ESTIMATING THE COLLEGE INCOME PREMIUM ACCOUNTING FOR COGNITIVE AND NON-COGNITIVE LATENT SKILLS[1]

Marcos Ki Hyung Lee[2]
Bruno Kawaoka Komatsu[3]
Naercio Menezes Filho[4]

In this paper, we estimate the returns to higher education in Brazil, controlling for the effects of latent abilities. We leverage a unique data that measures cognitive and noncognitive abilities to estimate a structural factor model that accounts for the effects of latent cognitive and non-cognitive skills. We find that college returns are on average 43% and they are positive even for the lowest levels of cognitive and non-cognitive skills. Additionally, both cognitive and non-cognitive skills influence the probability of having college education, but cognitive abilities have a larger effect. Among the college educated workers, cognitive skills are more important for earnings, but among the unskilled workers, non-cognitive skills are more rewarded.

Keywords: cognitive skill; non-cognitive skill; education; unobserved heterogeneity; income.

ESTIMANDO O PRÊMIO DE RENDA DO ENSINO SUPERIOR CONSIDERANDO AS HABILIDADES LATENTES COGNITIVAS E NÃO COGNITIVAS

Neste artigo, estimamos os retornos de renda da educação superior no Brasil, controlando pelos efeitos das habilidades latentes. Utilizamos um conjunto de dados único que mede habilidades cognitivas e não cognitivas para estimar um modelo estrutural de fatores que leva em conta esses efeitos. Nossos resultados indicam que o retorno da educação superior é, em média, de 43%, sendo positivo mesmo para os menores níveis de habilidades cognitivas e não cognitivas. Além disso, ambas as habilidades influenciam a probabilidade de obter ensino superior, mas as habilidades cognitivas têm um efeito maior. Entre os trabalhadores com ensino superior, as habilidades cognitivas são mais relevantes para os rendimentos, enquanto entre os trabalhadores não qualificados, as habilidades não cognitivas são mais recompensadas.

Palavras-chave: habilidades cognitivas; habilidades não cognitivas; educação; heterogeneidade não observada; renda.

JEL: J24; D31; I24.

1 INTRODUCTION

The college wage premium is large and increasing in developing countries. In 2014, a worker with college degree earned on average 2.4 times the wage of a worker with high school degree in Brazil and Colombia, while in Organisation for Economic Co-Operation and Development (OECD) countries that ratio

1. DOI: http://dx.doi.org/10.38116/ppe55n2art1.

2. Post-doctoral research fellow at Department of Economics of the Queen's University. E-mail: marcos.k@queensu.ca.

3. Researcher at Insper. E-mail: brunokk@uol.com.br.

4. Ruth Cardoso Chair professor at Insper; and associate professor at Faculdade de Economia, Administração, Contabilidade e Atuária of the Universidade de São Paulo (FEA/USP). E-mail: naercioamf@insper.edu.br.

was 1.6 (OECD, 2022). By 2020, those disparities grew even larger, increasing even further the individual incentive for high-school graduates to continue their studies into the higher education. Despite these trends, in Brazil, higher educational attainment is still relatively low. In 2022, 23% of the adults have tertiary education, below Latin American countries as Chile, Colombia, and Mexico – whose percents are, respectively, 41%, 34%, and 27% –, and well below the OECD average of 47% (OECD, 2023). Brazilian students who could go to higher education might earn a lot more income, but they either fail to clearly evaluate the college wage premium in the future (for examples, because of uncertainties about future earnings), or some binding constraint might prevent them from continuing the education path (Binelli and Menezes-Filho, 2019).

