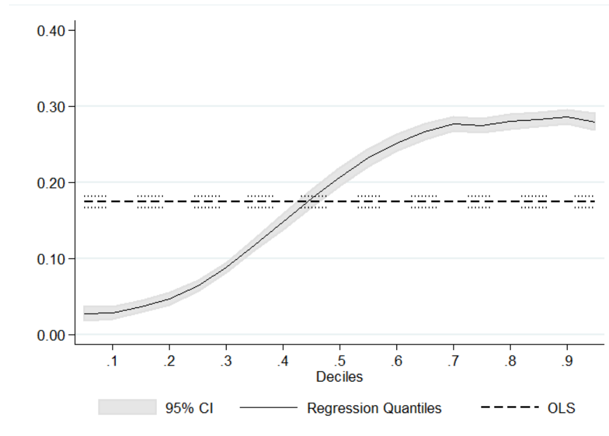
[25]We additionally estimate a latent variable model that implements alternative estimation methods to deal with attenuation bias. The results are presented in appendix D.

Figure 1: Returns to Types of Degrees at Conditional Quantiles of the Wage Distribution

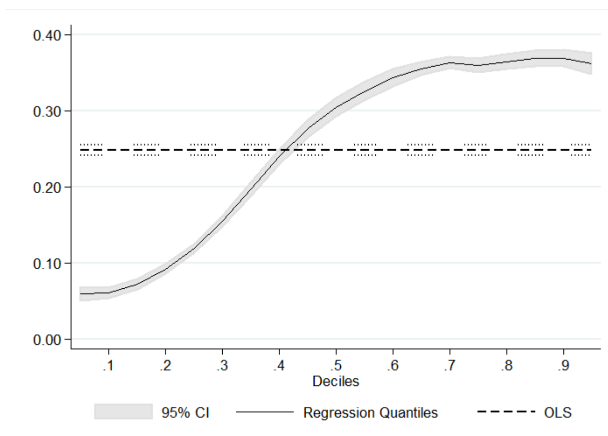
(a) Two-Year Private



(b) Four-Year Public



(c) Four-Year Private



Notes. The solid lines represents the regression quantiles using Koenker and Bassett (1978) estimator. The dashed lines correspond to an OLS specification. Quantile and OLS estimations for 2-3 years private, 4 years public and 4 year private programs used specification (4) in Table 1. Figures 1a, 1b, and 1c used the full sample of students who graduated from 2011 to 2016. OLS standard errors are clustered at the municipality level. Regression quantiles standard errors computed using 20 bootstrap replications. Confidence intervals of 95% are presented for all estimates.

In Figure 2 we present the results of conditional regression quantiles for each measure of skills.[26] We use again the main specification as in column (4) of Table 1 and contrast the quantile point estimates with the OLS.[27] Figures 2a, 2b, and 2c use the full sample of students who took the college exit exam between 2011 and 2016. Figure 2d uses the reduced sample of students with information about field-specific tests scores. Figure 2e shows the results within the sample of surveyed workers with measures of non-cognitive skills.

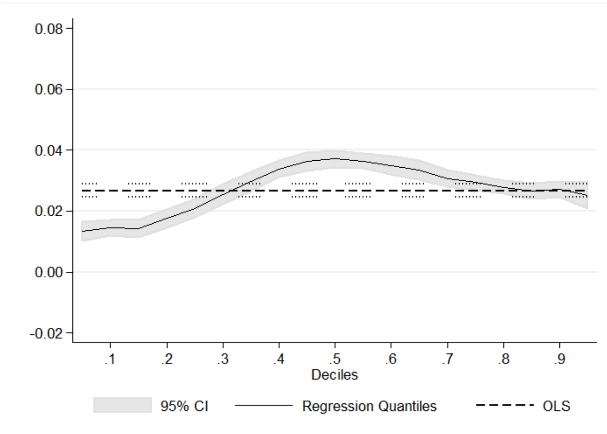
[26]We also estimate unconditional regression quantiles as in Firpo et al. (2009). Results are very similar and available upon request.

[27]For the case of specific and non-cognitive skills we use specifications (5) and (6) of Table 1, estimating the models in the corresponding restricted samples.

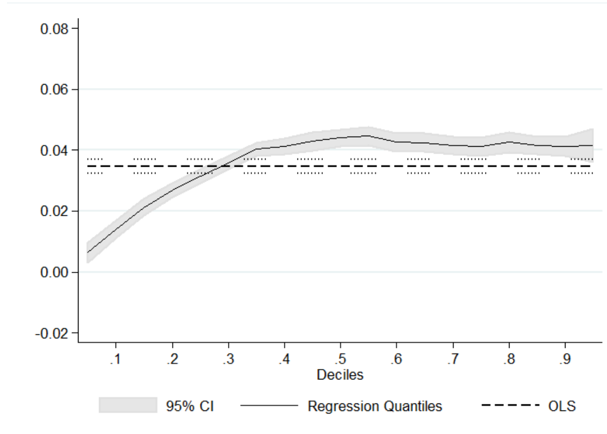
The returns to all the measures of skills are mostly positive along all the deciles. They are lowest among people in the lowest percentiles, and increase monotonically until, roughly, the 40th percentile. The returns to foreign language skills, however, are strictly increasing in the whole distribution, ranging from zero to 0.2. The returns to numeracy remain stagnant beyond the 40th percentile, although these are the highest returns we observe. The returns to literacy and major-specific skills decrease slightly beyond the 50th percentile, but such decrease is not statistically significant. The returns to non-cognitive skills are not precise because of the small sample size. However, they increase until the 20th percentile, remain steady until the 80th percentile, and increase again for the top percentiles. One implication of these results is that inequality in the distribution of skills can explain part of the observed wage inequality in Colombia.

Figure 2: Returns to Skills at Conditional Quantiles of the Wage Distribution

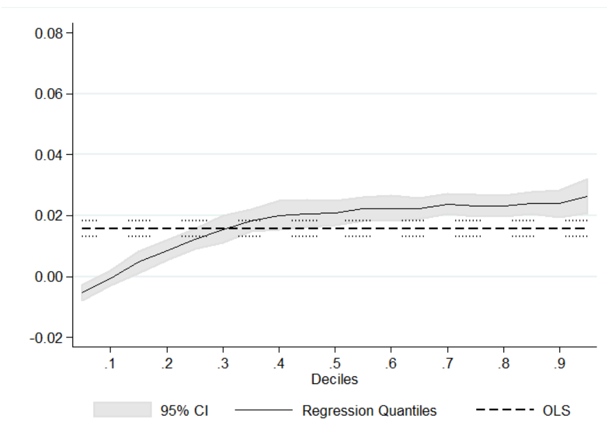
(a) Literacy



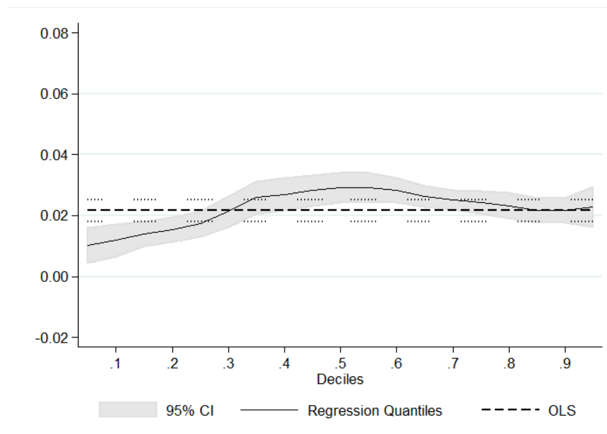
(b) Numeracy



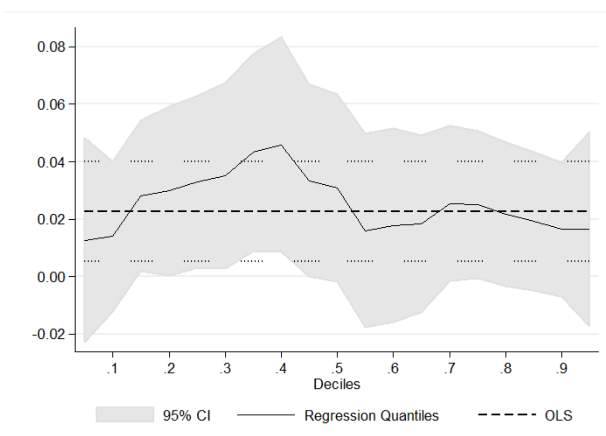
(c) Foreign Language



(d) Field Specific



(e) Non-Cognitive

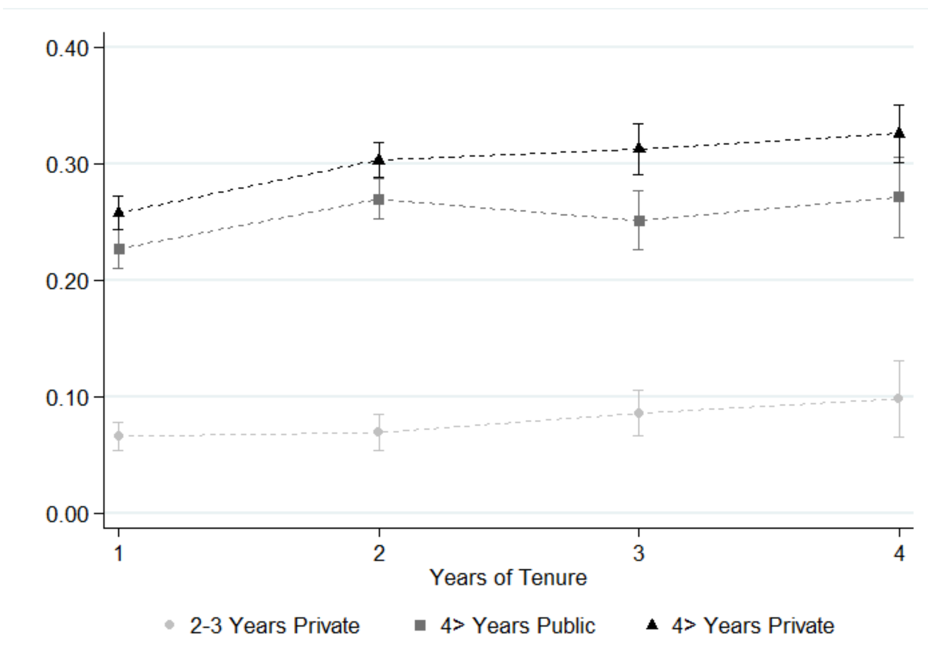


Notes. The solid lines represents the regression quantiles using Koenker and Bassett (1978) estimator. The dashed lines correspond to an OLS specification. Quantile and OLS estimations for numeracy, literacy and foreign language used specification (6) in Table 1. Figures 2a, 2b, and 2c used the full sample of students who graduated from 2011 to 2016. Figures 2d and fig:qregses used, respectively, the sample of students with major specific test scores available and the sample of individuals in graduate follow-up survey. OLS standard errors are clustered at the municipality level. Regression quantiles standard errors computed using 20 bootstrap replications. Confidence intervals of 95% are presented for all estimates.