In this paper, we investigate the role of latent skills in the educational decision. An individual's cognitive and non-cognitive skills affect their schooling decisions by influencing both its costs and the expected benefits (Carneiro, Heckman and Vytlacil, 2011). Psychic costs of doing higher education are higher to those with low levels of non-cognitive skills (Carneiro, Hansen and Heckman, 2003). Moreover, these skills have a direct impact on earnings, as individuals use them at work and other activities, increasing their productivity, and are paid accordingly in the labor market (Carneiro, Heckman and Vytlacil, 2011; Cunha and Heckman, 2007; Edin et al., 2022; Heckman, Stixrud and Urzúa, 2006; Urzúa, 2008). Skills also influence behaviors related to other sources of income as business investment decisions (Abay, Blalock and Berhane, 2017), personal finance decisions (Luik and Steinhardt, 2016; Strömbäck et al., 2017), wealth and retirement saving (Banks, O'Dea and Oldfiled, 2011), and risky behaviors (Gong and Zhu, 2019). Despite this significant role, there is still scarce evidence on the impacts of skills on returns to education, especially in developing countries (Krishnakumar and Nogales, 2020).

This paper contributes to fill that gap. We employ a structural model that accounts for the effects of latent factors on a unique dataset with measures of cognitive and non-cognitive skills, to estimate the average treatment effect (ATE) of college education on individual income over the distributions of cognitive and non-cognitive skills. We first estimate the effects of cognitive and non-cognitive skills on the probability of college education and estimate the effects of these skills on wage for workers with and without a college degree. We then use the estimated parameter to simulate to estimate the ATE for multiple levels of skills. That model has three main advantages. First, it provides a parsimonious factor structure to identify the effect of unobserved heterogeneity that do not depend on linearity or distributional assumptions. This structure allows us to adjust for biases due to measurement errors in scores of cognitive and non-cognitive skills that can be large (Cunningham, Torrado and Sarzosa, 2016; Heckman, Stixrud and Urzúa, 2006; Kejriwal, Li and Totty, 2024). Moreover, it directly models the schooling choice, accounting for the unobserved heterogeneity related to latent skills. Third, it allows us to estimate the effects of observable characteristics and latent factors for each potential outcome of the treatment. To our knowledge, this is the first study to estimate the effects of attending college on income that explicitly controls for latent cognitive and non-cognitive skills in Brazil.

Our main finding is an overall ATE of 43% of the college on income, after controlling for skill levels. To grasp an idea of this magnitude, it is similar to estimates of the returns to public and private college degree in Chile (respectively, 54.5% and 34.5%), in a study that also uses a model of non-observable heterogeneity (Rodríguez, Urzúa and Reyes, 2016).[5] We also find that the ATE of college education on income increases over the distributions of cognitive and non-cognitive skills, rising faster over the distribution of the former. The college ATE varies from 30% in the lowest decile of cognitive skills to 51% in the highest decile. In the distribution of non-cognitive skills, the correspondent values are very similar, 42% and 44%. These indicate that for both groups of skills, enrolling in college make individuals better off in terms of income, but this decision is more valuable for those with higher cognitive skills.

We also estimate the direct impacts of latent skills on educational choice and income. Our finding indicate that those skills influence the decision of entering higher education in Brazil, with a larger influence of cognitive skills. An increase of one standard deviation in the level of cognitive skills increases the probability of having college education by 81 percentage points (p.p.). A variation of the same magnitude in the level of non-cognitive skills raises the probability of having a college degree by 24 p.p. Our results also suggest that occupations taken up by college educated workers value skills more than those of unskilled workers. In this sense, a one standard deviation increases in the levels of cognitive and non-cognitive skills lead to increases in income of, respectively, 17% and 14% in the scenario with college education, but in the alternative scenario, the effects are respectively 10% and 14%.

This paper contributes mainly to the vast literature of access and returns to college education (Card, 2001; Oreopoulos and Petronijevic, 2013; Psacharopoulos and Patrinos, 2018), and more specifically to the recent and still scarce literature that estimates access and returns to education controlling for latent skills. A branch of this literature uses structural factor models to account for the unobserved heterogeneity of latent cognitive and non-cognitive skills. Heckman, Stixrud and Urzúa (2006) use United States longitudinal data and low dimension vectors of cognitive and

5. Our result is not strictly comparable to those of Rodríguez, Urzúa and Reyes (2016) as we assess the effects of having any college education, even not complete, as compared to lower levels of education, while that paper estimate the impact of having a college degree, compared to having a high school degree.

non-cognitive skills to show that non-cognitive skills are as important as cognitive skills to explain schooling decisions, wages and other labor market outcomes, and risky behaviors. Heckman, Humphries and Veramendi (2018) estimate the wages returns to education with a robust dynamic model of educational decisions that combine a structural factor model to reduced-form treatment effect models, finding that ATE of college degree (compared to some college education) is increasing in the distribution of cognitive and non-cognitive skills, but the variation over the latter is much smaller.