Returns Across Tenure. In some economics models, wages increase throughout time as a function of the capability of the employer to observe workers' abilities (Farber and Gibbons, 1996). Skills that are easily observable by the employer (such as foreign language) are predicted to have higher returns in the first year on the job, but then stay constant. The returns to less observable skills (such as numeracy or literacy), on the contrary, are expected to increase as the employer updates her beliefs about the worker.[28] Notice that the worker's type of postsecondary degree is very observable by employers and thus, following the previous argument, returns to degrees should not vary much with tenure.

We test that hypotheses between tenure and returns by exploiting the longitudinal feature of the data. Estimations use the same specification of column (4) in Table 1, but among workers with one, two, three, and four years of tenure. Figures 3 and 4 present the results. In Figure 3 we plot the returns to postsecondary degrees across years of tenure. Slight increases in the returns are observed, but the confidence intervals suggest that these slight increases are not in fact significant. We interpret this as evidence in favor of (Farber and Gibbons, 1996), suggesting that the returns to education degrees should not increase with tenure if they are observable for employers.

Figure 3: Returns to Degrees by Years of Tenure



Notes. The plotted circles, squares and triangles represent the point estimates of 2-3 year private programs, 4 years public programs and 4 years private programs, respectively. Separate regression were run among workers with different years of tenure to the estimate simultaneously the returns to different types of degrees. Estimations used the same specification as in column (4) of Table 1. The dependent variable is the log(Wage) for each year of tenure. Confidence levels of 95% are presented for all point estimates.

Returns across tenure for each measure of skills are presented in Figure 4. We additionally

[28]We assume that foreign language skills are more observable in a face-to-face interview –which is very common in a hiring process–, whereas literacy and numeracy skills are more difficult to observe.

include the results among students with field-specific scores for comparability. Non-cognitive skills are not included because of the small sample size of workers who have available scores for this measure. The returns using the complete sample of students are plotted with circles, and the corresponding confidence intervals are represented with a dark area. These results show that numeracy and literacy are increasing with tenure. However, the returns to numeracy increase much more steeply, growing from 1.5 percent to 3.6 percent. The returns to foreign language, on the contrary, do not increase as much; they remain relatively steady across years of tenure. We interpret this as evidence in favor of Farber and Gibbons (1996) model.

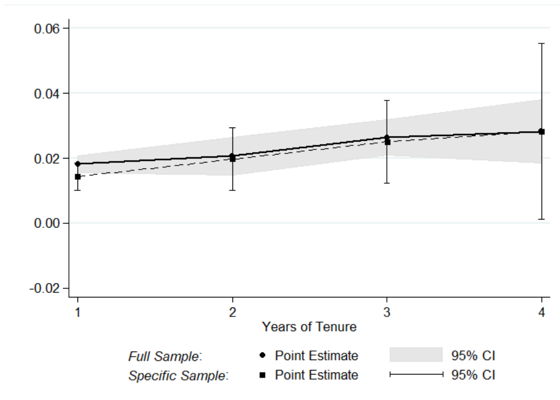
We find similar patterns if we restrict the sample to students with available field-specific tests scores. Estimations within this sample use specification (5) in Table 1. Returns are plotted using squares, with whiskers representing confidence intervals. Using this sample we are able to estimate returns to major-specific skills, and observe that these do not change across years of tenure. We explain this by considering that during hiring processes employers may apply tests that allow them to know their applicants level of specific abilities. On the other hand, applicants could also reveal their specific skills during interviews in order to increase their probability of being hired. Thus, specific skills could be more observable for employers than the rest of skills - at least in the labor market for college graduates.

Returns to Specialization: Field of Study and Economic Sector. Returns to degrees and skills can in part reflect specialization. For instance, individuals with better mathematics skills can choose careers that place greater emphasis on those abilities. Then, they can find jobs that value more their skills, receiving higher payments for greater levels. We explore whether the magnitude of returns varies with the person's specialization (either in the field of study or in the economic sector). Estimations of returns across different areas of study use specification (4) and (5) in Table 1, without including field of study fixed effects. We exclude the estimation of returns to non-cognitive skills because of low sample sizes.

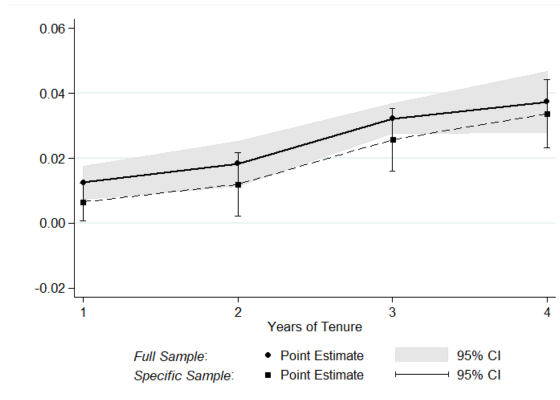
In Table 4 we present the returns to degrees and skills across groups of similar postsecondary programs or study areas. We estimate the returns for graduates of STEM, business and economics, social sciences and humanities, and health and education.

Figure 4: Returns to Skills by Years of Tenure

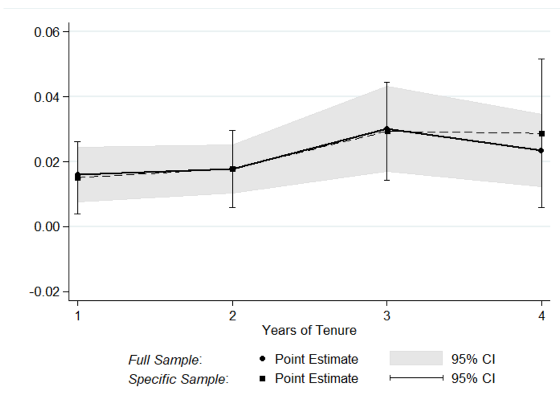
(a) Literacy



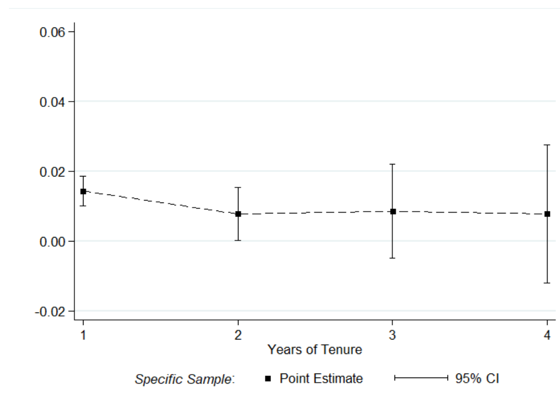
(b) Numeracy



(c) Foreign Language



(d) Field Specific



Notes. The plotted circles and squares represent the point estimates for each measure of skills using, respectively specification (6) and (8) of 1. Separate regression were run among workers with different years of tenure to the estimate simultaneously numeracy, literacy, foreign language and major-specific skills. The dependent variable is the log(Wage) for each year of tenure. Confidence levels of 95% are presented for all point estimates.

We find three main results. First, the positive gradient in the returns to degrees (by which 4-year degrees pay more than 2-year degrees and graduates from private school earn more than those of public schools) is observed in all fields of study. However it is more pronounced in STEM, business and social sciences. Second, literacy and numeracy skills matter in all fields. In that sense, these basic skills are always remunerated probably because they are highly transferable and ubiquitous across fields. Third, there is some evidence of positive returns to specialization in Table 4. Specific skills tend to be homogeneous across fields. Of course, different specific skills would be tested for each program. To shed some light on what those specific skills capture we estimate the model in the full sample without including specific skills (odd columns). Numeracy skills are have the highest return in STEM, business and economics for but not in social sciences and humanities; for these individuals, largest return is to literacy skills. Workers with STEM degrees have the lowest return to literacy skills (2.3 percent). Graduates from health and education degrees have similar returns across all skills. Furthermore, we see that including field-specific skills decreases remarkably the returns to numeracy for most fields suggesting a large complementarity between specific and numeracy skills.

Table 4: Heterogeneous Effects: Study Area

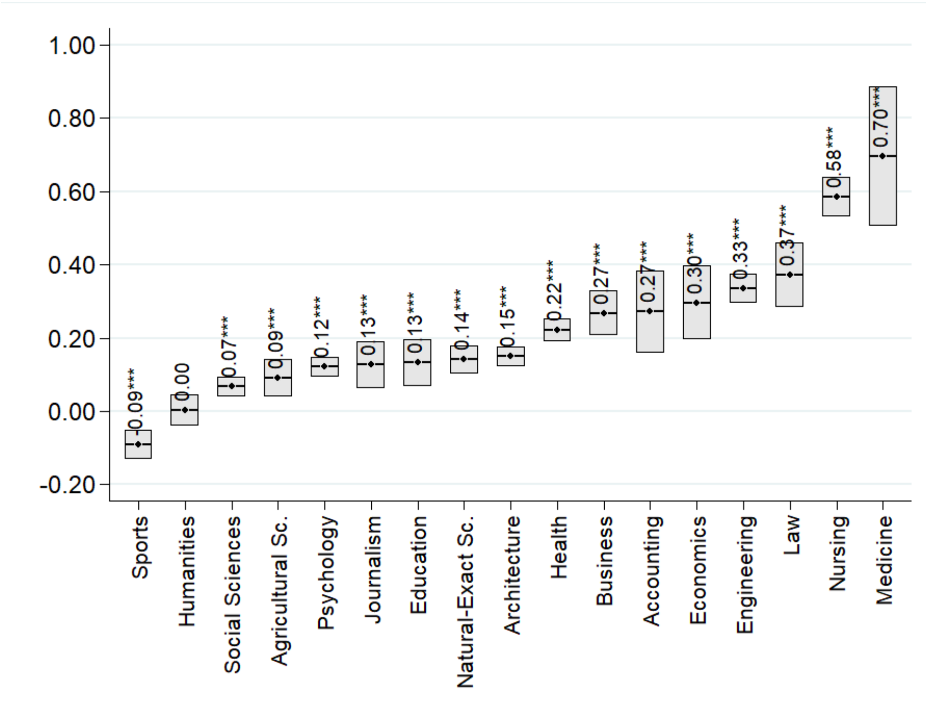
	Dependent Variable: log(Current Wage)									
	Study Area:									
	STEM		Business and Economics		Social Sciences and Humanities		Health and Education		Other	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Two-Year Private	0.062	0.052	0.029	0.031	-0.036	0.307	0.131	0.170	0.114	0.194
	[0.022]	[0.029]	[0.014]	[0.020]	[0.038]	[0.094]	[0.032]	[0.033]	[0.032]	[0.116]
Four-Year Public	0.234	0.218	0.091	0.078	0.359	0.522	0.228	0.205	0.140	0.173
	[0.024]	[0.025]	[0.026]	[0.028]	[0.028]	[0.068]	[0.025]	[0.021]	[0.029]	[0.086]
Four-Year Private	0.311	0.314	0.204	0.187	0.427	0.595	0.216	0.201	0.168	0.251
	[0.039]	[0.038]	[0.012]	[0.021]	[0.039]	[0.069]	[0.023]	[0.023]	[0.018]	[0.072]
Graduate Studies	0.085	0.115	0.177	0.238	0.163	0.186	0.142	0.144	0.042	-0.035
	[0.033]	[0.032]	[0.021]	[0.020]	[0.018]	[0.021]	[0.025]	[0.029]	[0.034]	[0.079]
College Reputation	0.024	0.023	0.052	0.051	0.044	0.049	0.012	0.009	0.017	0.020
	[0.004]	[0.003]	[0.013]	[0.013]	[0.003]	[0.007]	[0.007]	[0.007]	[0.007]	[0.013]
Literacy	0.023	0.020	0.033	0.024	0.032	0.033	0.026	0.015	0.013	0.024
	[0.002]	[0.002]	[0.002]	[0.004]	[0.004]	[0.004]	[0.003]	[0.004]	[0.005]	[0.013]
Numeracy	0.038	0.023	0.042	0.036	0.019	0.013	0.021	0.018	0.025	0.039
	[0.002]	[0.005]	[0.004]	[0.005]	[0.004]	[0.008]	[0.007]	[0.008]	[0.006]	[0.014]
Foreign language	0.010	0.012	0.020	0.019	0.012	0.003	0.023	0.020	0.004	-0.000
	[0.005]	[0.004]	[0.008]	[0.009]	[0.008]	[0.011]	[0.011]	[0.012]	[0.007]	[0.010]
Field Specific		0.024		0.024		0.022		0.022		-0.003
		[0.007]		[0.004]		[0.004]		[0.005]		[0.008]
Sample:	Full	Specific	Full	Specific	Full	Specific	Full	Specific	Full	Specific
Observations	95,270	40,668	112,184	42,572	56,568	23,697	78,526	43,690	20,782	5,312
R-squared	0.213	0.266	0.251	0.300	0.232	0.341	0.316	0.365	0.242	0.459
Controls:										
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Types of Degrees	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The dependent variable and measures of skills are as described in the notes of Table 1. Estimations using the full sample follow the same specification of column (4) of Table 1, but within the set of individuals belonging to the population of interest –defined by the categories on top of the table. In the study area vector, the STEM samples in column (1) includes individuals graduated from engineering, mathematics and natural sciences. Column (2) includes individuals graduated from business and economics, columns (3) from social sciences and humanities, columns (4) from health and education sciences, and column (5) from agronomy and arts. Standard errors clustered at the municipality level are in brackets.

Using model 1 we can additionally estimate heterogeneity among fields of study. Figure 5 shows the field of study fixed effects estimated in Column (4) of Table 1.[29] The returns to different fields of study vary dramatically. Sports have the lowest returns while medicine is more than 70 percent higher. These results suggests that what a person studies is as important as in which place those studies take place.

[29]The field of study fixed effects estimated using the specifications in the other columns of Table 1 are very similar to the ones reported in Figure 5.

Figure 5: Returns to Fields of Study



Notes. Every plotted dot is the point estimate for each category of the field of study fixed effects, taking Arts as base for comparison. Confidence levels of 95% are represented by boxes for all point estimates. The estimation uses the same specification as in column (4) of Table 1. Standard errors were clustered at the municipality level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

In Table 5 we then present the returns among workers in different economic sectors. To save space we present only the results for the sample of individuals for whom we have a measure of field-specific skills. Results for the full sample are very similar.[30] Again, we find clear returns to specialization, in particular for workers in the trade and tourism industries, in which the highest return is to foreign language skills (6.9 percent and 7.5 percent, respectively). Field-specific skills have a sizable and precise return in manufacturing, trade, and services. Numeracy skills show the biggest return for manufacturing and services. We also observe some heterogeneity in the gradient of returns to degrees. In the tourism and in the trade sector the differences between degrees is much more attenuated. Among the other economic sectors the difference between types of degrees remain relatively constant.

[30]The results for the full sample are available from the authors upon request.

Table 5: Heterogeneous Effects: Economic Activity

	Dependent Variable: log(Current Wage)								
	Economic Activity:					Gender:		Firm Size:	
	Manufacture	Trade	Services	Turism	Retail	Female	Male	Small	Large
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2-3 Years Private	0.009	-0.021	0.118	0.077	0.135	0.087	0.078	0.070	0.070
	[0.037]	[0.046]	[0.012]	[0.052]	[0.043]	[0.013]	[0.016]	[0.015]	[0.013]
4> Years Public	0.187	0.126	0.192	0.056	0.182	0.154	0.175	0.191	0.257
	[0.023]	[0.054]	[0.017]	[0.069]	[0.046]	[0.011]	[0.021]	[0.020]	[0.019]
4> Years Private	0.243	0.189	0.275	0.222	0.299	0.237	0.245	0.257	0.326
	[0.022]	[0.069]	[0.018]	[0.058]	[0.038]	[0.016]	[0.032]	[0.016]	[0.020]
Graduate Studies	0.195	0.202	0.192	0.214	0.089	0.190	0.172	0.175	0.170
	[0.045]	[0.105]	[0.021]	[0.098]	[0.083]	[0.016]	[0.035]	[0.023]	[0.027]
College Reputation	0.037	0.040	0.032	0.066	0.048	0.037	0.022	0.040	0.034
	[0.018]	[0.016]	[0.004]	[0.013]	[0.017]	[0.005]	[0.004]	[0.005]	[0.007]
Literacy	0.031	0.011	0.017	0.017	0.039	0.020	0.020	0.018	0.019
	[0.008]	[0.016]	[0.002]	[0.012]	[0.012]	[0.003]	[0.002]	[0.002]	[0.003]
Numeracy	0.051	0.021	0.018	0.003	0.041	0.030	0.023	0.020	0.028
	[0.007]	[0.018]	[0.003]	[0.019]	[0.015]	[0.003]	[0.003]	[0.005]	[0.003]
Foreign language	0.027	0.069	0.010	0.075	0.048	0.020	0.011	0.030	0.011
	[0.010]	[0.016]	[0.005]	[0.019]	[0.021]	[0.006]	[0.006]	[0.007]	[0.008]
Field Specific	0.034	0.027	0.021	0.019	-0.011	0.019	0.027	0.025	0.021
	[0.007]	[0.014]	[0.005]	[0.022]	[0.010]	[0.004]	[0.004]	[0.004]	[0.003]
Sample:	Specific	Specific	Specific	Specific	Specific	Specific	Specific	Specific	Specific
Observations	9,610	4,907	86,051	1,410	5,300	98,182	57,757	39,626	82,124
R-squared	0.528	0.642	0.274	0.175	0.633	0.246	0.247	0.387	0.297
Controls:									
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Types of Degrees	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The dependent variable and measures of skills are as described in the notes of Table 1. Estimations using the full sample follow the same specification of column (5) of Table 1, but within the set of individuals belonging to the population of interest –defined by the categories on top of the table. Economic activity categories were defined by grouping four-digit industry codes. Column (6) and (7) use the subsample of women and men, respectively. columns (8) and (9) include individuals working for small and large firms, respectively. Individual, college and occupation controls as described in the notes of Table 1. Standard errors clustered at the municipality level and in brackets.

Returns by Gender and Firm Size. We also explore heterogeneity by gender and firm size. Columns (6) and (7) of Table 5 present the returns by gender, and suggest that female workers have, for numeracy and foreign languages measures, slightly higher returns than male workers. However, males have higher returns to specific skills than females (2.7 percent vs. 1.9 percent). We do not observe heterogeneity by gender in terms of returns to degrees.

Columns (8) and (9) of the same table explore heterogeneity among firm size. [31] We find that numeracy skills and returns to postsecondary degrees are higher among people working

[31] The firm size is computed using the number of workers observed by establishment in the earnings records. These records only include formal workers who have been enrolled in a postsecondary education program, they could be active students, dropouts or have a college degree. Although this is not a perfect measure of firm size, it stills proxies fairly well for firm size.

in large firms compared to those in smaller firms.

6 Conclusion

Understanding the reasons that make some workers earn higher wages than others is a key question in labor economics. In this paper, we investigate how skills and postsecondary degrees relate to labor market outcomes later in life. Our paper differs from the previous literature in that we are able to jointly measure the importance of a broad set of skills (literacy, numeracy, foreign language, field-specific and non-cognitive) and different types of postsecondary degrees (by length, quality and field of study).

Exams required for graduation from high school and college in Colombia evaluate students in multiple areas including mathematics, language (Spanish) and foreign language (English). A follow-up survey then assesses the non-cognitive abilities of former graduates. We combine a vast vector of test scores from these exams with college enrollment and graduation data and with Social Security records to create a data set that follows individuals from high school to college, and into the first years in the labor market. This uniquely rich dataset allows us to unbundle the returns to postsecondary education: we jointly estimate returns to different skills, to different types of degrees; accounting for the length, quality and field in which those degrees were granted.

We view the results presented in this paper as a building block toward better understanding which skills and degrees are valued in today's labor market. Our evidence confirms that the return to basic skills (like mathematics, literacy, foreign language, and non-cognitive) are sizable even for people that graduated in, essentially, the same program. On average a one standard deviation increase in numeracy skills have a return between 1.7 percent and 14 percent, in literacy skills between 1.4 percent and 11.2 percent, and in foreign language between 1 percent and 12.3 percent. Specific skills are also important ranging between 2.1 percent and 12.3 percent. Returns to non-cognitive skills ranges from 1.2 to 2.6 percent. This is equivalent to half to one year extra of education in a 4-year private university. We also find that the type of postsecondary degrees individuals obtain are associated to very different wage returns. First, the returns to college reputation are as high as the returns to skills. Second, the annual returns to a four-year degree public program is 3.9 percent higher than returns to two-year public (and 5.9 higher in a private school). Third, what field people study can be as important as to where they earn those degrees: students of engineering and medicine expect 33 to 70 percent higher wages than students of arts or education.