Other recent studies estimate returns to marginal expansions in higher education (MTE) and deal with the endogeneity of college schooling decision with a local instrumental variables strategy (Carneiro, Heckman and Vytlacil, 2011; Nybom, 2017). They use measures of cognitive and non-cognitive skills to control for heterogeneity in skills, even if those measures have measurement errors (Cunningham, Torrado and Sarzosa, 2016; Heckman, Stixrud and Urzúa, 2006; Kejriwal, Li and Totty, 2024). Our main contribution consists in estimating returns to education in a developing country, controlling for unobserved heterogeneity of latent skills and dealing directly with the endogeneity of schooling choices.

We also contribute to a recent and growing branch of the literature that studies the returns to cognitive and non-cognitive skills in developing countries, using structural factor models. Acosta, Muller and Sarzosa (2015) use Colombian data to examine how cognitive and non-cognitive skills correlate with various labor market outcomes and find that while non-cognitive skills correlate more strongly with probabilities of labor market participation and employment, cognitive skills strongly correlates to earnings and the probabilities of having a formal or qualified job. Cunningham, Torrado and Sarzosa (2016) investigate a similar question with Peruvian data and further verify that some dimensions of cognitive and non-cognitive skills have stronger correlations with labor market outcomes. Krishnakumar and Nogales (2020) adapts the framework of technology of skill formation to estimate the impacts of cognitive and non-cognitive skills on work-related well-being with data from Bolivia. Their structural factor model also specifies the outcome variable as a latent variable related to multiple aspects of a job. They find that cognitive and non-cognitive skills affect job-related well-being, especially cognitive skills. We also estimate the direct impacts of cognitive and non-cognitive skills on income in Brazil, separately for counterfactual scenarios with and without college education.

Finally, we dialog with the literature of returns to higher education in Latin-American countries, which experienced rapid expansions of college enrollments in the last decades. González-Velosa et al. (2015) verifies that the rapid expansion in higher education in Colombia and Chile generated an optimistic expectation that, however, was not met by labor productivity, economic and social equity gains. As access to higher education expanded, authors hypothesize that the quality of new higher education institutions and students might have declined and realized lower or negative gains. Carranza and Ferreyra (2019) confirms that in Colombia, expansion of higher education was driven by higher education supply side and allowed the inflow of low-ability students. Consistent with these developments, Montoya, Noton and Solis (2017) use an regression discontinuity design (RDD) strategy and find no effect of college versus vocational education in Chile on labor market outcomes and conclude that given higher costs of college education, students that went to college might have incurred income losses. We contribute to this literature by estimating the ATE of college education in Brazil, and verify to which levels of abilities higher education bring benefits.

The rest of this paper is organized as follows. Section 2 describes the structural factor model and section 3, our database. In section 4, we present our results and in section 5, some conclusions.

2 METHODOLOGY

In this paper, we aim to estimate the average income returns to tertiary education, using the routine developed by Sarzosa and Urzúa (2016). Ordinary Least Squares (OLS) regressions of income on a dummy variable for college education usually do not identify the income return of college education, because of the endogeneity introduced by the presence, in the error term, of cognitive and non-cognitive abilities that influence schooling decisions and have direct impacts on earnings. Augmenting the equation by using measures of those latent skills as proxies might still generate biased estimates, as test scores of skills incorporate large measurement errors, specially the scores for non-cognitive skills (Cunningham, Torrado and Sarzosa, 2016; Heckman, Stixrud and Urzúa, 2006; Kejriwal, Li and Totty, 2024). Moreover, if the measures of skills are taken after schooling decisions are made, reverse causality might also be present, as levels of schooling affect the scores (Heckman, Stixrud and Urzúa, 2006).