The results in this paper can be useful for a number of related literatures. There is a large literature in economics evaluating the effect of education interventions aimed at improving learning (most frequently of numeracy and literacy skills) including: (i) school choice (e.g., competition, vouchers); (ii) human resources policies (e.g., teacher pay, incentives, and training); (iii) school and classroom management policies (e.g., class sizes and student tracking) and; (iv) school resources (e.g., spending, computers, remedial teaching, student incentives). Many of these studies lack cost-benefit analysis in part because it is difficult to monetize the benefits. Our paper provide wage returns that could be used for that purpose.

Finally, our results also provide some insights into a number of policy debates. First, wage gains associated with admission to some schools and fields can be sizable and, in fact, explain a substantial part of the variation of the wage variance. This suggests that interventions aimed at helping low-income and qualified students gaining admission to certain fields and to four-year colleges can improve welfare (Hoxby and Turner, 2013). Second, returns to skills are positive and large both on average and for most wage quantiles. This suggests that, in the presence of resources constraints, policies that aim at improving the quality of education of low-income individuals can be expected to reduce wage inequality.

References

Acosta, P., Muller, N., and Sarzosa, M. (2015). Beyond Qualifications: Returns to Cognitive and Socio-Emotional Skills in Colombia. Working Paper 7430, The World Bank.

Azam, M., Chin, A., and Prakash, N. (2013). The Returns to English-Language Skills in India. *Economic Development and Cultural Change*, 61(2):335–367.

Bassi, M., Busso, M., Urzúa, S., and Vargas, J. (2012). Disconnected. Skills, Education and Employment in Latin America. IDB Publications.

Behrman, J. R., Birdsall, N., and Székely, M. (2007). Economic policy changes and wage differentials in latin america. *Economic Development and Cultural Change*, 56(1):57–97.

Bleakley, H. and Chin, A. (2004). Language Skills and Earnings: Evidence from Childhood Immigrants. *The Review of Economics and Statistics*, 86(2):481–496.

Bowles, S., Gintis, H., and Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature*, 39(4):1137–1176.

Budría, S. and Swedberg, P. (2015). The Impact of Language Proficiency on Inmigrants. *Revista de Economía Aplicada*, XXIII(67):62–91.

Busso, M., Cristia, J., Hincapie, D., Julian, M., and Ripani, L. (2017a). *Learning Better: Public Policy for Skills Development*, chapter 3, pages 45–68. Development in the Americas. IDB Publications.

Busso, M., Dinkelman, T., Martínez, A. C., and Romero, D. (2017b). The effects of financial aid and returns information in selective and less selective schools: Experimental evidence from chile. *Labour Economics*, 45:79 – 91. Field experiments in labor economics and social policies.

Carneiro, P., Hansen, K., and Heckman, J. (2003). 2001 Lawrence R. Klein Lecture. Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economics Review*, 44(2):361–422.

Christofides, L. N. and Swidinsky, R. (2010). The Economic Returns to the Knowledge and Use of a Second Official Language: English in Quebec and French in the Rest-of-Canada. *Canadian Public Policy / Analyse de Politiques*, 36(2):137–158.

Cunha, F. and Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2):31–47.

Cunha, F., Heckman, J. J., and Lochner, L. (2006). *Interpreting the Evidence on Life Cycle Skill Formation*, volume 1 of *Handbook of the Economics of Education*, chapter 12, pages 697–812. Elsevier.

Dasgupta, U., Mani, S., Sharma, S., and Singhal, S. (2017). Cognitive, socioemotional, and behavioural returns to college quality. Discussion Paper 10701, IZA Institute of Labor Economics.

De Coulon, A., Marcenaro-Gutierrez, O., and Vignoles, A. (2008). The Value of Basic Skills in the British Labour Market. Discussion Paper 77, Centre for the Economics of Education, LSE.

Di Paolo, A. and Tansel, A. (2015). Returns to Foreign Language Skills in a Developing Country: The Case of Turkey. *The Journal of Development Studies*, 51(4):407–421.

Farber, H. S. and Gibbons, R. (1996). Learning and wage dynamics*. *The Quarterly Journal of Economics*, 111(4):1007–1047.

Fasih, T., Patrinos, H. A., and Sakellariou, C. (2013). Functional Literacy, Heterogeneity and the Returns to Schooling: Multi-Country Evidence. Working Paper 6677, The World Bank.

Firpo, S., Fortin, N., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.

Guo, Q. and Sun, W. (2014). Economic Returns to English Proficiency for College Graduates in Mainland China. *China Economic Review*, 30:290 – 300.

Hanushek, E. A. and Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100(2):267–71.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., and Woessmann, L. (2015). Returns to Skills Around the World: Evidence from {PIAAC}. *European Economic Review*, 73:103 – 130.

Hastings, J. S., Neilson, C. A., and Zimmerman, S. D. (2013). Are some degrees worth more than others? evidence from college admission cutoffs in chile. Working Paper 19241, National Bureau of Economic Research.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Hoxby, C. and Turner, S. (2013). Expanding college opportunities for high-achieving, low income students. *Discussion Papers, Stanford Institute for Economic Policy Research*, 12-014.

Icfes (2017). Informe Nacional de Resultados: Examen Saber Pro 2016. Informe Nacional.

Ishikawa, M. and Ryan, D. (2002). Schooling, basic skills and economic outcomes. *Economics of Education Review*, 21.

James, J. (2013). The Surprising Impact of High School Math on Job Market Outcomes. *Economic Commentary.*

Joensen, J. S. and Nielsen, H. S. (2009). Is there a causal effect of high school math on labor market outcomes? *The Journal of Human Resources*, 44:171–198.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of study, earnings, and self-selection*. *The Quarterly Journal of Economics*, 131(3):1057–1111.

Koenker, R. and Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1):33–50.

Levine, P. B. and Zimmerman, D. J. (1995). The benefit of additional high-school math and science classes for young men and women. *Journal of Business and Economic Statistics*, 13(2).

Lindqvist, E. and Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics*, 3(1):101–28.

Lustig, N., Lopez-Calva, L. F., and Ortiz-Juarez, E. (2012). Declining inequality in latin america in the 2000s. the cases of argentina, brazil, and mexico. Working Paper 6248, The World Bank.

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

Murnane, R., Willett, J., and Levy, F. (1995). The growing importance of cognitive skills in wage determination. *The Review of Economics and Statistics*, 77(2).

Psacharopoulos, G. and Ng, Y. C. (1994). Earnings and education in latin america. *Education Economics*, 2(2):187–207.

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in pisa, timss, pirls and iq-tests across nations. *European Journal of Personality*, 21(5):667–706.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2):417–458.

Rodríguez, J., Urzúa, S., and Reyes, L. (2015). Heterogeneous economic returns to postsecondary degrees: Evidence from chile. *Journal of Human Resources*.

Saiz, A. and Zoido, E. (2005). Listening to What the World Says: Bilingualism and Earnings in the United States. *The Review of Economics and Statistics*, 87(3):523–538.

Sanders, C. (2016). Reading Skills and Earnings: Why Does Doing Words Good Hurt Your Wages? *mimeo*.

Sarzosa, M. and Urzúa, S. (2016). Implementing Factor Models for Unobserved Heterogeneity in Stata. *The Stata Journal*, 16(1):197–228.

Song, M., Orazem, P., and Wohlgemuth, D. (2008). The role of mathematical and verbal skills on the returns to graduate and professional education. *Economics of Education Review*, 27(6):664–675.

Stöhr, T. (2015). The Returns to Occupational Foreign Language Use: Evidence from Germany. *Labour Economics*, 32:86 – 98.

Tyler, J. (2004). Basic skills and the earnings of dropouts. *Economics of Education Review*, 23:221–235.

A Appendix: Additional Tables

Appendix Table 1: Estimates of the Effect of Cognitive Skills from Previous Literature

Skills	Reference	Country/ City	Population in Estimating Sample	Identification Strategy	Estimation	Dependent Variable (log)	Range[a] Lowest Estimate	Range[a] Highest Estimate
Panel A: Numeracy								
	Levine and Zimmerman (1995)	USA	1980 HS graduates	Selection on Observables	OLS	1990 & 1986 weekly wage	0.028	0.030[b]
	Murnane et al. (1995)	USA	1972 and 1980 HS graduates	Selection on Observables	OLS	1978 & 1986 hourly wage	0.026	0.069[c]
	Tyler (2004)	Florida (USA)	HS dropouts	Selection on Observables	OLS	1995-1999 quarterly earnings	0.063	0.074
	Song et al. (2008)	USA	College graduates	Selection on Observables	IV	1993 earnings	0.181	0.210[d]
	Joensen and Nielsen (2009)	Denmark	1986 and 1987 HS graduates	Quasi-experiment	IV	1999-2002 annual earnings	0.12	0.32[e]
	Hanushek et al. (2015)	Several	Adults aged 20 to 50 from 23 OECD countries	Selection on Observables	OLS	2011-2012 hourly earnings	0.079	0.178[f]
Panel B: Literacy								
	Ishikawa and Ryan (2002)	USA	Adults above 16	Selection on Observables	OLS	weekly wages	0.001	0.008[g]
	Fasih et al. (2013)	Several	Males aged 22 to 65 from 20 countries, mostly OECD	Selection on Observables	OLS	hourly wage	0.021	0.210[h]
	Hanushek et al. (2015)	Several	Adults aged 20 to 50 from 23 OECD countries	Selection on Observables	OLS	1993 earnings	0.068	0.171
	Sanders (2016)	USA	Populations represented in 5 longitudinal surveys	Selection on Observables	OLS	real wages	-0.056	-0.024

[a] Estimates points correspond to standardized test scores, unless another interpretation is suggested. [b] Estimations correspond to the number of mathematics classes taken during high school. [c] Point estimates are given originally for levels of a mathematics score. Since a one standard deviation is 6.25, then coefficients are translated into this scale. Lower and upper bounds correspond, respectively, to 1972 and 1980 high school graduates. [d] The mathematics score is estimated in levels. [e] Numeracy is a dummy valued 1 if individual's took a high level mathematics course during high school. Reported bounds correspond to the pilot school sample. [f] Numeracy is also estimated using literacy as insturment and the coefficient found is 0.201. [g] Lower and Upper bounds correponds to the point estimates for black men and hispanic men samples, which respectively are the lowest and highest point estimes. Literacy was estimated in levels. [h] Lower and Upper bounds correspond to the point estimates for Denmark and Bermuda, which are respectively the lowest and largest estimates found. Check the paper for more details.