We deal with each of those issues. First, we model the endogenous schooling choice, accounting for the effects of cognitive and non-cognitive skills. We also estimate the direct impacts of those skills on income, given the schooling levels, in counterfactual scenarios for each of those levels. Moreover, we use a latent factor structural model to deal with measurement error in the levels of skills, based on a measurement system of test scores (Carneiro, Hansen and Heckman, 2003; Heckman, Stixrud and Urzúa, 2006; Sarzosa and Urzúa, 2016), leveraging the cognitive and non-cognitive test results in our data. Finally, we simulate a counterfactual scenario of each choice, based on the results on parameters' distributions, to estimate the ATE of tertiary education on income. We also examine the heterogeneity of ATE over the distribution of cognitive and non-cognitive skills. This approach allows us to deal with the most common OLS biases when estimating the returns of tertiary education.

We first model the attending tertiary education choice as a Roy model (Roy, 1951). Formally, let D be an indicator variable that is equal to 1 if the individual has tertiary education, and let Y_1 and Y_0 be the potential income that would be observed if the individual, respectively, has ($D = 1$) or do not have ($D = 0$) tertiary education. Individuals choose their schooling level comparing the life-time utility levels of each of those treatment states, which might depend on observable characteristics (X_D) and latent skills (we denote cognitive and non-cognitive latent factors, respectively, as θ^C and θ^N). We model that relationship as a linear combination and define the utility threshold for attending college to be 0, for each individual $i = 1, \ldots, M$:

$$Di = 1\left(X_{Di}\beta^D + \alpha^{D,C}\theta_i^C + \alpha^{D,N}\theta_i^N + e_i^D > 0\right) \tag{1}$$

In equation (1), $1(\cdot)$ is the indicator function, β^D, $\alpha^{D,C}$ and $\alpha^{D,N}$ are coefficients that relate observable characteristics and latent factors to the level of individual utility, and e_i^D is the error term. Here, we assume $e_i^D \perp \left(X_i^D, \theta_i^C, \theta_i^N\right)$.

The main equation in the model relates individual income to observable characteristics (X_Y) and to the latent factors. Moreover, the loadings that mediate that relationship might differ for each treatment state D:

$$Y_{1i} = \left(X_{Yi}\beta^{Y_1} + \alpha^{Y_1,C}\theta_i^C + \alpha^{Y_1,N}\theta_i^N + e_i^{Y_1}\right) \times 1(D_i = 1) \tag{2}$$
$$Y_{0i} = \left(X_{Yi}\beta^{Y_0} + \alpha^{Y_0,C}\theta_i^C + \alpha^{Y_0,N}\theta_i^N + e_i^{Y_0}\right) \times 1(D_i = 0)$$

In equation (2), β^{Y_1}, $\alpha^{Y_1,C}$, $\alpha^{Y_1,N}$, β^{Y_0}, $\alpha^{Y_0,C}$, $\alpha^{Y_0,N}$ are the coefficients that relate the observable characteristics and latent factors to personal income, and $e_i^{Y_1}$ and $e_i^{Y_0}$ are error terms. We assume $e_i^{Y_1}$ and $e_i^{Y_0}$ orthogonal to θ_i^C, θ_i^N, and X_{Yi}, and $e_i^{Y_l} \perp e_j^{Y_l}, \forall i, j = 1, \ldots, M$ and $\forall l = \{0,1\}$. Moreover, we assume $\left(e_i^{Y_1}, e_i^{Y_0}\right) \perp e_i^D$, $(X_{Di}, X_{Yi}) \perp \left(\theta_i^C, \theta_i^N, e_i^{Y_1}, e_i^{Y_0}, e_i^D\right)$.