Skills	Reference	Country/ City	Population in Estimating Sample	Identification Strategy	Estimation	Dependent Variable (log)	Range	
							Lowest Estimate	Highest Estimate
Panel C: Foreign Language								
	Bleakley and Chin (2004)	USA	1960-1974 Young immigrants	Quasi-experiment	IV	1990 annual wage	0.222	0.334[i]
	Saiz and Zoido (2005)	USA	College graduates	Selection on Observables	OLS	1997 hourly wage	0.025	0.028[j]
	Christofides and Swidinsky (2010)	Quebec (CA)	Fulltime native workers aged 15 to 64	Selection on Observables	OLS	2000 earnings	0.109	0.139[k]
	Azam et al. (2013)	India	Male workers aged 18 to 65	Selection on Observables	OLS	2005 earnings	0.345	0.603[l]
	Guo and Sun (2014)	China	College graduates	Selection on Observables	OLS	2010 monthly wage	0.033	0.131[m]
	Budría and Swedberg (2015)	Spain	Male immigrants aged 18 to 65	Quasi-experiment	IV	2006-2007 hourly wages	0.049	0.204[n]
	Di Paolo and Tansel (2015)	Turkey	Male workers	Selection on Observables	OLS	2007 wage	0.107	0.072[o]
	Stöhr (2015)	Germany	Fulltime workers	Selection on Observables	OLS	2005-2006 gross monthly wage	0.033	0.093[p]

[i] The independent variable takes 1 as value if individual speak english very well. [j] IV, Panel and PSM estimations are also considered. For instance, PSM point estimates ranged from 0.020 to 0.021. [k] Point estimates correspond to a subsample of only men. The independent variable takes 1 as value if individual uses english in his/her workplace. Estimations are also carried for women and ranged from 0.068 to 0.076. [l] Point estimates for a dummy varibles that takes 1 as value if the individual is fluent in english. Estimations for knowing little english can be seen in the paper. [m] The english score of CET-4 test is used to measure foreign language proficiency, check the paper for more details. [n] Lower bound corresponds to OLS estimation and Upper bound corresponds to IV estimation using simultaneously the following instruments: 1(arrived before 12), 1(has a child proficient in spanish) and 1(willingness to stay in Spain). [o] The independent variable takes 1 as value if the individual knows english. Other languages are estimated (french, german, arabic and russian), but those who know english account for 76%. [p]The independent variable takes 1 as value if the individual's occupation requires expertise in a foreign language.

Appendix Table 2: Non-Cognitive Skills in the Survey of Graduates

Questionnaire item	**Factor Loadings**
Accept the difference and work under multicultural contexts	0.565
Learn and keep updated	0.593
Work under pressure	0.504
Adopt a coexistence culture	0.589
Work independently without permanent supervision	0.557
Convince and persuade others	0.532
Identify and use communication symbols (i.e. non-verbal or iconic language)	0.344
Abstraction capacity, analysis and synthesis	0.625
Identify, plan and solve problems	0.680

Notes. The answers for each item ranges from 1 to 4: (1) Very unsatisfied, (2) Unsatisfied, (3) Satisfied and (4) Very satisfied. The largest eigenvalue from a factor analysis model is 2.836, and the second largest is 0.109. For each item, the factor loadings in this table correspond to the factor with the largest eigenvalue. The scale reliability coefficient, α, is 0.80.

Appendix Table 3: Non-Cognitive Skills used as Instrumental Variable

Questionnaire item	**Factor Loadings**
To be creative and innovative	0.5836
Look for, analyze, manage and share information	0.6014
Understand the surrounding reality	0.6350
Assume responsibilities and make decisions	0.6959
Plan and use time effectively to achieve the goals	0.6294
Formulate and execute projects	0.5479
Work in teams to achieve common goals	0.6575
Consider values and professional ethics to perform tasks	0.6479
Adapt to changes	0.6486
Take risks	0.6712
Identify opportunities and resources	0.6660

Notes. The answers for each item ranges from 1 to 4: (1) Very unsatisfied, (2) Unsatisfied, (3) Satisfied and (4) Very satisfied. The largest eigenvalue from a factor analysis model is 4.453, and the second largest is 0.109. For each item, the factor loadings in this table correspond to the factor with the largest eigenvalue. The scale reliability coefficient, α, is 0.88.

Appendix Table 4: Descriptive Statistics

	Estimation Sample						Mean Difference	
	Full (N = 363,330)		Specific (N = 155,939)		Survey (N = 2,401)			
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	*p*-value	
	(1)		(2)		(3)		(1) - (2)	(1) - (3)
Panel A: Socioeconomic Statistics								
Age	26.93	3.44	26.86	3.5	28.74	4.57	0.00	0.00
Share female	0.61	0.49	0.63	0.48	0.59	0.49	0.00	0.08
Share living in big urban areas	0.73	0.45	0.71	0.45	0.80	0.40	0.00	0.00
Share living in low income households	0.45	0.50	0.43	0.50	0.46	0.50	0.00	0.40
Share graduate students	0.02	0.14	0.03	0.17	0.04	0.19	0.00	0.00
Panel B: Education Statistics								
Share graduated from:								
STEM	0.26	0.44	0.26	0.44	0.25	0.44	0.29	0.36
Business and Economics	0.31	0.46	0.27	0.45	0.27	0.44	0.00	0.00
Social Sc. and Humanities	0.16	0.36	0.15	0.36	0.12	0.33	0.00	0.00
Health and Education	0.22	0.41	0.28	0.45	0.22	0.41	0.00	0.66
Share postsecondary degrees:								
Two-Year Private	0.09	0.29	0.04	0.20	0.40	0.49	0.00	0.00
Two-Year Public	0.10	0.30	0.05	0.21	0.10	0.3	0.00	0.64
Four-Year Private	0.50	0.50	0.56	0.50	0.32	0.47	0.00	0.00
Four-Year Public	0.32	0.47	0.36	0.48	0.18	0.38	0.00	0.00
Panel C: Occupation Statistics								
Share working in:								
Manufacture	0.07	0.26	0.07	0.26	0.08	0.27	0.00	0.60
Commerce	0.03	0.18	0.03	0.17	0.04	0.19	0.00	0.27
Services	0.54	0.50	0.55	0.50	0.60	0.49	0.00	0.00
Turism	0.01	0.10	0.01	0.09	0.01	0.11	0.00	0.83
Retail	0.04	0.20	0.04	0.19	0.04	0.20	0.00	0.95
Share working as:								
Public Employee	0.03	0.18	0.04	0.21	0.04	0.19	0.00	0.12
Independent Workers	0.07	0.25	0.07	0.26	0.08	0.27	0.09	0.03
Share working in firms that are:								
$[1-10]$	0.15	0.36	0.14	0.35	0.12	0.32	0.00	0.00
$[11-50]$	0.18	0.39	0.18	0.39	0.16	0.37	0.97	0.02
$[51-100]$	0.08	0.28	0.08	0.28	0.09	0.29	0.32	0.13
$[101-500]$	0.21	0.40	0.21	0.41	0.21	0.41	0.15	0.50
$[501-1000]$	0.10	0.30	0.10	0.30	0.09	0.29	0.02	0.26
$[1001+]$	0.28	0.45	0.28	0.45	0.33	0.47	0.01	0.00
Panel D: Labor Statistics								
Current Wage	16.86	11.24	17.48	11.59	19.75	10.79	0.00	0.00
First Wage	13.42	8.48	14.07	8.88	12.85	7.72	0.00	0.00
Average Wage	15.17	8.86	15.8	9.21	16.19	7.92	0.00	0.00
Current Tenure	1.77	1.02	1.78	1.11	2.28	1.3	0.00	0.00

Notes. Descriptive statistics of students who took the college exit exam from 2011 to 2015 for whom data were matched to earnings and college records. Big urban area refer to the largest 13 cities in Colombia. Low-income households refer to individuals in the first two incom strata designated by the place of living. Wages are presented in nominal 2016 USD currency (1 USD = 3050.98 COP).

Appendix Table 5: Correlation Matrix Across Test Scores

	High School Exit Exams					College Exit Exams		
	Average	Subject	Literacy	Numeracy	Foreign	Literacy	Numeracy	Foreign
Panel A: Full Sample (N = 363,330)								
High School Exit Exams:								
Subject	**0.941***							
Literacy	**0.708***	**0.595***						
Numeracy	**0.671***	**0.546***	**0.394***					
Foreign Language	**0.726***	**0.571***	**0.460***	**0.428***				
College Exit Exams:								
Literacy	0.537*	0.497*	0.438*	0.314*	0.390*			
Numeracy	0.584*	0.553*	0.394*	0.487*	0.401*	**0.449***		
Foreign Language	0.611*	0.515*	0.414*	0.393*	0.678*	**0.458***	**0.464***	
Panel B: Specific Sample (N = 155,939)								
High School Exit Exams:								
Subject	**0.941***							
Literacy	**0.705***	**0.593***						
Numeracy	**0.676***	**0.553***	**0.396***					
Foreign Language	**0.727***	**0.573***	**0.459***	**0.431***				
College Exit Exams:								
Literacy	0.553*	0.513*	0.446*	0.331*	0.401*			
Numeracy	0.611*	0.579*	0.409*	0.509*	0.424*	**0.470***		
Foreign Language	0.623*	0.526*	0.419*	0.405*	0.687*	**0.472***	**0.490***	
Major-Specific	0.520*	0.496*	0.394*	0.344*	0.345*	**0.519***	**0.511***	**0.413***
Panel C: Survey Sample (N = 2,401)								
High School Exit Exams:								
Subject	**0.933***							
Literacy	**0.708***	**0.586***						
Numeracy	**0.620***	**0.483***	**0.369***					
Foreign Language	**0.668***	**0.508***	**0.420***	**0.347***				
College Exit Exams:								
Literacy	0.505*	0.461*	0.430*	0.278*	0.335*			
Numeracy	0.513*	0.483*	0.373*	0.417*	0.321*	**0.395***		
Foreign Language	0.597*	0.500*	0.434*	0.352*	0.636*	**0.463***	**0.400***	
Socioemotional:								
Factor Score	0.054*	0.053*	0.038†	0.047*	0.039†	0.026	0.025	0.058*

Notes. Pairwise correlations are estimated using the Pearson's formula. For both, the college exit exam (*Saber Pro*) and the high school exit exam (*Saber 11*), individual's scores are standardized with respect to the corresponding average in each test edition. The specific scores from the college exit exam is standardized with respect to average of the test edition and the corresponding group of related programs. The subject score from the high school exit exam is computed as the standardized average of biology, philosophy, physics, chemistry, and social science tests. The non-cognitive scores were computed as the predictions from a factor model considering categorical answers to nine questions (see Appendix Table 2). † $p<0.1$, * $p<0.05$.

Appendix Table 6: Other Outcomes

	Dependent Variable: log(First Wage After Graduation)			*Dependent Variable*: log(Avg. Wage Since Graduation)			*Dependent Variable*: log(Current Earnings)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Two-Year Private	0.068	0.061	0.144	0.078	0.077	0.137	0.080	0.084	0.130
	[0.009]	[0.010]	[0.026]	[0.011]	[0.009]	[0.036]	[0.014]	[0.013]	[0.050]
Four-Year Public	0.149	0.156	0.287	0.174	0.171	0.263	0.177	0.159	0.199
	[0.010]	[0.013]	[0.029]	[0.010]	[0.012]	[0.029]	[0.011]	[0.014]	[0.041]
Four-Year Private	0.211	0.221	0.216	0.244	0.244	0.225	0.251	0.239	0.218
	[0.017]	[0.019]	[0.027]	[0.018]	[0.018]	[0.025]	[0.020]	[0.022]	[0.035]
Literacy	0.016	0.010	0.008	0.023	0.016	0.017	0.027	0.020	0.015
	[0.001]	[0.002]	[0.009]	[0.001]	[0.002]	[0.006]	[0.001]	[0.002]	[0.008]
Numeracy	0.025	0.018	0.032	0.031	0.023	0.020	0.035	0.026	0.022
	[0.001]	[0.002]	[0.011]	[0.001]	[0.002]	[0.010]	[0.002]	[0.003]	[0.010]
Foreign language	0.011	0.013	0.047	0.015	0.016	0.032	0.015	0.015	0.025
	[0.003]	[0.003]	[0.008]	[0.004]	[0.004]	[0.006]	[0.005]	[0.005]	[0.007]
Field Specific		0.016			0.020			0.022	
		[0.003]			[0.004]			[0.004]	
Socioemotional			0.010			0.021			0.021
			[0.011]			[0.008]			[0.006]
College Reputation	0.026	0.029	0.030	0.031	0.032	0.050	0.032	0.031	0.059
	[0.005]	[0.005]	[0.017]	[0.004]	[0.005]	[0.020]	[0.004]	[0.005]	[0.023]
Sample:	Full	Specific	Survey	Full	Specific	Survey	Full	Specific	Survey
Observations	363,330	155,939	2,401	363,330	155,939	2,401	363,330	155,939	2,401
R-squared	0.196	0.237	0.244	0.233	0.264	0.324	0.195	0.220	0.244
Controls:									
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field of Study	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Estimations in columns (1), (4) and (7) follow the specification of column (4) in Table 1. Estimations in columns (2), (5) and (8) follow the specification of column (5) in Table 1. Estimations in columns (3), (6) and (9) follow the specification of column (6) in Table 1. The dependent variable for columns (1) to (3) is the log of the first wage observed after the individual's graduation year. The dependent variable for columns (4) to (6) is the log of the average wage from 2011 to 2016 for each individual, taking into account only wages after the year of graduation. The dependent variable for columns (7) to (9) is the log of the last observed annual earnings for each individual. Individual and field of study controls as described in the notes of Table 1. Standard errors are clustered at the municipality level and in brackets.

Appendix Table 7: Extensions: Estimates' Ranges and Measurement Error

	Dependent Variable: log(Current Wage)										
	OLS						IV: Measurement Error Correction				
	Simultaneous			Stacked			Simultaneous		Stacked		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Literacy	0.026	0.020	0.015	0.036	0.034	0.022	0.075	0.010	0.099	0.100	0.110
	[0.001]	[0.002]	[0.008]	[0.001]	[0.002]	[0.009]	[0.010]	[0.059]	[0.012]	[0.014]	[0.046]
Numeracy	0.034	0.025	0.017	0.043	0.040	0.024	0.048	0.114	0.085	0.081	0.127
	[0.003]	[0.003]	[0.009]	[0.003]	[0.003]	[0.009]	[0.007]	[0.044]	[0.010]	[0.013]	[0.029]
Foreign Language	0.010	0.011	0.024	0.024	0.026	0.031	-0.009	0.021	0.034	0.036	0.060
	[0.005]	[0.005]	[0.013]	[0.006]	[0.006]	[0.014]	[0.004]	[0.028]	[0.006]	[0.006]	[0.022]
Field Specific		0.021			0.037					0.119	
		[0.004]			[0.002]					[0.019]	
Socioemotional			0.012			0.012		0.012			0.007
			[0.006]			[0.006]		[0.007]			[0.007]
Sample:	Full	Specific	Survey	Full	Specific	Survey	Full	Survey	Full	Specific	Survey
Observations	363,330	155,939	2,401	363,330	155,939	2,401	363,330	2,401	363,330	155,939	2,401
Controls:											
Individual & Field	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
College	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The dependent variable and measures of skills are as described in the notes of Table 1. Estimations within the full sample use the same specification of column (6) of Table **??**. Likewise, estimations within the specific and survey samples use respectively the specifications of columns (8) and (10) of Table **??**. Stacked results refer to separe regressions for each ability measure. In columns (7) to (9), cognitive skills were instrumented using each time numeracy, literacy and foreign language standardized scores computed from the high school exit exam. Individual and types of degrees area as described in the notes of Table 1. Standard errors clustered by municipality and in brackets.

Appendix Table 8: Extensions: Estimates' Ranges and Measurement Error

	Dependent Variable: log(Current Wage)													
Panel A:	Ordinary Least Squares Estimations													
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Two-Year Private	0.080	0.084	0.128	0.081	0.078	0.078	0.086	0.083	0.083	0.084	0.129	0.129	0.126	0.132
	[0.014]	[0.012]	[0.049]	[0.014]	[0.014]	[0.014]	[0.012]	[0.012]	[0.012]	[0.012]	[0.053]	[0.052]	[0.051]	[0.052]
Four-Year Public	0.174	0.156	0.197	0.176	0.175	0.173	0.159	0.156	0.161	0.164	0.197	0.199	0.206	0.208
	[0.011]	[0.013]	[0.042]	[0.011]	[0.011]	[0.011]	[0.013]	[0.013]	[0.013]	[0.013]	[0.043]	[0.043]	[0.042]	[0.044]
Four-Year Private	0.249	0.236	0.212	0.247	0.249	0.240	0.235	0.234	0.232	0.240	0.215	0.221	0.211	0.222
	[0.020]	[0.021]	[0.036]	[0.020]	[0.020]	[0.019]	[0.021]	[0.021]	[0.021]	[0.021]	[0.039]	[0.038]	[0.037]	[0.038]
Literacy	0.027	0.020	0.017	0.038			0.036				0.026			
	[0.001]	[0.002]	[0.007]	[0.002]			[0.002]				[0.007]			
Numeracy	0.035	0.025	0.019		0.044			0.041				0.027		
	[0.002]	[0.003]	[0.010]		[0.003]			[0.003]				[0.010]		
Foreign language	0.016	0.016	0.026			0.030			0.031				0.035	
	[0.006]	[0.006]	[0.006]			[0.006]			[0.006]				[0.007]	
Field Specific		0.022								0.038				
		[0.004]								[0.003]				
Socioemotional			0.023											0.023
			[0.006]											[0.006]
Sample:	Full	Specific	Survey	Full	Full	Full	Specific	Specific	Specific	Specific	Survey	Survey	Survey	Survey
Observations	363,330	155,939	2,401	363,330	363,330	363,330	155,939	155,939	155,939	155,939	2,401	2,401	2,401	2,401
R-squared	0.194	0.219	0.247	0.191	0.192	0.189	0.216	0.217	0.215	0.217	0.243	0.243	0.243	0.243
Panel B:	IV: Measurement Error Correction													
	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	
Two-Year Private	0.083	0.128	0.085	0.079	0.078	0.090	0.083	0.083	0.083	0.126	0.126	0.124	0.132	
	[0.015]	[0.045]	[0.016]	[0.014]	[0.014]	[0.013]	[0.011]	[0.013]	[0.012]	[0.054]	[0.047]	[0.050]	[0.052]	
Four-Year Public	0.176	0.169	0.177	0.174	0.173	0.153	0.148	0.160	0.168	0.165	0.170	0.206	0.208	
	[0.012]	[0.045]	[0.012]	[0.011]	[0.011]	[0.014]	[0.013]	[0.013]	[0.015]	[0.040]	[0.041]	[0.042]	[0.044]	
Four-Year Private	0.257	0.225	0.255	0.254	0.238	0.235	0.232	0.231	0.252	0.197	0.222	0.206	0.222	
	[0.022]	[0.034]	[0.022]	[0.021]	[0.020]	[0.022]	[0.021]	[0.021]	[0.023]	[0.042]	[0.035]	[0.036]	[0.038]	
Literacy	0.075	-0.004	0.105			0.106				0.112				
	[0.011]	[0.057]	[0.012]			[0.013]				[0.046]				
Numeracy	0.045	0.154		0.087			0.083				0.140			
	[0.006]	[0.046]		[0.008]			[0.010]				[0.028]			
Foreign language	-0.002	0.002			0.040			0.041				0.053		
	[0.004]	[0.027]			[0.006]			[0.005]				[0.022]		
Field Specific									0.123					
									[0.017]					
Socioemotional		0.026											0.022	
		[0.007]											[0.006]	
Sample:	Full	Survey	Full	Full	Full	Specific	Specific	Specific	Specific	Survey	Survey	Survey	Survey	
Observations	363,330	2,401	363,330	363,330	363,330	155,939	155,939	155,939	155,939	2,401	2,401	2,401	2,401	
R-squared	0.042	0.046	0.034	0.042	0.043	0.028	0.036	0.037	0.023	0.071	0.054	0.097	0.097	
Controls:														
Individual & Field	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Types of Degrees	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes.*The dependent variable and measures of skills are as described in the notes of Table 1. Estimations within the full sample use the same specification of column (1) of Table **??**. Likewise, estimations within the specific and survey samples use, respectively, the specifications of columns (3) and (5) of Table **??**. Measures of numeracy, literacy and foreign language computed from the high school exit exam are used as instruments within the full and specific samples. We additional include an alternative measure of non-cognitive skill as instrument within the survey sample. Individual and types of degrees area as described in the notes of Table 1. Standard errors clustered by municipality and in brackets.

B Appendix: Skills Measures

In this appendix, we describe the cognitive abilities evaluated in the college exit exam and the high school exit exam. We also describe the non-cognitive abilities collected in the follow-up survey of graduates, and how we compute the different measures of skills used in the paper.

The abilities tested in the college exit exam (Saber Pro) are divided into two sections. The first section has five mandatory general tests and consists of 160 multiple choice questions and 1 open question, lasting a maximum of 4 hours and 40 minutes. The main objective of this section is to evaluate common abilities that students from the wide range of fields should be able to apply in non-specialized tasks. On the other hand, some students also take a second section with specific tests. This section is only available if the student's college previously decided which specific tests will be applied to their undergraduate programs. There are 40 specific tests and combinations of theses suggested for each field of study or group of related programs.[32] Following those combinations, colleges can decide up to three specific tests. The maximum time allowed for students taking one specific test is 1 hour and 30 minutes, while students taking 2 or 3 specific tests have a maximum of 4 hours and 30 minutes. Appendix Table 9 presents more details on the abilitites evaluated and a sample question for each general test. It also shows one question for the Economic Analysis test, to present an example for one of the specific tests evaluated in the second part of the college exit exam.

We now describe how the different measures of skills were computed. Let η_{t_i} be student i's scores for η_t, which is test η applied in time t. Let μ_{η_t} be the mean of η_t, this is:

$$\mu_{\eta_t} = \frac{1}{|I_t|} \sum_{i \in I_t} \eta_{t_i}$$

where I_t is the set of students who took η_t. Thus, the standardized score of η_{t_i} with respect to μ_{η_t} is $\tau_{t_i} = \frac{(\eta_{t_i} - \mu_{\eta_t})}{\sigma_{\eta_t}}$, where the standard deviation of η_t is defined as:

$$\sigma_{\eta_t} = \sqrt{\frac{1}{|I_t| - 1} \sum_{i \in I_t} (\eta_{t_i} - \mu_{\eta_t})^2}.$$

Let now $\mu^s_{\eta_t}$ be the mean of η_t, but computed within a sample $s \subseteq |I_t|$. If $i \in s$, then $\tau^s_{t\,i}$ is the standardized score of η_{t_i} within sample s.

Taking the previous framework into account, we use the test scores from Saber Pro exam to compute literacy, numeracy and foreign language measures as the standardized scores within the four samples mentioned in Appendix C. Hence, for each individual in a sample, the numeracy ability level is the standardized score of the quantitative reasoning test with

[32]Throughout the years there have been changes in some of these specific evaluations and, thus, our data contains a total of 87 specific tests. For instance, some specific tests were divided into others and some have disappeared.

respect to the mean in the time period in which the test was taken, and considering only individuals within the sample. Likewise, the foreign language ability level is computed as the standardized score of the english proficiency test within each sample. On the other hand, to define the literacy measure we first compute the average score of written communication and critical reading tests, and then standardized the resulting vector in the same way as we did for numeracy and foreign language. To compute the specific skills measure, we first average all field-specific test scores available for each student and then proceed to standardize the resulting vector, this time considering both the sample and group of related majors.

Appendix Table 9: Description of Tests and Abilities evaluated in Saber Pro exam

Test	Evaluated Abilities	Sample Question
Section 1:		
Critical Reading (35 Questions)	Abilities that allow individuals to understand, interpret and analyze texts found in both, common life and non-specialized academic scenarios.	The text's author states that "seeking justice is the eternal seeking of human happiness". This statement: A) implies that "every human pursues hapiness", B) does not imply that "seeking justice is seeking happiness"
Quantitative Reasoning (35 Questions)	Mathematic abilities that every citizen should have, independently of their profession or occupacion, such as: algebra, calculus, geometry, statistics, interpretation of numeric information, use of mathematics to formulate and execute plans, and use of mathematics to solve problems.	Four utility companies estimated their daily efficiency to solve customer complaints: Electricity: 2 out of 3 solved complaints per day. Aqueduct: 5 out of 6. Telephone: 9 out of 10. The average efficiency for one of these companies is : A) 72%, B) 79%
Citizenship Abilities (35 Questions)	Knowledge and abilities to understand the social environment and its issues, as well as abilities to analyze positions taken by different parties involved in a conflicting situation.	To reduce traffic jams within a city, a major decided to restrict the free circulation of vehicles using the last digit of license plates. The offer of public transporation in the city is limited and has low quality. What undesired effects may cause to the mentioned policy for citizens using the public transportation?
Written Communication (1 Open Question)	Abilities to communicate ideas in writing, regarding a given topic. Students are asked to produce a text in response to a non-specialized problem.	Some consider that national and international sport competitions are used for political and commercial means. Do you agree or disagree with this opinion? Discuss.
English Proficiency (55 Questions)	Communication abilities in English throughout reading, grammar and vocabulary tasks.	The Ozone Layer is a "blanket" (1) _____ earth. It protects (2) _____ from the sun's UV rays. Fill the blanks: (1) A) around, B) through; (2) A) our, B) us.
Section 2:		
Specific Tests (30 - 60 Questions)	Abilities that different postsecondary programs must provide to its students. These abilities have been defined between the Ministry of Education, the academic and professional community, and the industry.	(Economic Analysis Test:) Consider a linear model $y = X\beta + \varepsilon$, where ε is an error term. Assuming that $E(X'\varepsilon) = 0$, then: A) OLS are consistent, and 2SLS are consistent and efficient. B) OLS are inconsistent and 2SLS are consistent and efficient. (OLS: Ordinary Least Squares; 2SLS: Two-Stage Least Squares)

Notes. Information adapted by the authors from Icfes (2017).

To compute the non-cognitive skills measure, we rely on a module of abilities contained in the follow-up survey of graduates. The questionnaire to this module emphasizes, previous to ask any question, that an acquired ability allows individuals to use their knowledge to solve problems in different contexts, and to perform efficiently in their social, academic and working activities. The items we use from the module of abilitites are presented in Appendix Table 2. Surveyed individuals are asked to pinpoint the satisfaction level, regarding the academic and labor impact of each item or ability adquired during their undergradute studies. Additionally, we present the loadings from a factor analysis model performed using the nine items in this table. Using this results we predict a vector of non-cognitive scores, which is then standardized to have mean zero and standard deviation one.

Following MacLeod et al. (2017), we used the administrative records of undergraduate students to build a measure of college reputation or quality. This measure is defined as the mean admission score of graduates, then for college c, in time t, the reputation measure is:

$$R_{ct} = \frac{1}{|G_{ct}|} \sum_{i \in G_{ct}} \tilde{\eta}_i$$

where G_{ct} is the set of students graduating from college c in time t, and $\tilde{\eta}_i$ is the percentile score of individual $i \in G_{ct}$ in the high school exit exam. We then standardized R_{ct} to have mean zero and standard deviation one.

C Appendix: Data Construction

In this appendix, we explain in detail the construction of samples used in the paper. The starting point is to append the individual records of Saber Pro exams applied between the second semester of 2011 and 2015. This is our population.[33] The Colombian Institute for the Evaluation of Education (ICFES, in Spanish) applies Saber Pro twice per year to undergraduate students who have approved at least three-fourths of their coursework. For theses students, the exam constitutes a graduation requirement. Our data contain information for around 1,448,000 students who have been tested in critical reading, written communication, English proficiency, and quantitative reasoning. A test in citizenship abilities has been also evaluated since 2012, and almost 90% of the students have available scores for this test.

For almost 42% of the population, the data has scores from different field-specific tests. The specific exam taken by each student depends exclusively on the college decision, that has to be supported by ICFES. For each program, ICFES suggests a list of potential specific tests per program. Then, each college decides which specific test(s) are their students taking. Students can take up to three specific tests. For more details about measures of skills, and the test scores used to compute them, check Appendix B.

The second step in our data construction is to link the data from Saber Pro with college and labor market records from different sources. We merge individuals using the national identification number. For the subset of individuals that are unable to merge by these means,

[33] It is worth noting that the decision to choose the mentioned period of time was driven by the fact that students attending certain undergraduate programs, were exempted of taking Saber Pro during previous editions of the exam.

we conduct a record linkage procedure using names and dates of birth.[34] Appendix Table 10 shows the merging results by technique, and a measure of reliability for observations merged using record linkage.

We merge the population of students in Saber Pro with three other datasets. First, we merge it with the corresponding results in the high school exit examination, known as Saber 11. We merge 83% observations with a file containing high school test takers from 1996 to 2013. Saber 11 data include socioeconomic variables such as gender, household stratum and the high school where the student graduated from.[35] Most importantly, these data allow us to obtain a proxy of initial abilities, since individuals taking Saber 11 were evaluated in mathematics, physics, chemistry, biology, language, philosophy, geography, history, social sciences and foreign language.[36]

Second, we merge individuals in Saber Pro with a file containing records of individuals enrolled in college between 1998 and 2016. This dataset is known as Spadies (in Spanish "Sistema para la prevención de la deserción de la Educación Superior"). The information in this file is collected each semester by the Ministry of Education directly from all colleges. 82% of the students in Saber Pro were linked to this administrative data, which allowed us to compute a measure of college quality proposed by MacLeod et al. (2017). By these means, we obtain information regarding college attendance, students' major, and access to college loans.

Third, we merge records on monthly wages of graduates from college. These records are gathered by the Colombian social security administration office.[37] We merge 64% of Saber Pro students with these records. We use a sample of graduates that received their degrees between 2012 and 2016, and whose monthly wage is observed between the second and third quarters after their corresponding semester of graduation. The merged data provided us with an observation of the monthly wage after graduation for each year between 2012 and 2016. It also allow us to have information about the city or municipality where individuals work, four-digit industrial codes, and the tax-identification of employers or firms. Notice that some individuals, among Saber Pro test takers, may have received their diplomas after the period of time for our sample of graduates, or may have dropped out of college. In these two cases, we are not able to link merge them into our sample, explaining our merging results.

[34] In Colombia, individuals change their identification number when they turn 18. A large portion of individuals graduate from high school before this age and, thus, using the identification number is not enough to merge information from postsecondary education, or labor market information, with high school information. We use a crosswalk of national identification numbers between youth ids (before they turn 18) and adult ids (after they turn 18) given by the Colombian registry's office (in Spanish "Registraduría General de la Nación"). We minimize the number of fuzzy matches using this crosswalk. For the remaining sample, we rely on other merging techniques to link our available datasets.

[35] In the text we explain that each household in Colombia is given a stratum number depending on the neighborhood they reside.

[36] Before the second semester of 2006, students taking Saber 11 were allowed to choose among three foreign language tests: English, French or German. But, despite having these options, 99% of students chose English as their test. Because of this reason, nowadays the unique foreign language tested is English.

[37] We were given access to these data by the Labor Observatory for Education (OLE, in Spanish), which is part of the Ministry of Education.

Appendix Table 10: Results by Merging Technique with Saber Pro dataset

Dataset	Identity Card	Record Linkage	
High School Exit Exam (Saber 11)	281,861	905,748	(0.98)
Undergraduate Enrollees (Spadies)	1,027,802	160,435	(0.97)
Graduates Social Security Records (OLE)	922,684	12,844	(0.96)

Notes. The numbers in this table correspond to individuals merged with a total of 1,448,395 Saber Pro test takers. Record linkage merging method used names and birthdates to pair individuals. Average Jaccard index of similarity between names in parenthesis. A score of 1 is assigned to names with perfect similarity, a score of 0 is assigned to names without similarity. Birthdates we required to be equal in order to merge individuals.

After assembling the mentioned datasets with Saber Pro population, we obtain a subset of 743,542 students that belong to the intersection of the three merging processes. However, we also make some selection of individuals in order to have more homogenous samples. For instance, some individuals may have taken the college exit exam in different opportunities. These could be individuals attending different undergraduate programs. For these, we keep the information of the first observed test scores, considering that those should reflect their level of skills at the time when they entered the labor market as graduates. It is also possible that some students took more than once the high school exit exam. In that case, we leave the information of the first observed results, having in mind that some individuals could have prepared after graduating from high school to apply for funding or a scholarship. It is also possible that students took the test more than once in order to get the minimum scores required to be admitted at some undergraduate programs. Also, we drop individuals who are under 19 or above the retirement age in Colombia (i.e. 57 years old for women and 62 for men). Individuals who graduated from other undergraduate programs before 2012 are dropped as well.

We exclude observations with missing values in any of the independent variables commonly used in our regressions, remaining with a pool of 437,673 test takers. We link them to individuals in a follow-up survey of graduates that contains information of non-cognitive abilities. The survey takes a stratified probabilistic sample from the students who received their diplomas during the past year, three years ago and five years ago. Using the 2013 and 2014 surveys, we obtain information for a random sample of individuals that belong to cohorts of graduates from 2011 to 2013.

Notice that for the described sample, we observe values for gender, age, socioeconomic stratum, and also for literacy, numeracy and foreign language skills coming from both, the college and high school exit exams. However, we may not have information about their high school, their specific skills, their non-cognitive skills or current wage. Thus, from this pool or *complete sample* ($N = 437,673$), we obtained three subsets: 1) a *full sample* with formal workers for who we observe high school and current wage ($N = 363,330$); 2) a *specific sample* with formal workers for who we observe high school, current wage, and a measure of specific skill (N = 155,939); and 3) a *survey sample* with formal workers for who we observe a measure of non-cognitive skills and current wage, but for who we may not observe their high school of graduation (N = 2,401).

D Appendix: Alternative Robustness on Returns to Skills

Following Heckman et al. (2006) and Acosta et al. (2015), we additionally address the potential attenuation bias by estimating the returns using a latent skills model. This model identifies the underlying latent parameters of a set of skills and estimates the associated economic returns. Assume the following model in which wages (W) are a function of latent skills (Θ) and other observables (X_W). The reduced equation is given by:

$$W = \alpha^W \Theta + X_W \beta^W + e^W \tag{2}$$

While the latent variable is unobservable, there are measures available in the data (i.e. tests scores) that represent realizations of the latent skills:

$$T = \alpha^T \Theta + X_T \beta^T + e^T \tag{3}$$

where T is an $L \times 1$ vector of test scores, and L corresponds to the number of associated test scores per latent skill. The identification assumption states that $e^T \perp e^W$ conditional on (Θ, X_W), and that $L \geq 3$.[38]. Parameters $\alpha^W, \alpha^T, \beta^W$ and β^T, from equations (2) and (3), can be jointly estimated by maximizing the following likelihood fuction:

$$\mathcal{L} = \prod_{i=1}^{N} \int \Big\{ f_{e^W}(X_W, W, \zeta^\Theta) \times f_{e^{T_1}}(X_{T_1}, T_1, \zeta^\Theta) \times ... \times f_{e^{T_L}}(X_{T_L}, T_L, \zeta^\Theta) \Big\} dF_\Theta(\zeta^\Theta) \tag{4}$$

The results of this model (α^W) should be interpreted as the economic return of a latent skill allowing for the presence of unobserved heterogeneity (Sarzosa and Urzúa, 2016; Carneiro et al., 2003). Since the return should be read in levels of the latent skill, we use the estimated distribution of the skill, $\widehat{F_\Theta}(\zeta^\Theta)$, to rescale α^W so it can be read in terms of standard deviations. This way we can compare the results from this model, with the ones previously obtained.

Identification requires at least three test scores for each latent skill. Thus, to estimate the effect of latent literacy skills we use writing and reading scores from the college exit exam and the language (Spanish) test scores from the high school exit exam. To identify the latent effect of numeracy we use the numeracy measures from the college exams and the standardized mathematics and physics scores from the high school exit exams. For foreign language we use the test scores from the college and high school exit exams, and the social science score from the high school exit exam, since it is strongly correlated with the results on foreign language (0.46). Among the survey sample we replace the social sciences scores with self-assessed information in the graduates' follow-up survey about the respondents' abilities in listening, speaking, reading and writing in a foreign language.[39] For major-specific

[38] More details about these identification assumptions are given in Sarzosa and Urzúa (2016).

[39] The follow-up survey includes a module in which respondents assess their own ability in listening, speaking, reading and writing in a foreign language. These information is only available for the sub-sample that was surveyed.

skills we use different scores from the set of field-specific tests in the college exit exam, and a combination of scores from the high school exist exams. [40] For non-cognitive skills we divided the nine questions used to built the non-cognitive measure into three dimensions. We, therefore, exploit the full extent of the data and the large number of available tests scores to estimate the returns to latent literacy, numeracy, major-specific, and non-cognitive skills.

This model is computationally demanding and estimation times increase dramatically with sample size, controls, and the number of factors or latent skills to estimate (Sarzosa and Urzúa, 2016). Therefore, we respectively took two random samples of 36,403 and 15,585 individuals from both, the complete set of test takers working formally and the set of formal workers with specific scores available.[41] Notice as well that we use all the individuals in the survey sample to estimate the model.

Appendix Table 11 presents the estimation of the latent skill model for each sample and, for comparison, it also presents OLS estimations. Again, simultaneous and stacked results could be interpreted as lower and upper bounds. Within the survey sample, the results suggest that latent numeracy skills have the largest returns (from 4.8 percent to 7.2 percent), followed by latent literacy skills (5.1 percent to 6 percent), foreign language skills (up to 1.8 percent) and, at last, non-cognitive skills (around 2 percent). On the other hand, using the random sample representing our complete set of students, we see that numeracy returns up to 6.4 percent, literacy latent skills returns from 3.7 percent to 7.1 percent, and foreign language returns up to 12 percent. The results from the sample representing the students with field specific scores show that the latent specific skills returns up to 6.7 percent.

[40] Given the constraint in the number of test scores, we do this only for individuals who have only taken one or two specific tests and exclude those that took three. For students in health and agronomy, we use biology and chemistry test scores. For education, social sciences, humanities, and economics and business, we use history and geography tests scores. For students in arts, we use philosophy and history tests scores. For engineering, mathematics and natural sciences we use chemistry and biology.

[41] We conducted a stratified probabilistic sampling within both sets of information: 1) the set of workers who took the college exit exam between 2011 and 2015, and 2) the workers for who it is available specific tests scores. Strata were defined using the nine test editions of the college exit exam, eighteen groups of related majors, six cohorts of graduates, four groups or quartiles defined using the average score in the college exit exam, and quartiles defined by the current wage of individuals. Within each stratum, we applied a simple random sampling methodology to select 10 percent of test takers. Expansion factors were computed as $\pi_{kh} = \frac{n_h}{N_h}$, with n_h as the number of sampled individuals in stratum h, and N_h as the original number of individuals in h.

Appendix Table 11: Structural Estimations: Measurement Error Correction

	Dependent Variable: log(Current Wage)									
	OLS in Random Sample				Unobserved Heterogeneity					
	Simultaneous		Stacked		Simultaneous			Stacked		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Literacy	0.029	0.013	0.040	0.030	0.037	0.035	0.051	0.071	0.070	0.060
	[0.002]	[0.007]	[0.003]	[0.006]	[0.006]	[0.009]	[0.012]	[0.004]	[0.006]	[0.011]
Numeracy	0.032	0.033	0.042	0.047	-0.002	0.014	0.048	0.064	0.073	0.072
	[0.004]	[0.007]	[0.005]	[0.007]	[0.007]	[0.010]	[0.020]	[0.004]	[0.006]	[0.012]
English	0.018	0.016	0.031	0.031	0.096	0.126	0.011	0.049	0.058	0.018
	[0.007]	[0.009]	[0.008]	[0.008]	[0.011]	[0.016]	[0.011]	[0.003]	[0.005]	[0.011]
Field Specific		0.023		0.040		0.031			0.067	
		[0.005]		[0.005]		[0.007]			[0.005]	
Socioemotional							0.023			0.022
							[0.012]			[0.011]
Sample	Full	Specific	Full	Specific	Full	Specific	Survey	Full	Specific	Survey
Observations	36,403	15,585	36,403	15,585	36,403	15,585	2,401	36,403	15,585	2,401
Controls:										
Individual & Field	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Types of Degrees	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The dependent variable is as described in the notes of Table 1. Columns (1) and (3) use the same specifications of columns (1) and (4) of Table 3, but within a stratified random sample. Columns (2) and (4) use the specifications of columns (2) and (5) of Table 3, within another random sample. Expansion factors for each random sample were used to estimate columns (1) to (4). The point estimates from the unobserved heterogeneity model are originally computed in levels of each latent skill. Using the standard deviation of these skills, we rescale the coefficients and the standard errors applying 'delta method'. From columns (5) to (10), to estimate literacy we used: 1) writing scores from the college exit exam, 2) reading scores from the college exam, and 3) language (Spanish) scores from the high school exit exam. Numeracy used: 1) quantitative reasoning scores from the college exam, 2) mathematics scores from high school exam, and 3) physics scores from the high school exam. Foreign language used: 1) foreign language scores from the college exam, 2) foreign language scores from high school exam, and 3) a factor score predicted using self-reported ability in a foreign language for columns (7) and (10). For columns (5), (6), (8) and (9) we used as the social science scores from the high school exam as the third score for foreign language. For non-cognitive estimations, we divided the nine categorical questions of Appendix Table 2 into three scores computed using each time a factor analysis model. Major specific skills used the specific test scores from the college exit exam. When a student has only taken one or two specific tests in the college exam, we used the set of test scores from the high school exam depending on the study area of the student (see footnote 26). Stacked results refer to separe regressions for each ability measure. Field of study and type of degrees controls as described in the notes of Table **??**, however individual controls in columns (5) to (10) do not include high school fixed effects and the initial ability proxy.