The need for the structural model comes from the fact that latent factors are not observable. The model incorporates a system of test scores of latent factors to estimate the distribution parameters of each factor. Let L be the number of those scores, so that we represent the vector of full set of test scores as $T = (T_1, T_2, \ldots, T_L)$. The measurement system of test scores is:

$$T_i = X_{Ti}\beta^T + \alpha^{T,C}\theta_i^C + \alpha^{T,N}\theta_i^N + e_i^T \tag{3}$$

In equation (3), $\alpha^{T,C}$ and $\alpha^{T,N}$ are vectors of the coefficients that relate latent factors to their scores and e_i^T is the error term. We assume $e_i^T \perp \left(\theta_i^C, \theta_i^N, X_T\right)$, elements in e_i^T are mutually independent ($e_i^h \perp e_j^h, \forall i, j = 1, \ldots, M$), and each one has an associated distribution $f_{e_i^h}(\cdot)$, $h \in \{1, \ldots, L\}$.

Estimation of the parameters of system 3 includes the distribution parameters of the latent factors and depend on three assumptions for identification. First, factors are orthogonal ($\theta_i^C \perp \theta_i^N$). Second, the number of test scores L is at least so that $L(L-1)/(2(L+1)) \leq k$, where k is the number of latent factors. In our case, we mainly deal with two factors (the cognitive and the non-cognitive factors), so that the minimum necessary number of test scores is $L = 6$. Third, it is necessary to normalize one of loadings to one for each set of scores of latent factors, as latent abilities have no intrinsic measurement unit (Carneiro, Hansen and Heckman, 2003; Heckman et al., 2014; Sarzosa and Urzúa, 2016).

Two stages maximum likelihood estimation (MLE) is used to fit model 3 and recover the parameters of the measurement system ($\beta^T, \alpha^{T,C}, \alpha^{T,N}$) and the distribution of latent factors in the first stage. Under the model hypothesis, distributional parameters of latent factors are recovered and their distributions are non-parametrically estimated as a mixture of two normal distributions. In the second stage, the likelihood function to estimate incorporates the estimated distributions of the latent factors and MLE estimation recovers the parameters of income equations 2 (β^{Y_1}, $\alpha^{Y_1,C}$, $\alpha^{Y_1,N}$, β^{Y_0}, $\alpha^{Y_0,C}$, $\alpha^{Y_0,N}$).

As latent skills are not observable, we are not able to predict outcomes for each individual to estimate the ATE. We then simulate the counterfactual scenarios based on estimates of the distributional parameters of latent skills and of the other parameters and, using these simulations, we estimate the ATE of attending college as (Heckman et al., 2014):

$$ATE = \int\int E\left(Y_1 - Y_0 \mid X_Y = x, \theta = \overline{\theta}\right) d\hat{F}_{X_Y,\theta}(x,\theta) \tag{4}$$

In equation (4), $\theta = (\theta^C, \theta^N)$ and $\hat{F}_{X_Y,\theta}$ is the estimated joint distribution of X_Y and θ. We use 1.000 repetitions in each simulation and calculate the overall ATE, as well as the ATE for each tenth of the distribution of the cognitive and non-cognitive skills.

3 DATA

This study uses the 2015 wave of the Indicator of Functional Literacy (Indicador de Alfabetismo Funcional – Inaf) household survey, which aims to measure the level of functional literacy of the Brazilian adult population. Inaf collects socioeconomic data and cognitive tests from a representative sample of 15 to 64 years-old people from all Brazilian regions. The cognitive tests consist of 32 reading, writing and numeracy items that capture how individuals use those abilities in their regular activities to assess functional literacy levels. As measures of cognitive skills, we use standardized literacy and numeracy scores, which are generated from collected items using the Theory of Response to Item, and also the self-reported ability in using a computer.[6]

6. To evaluate the level of ability using computer, Inaf uses a five-level scale, that varies from the lowest level of "incapable of using a computer" to "uses a computer with no difficulty".

Especially in the 2015 wave, Inaf also collected information on three dimensions of noncognitive abilities that more strongly influences the performance on cognitive tests – self-concept, self-regulation, and openness to experience (Lima, 2016). Self-concept measures the tendency to think about ourselves and about our abilities in a positive way, while self-regulation captures the tendency to be organized, focused, persistent, and self-disciplined. Openness relates to the tendency to be imaginative, curious, and interested in learning. Inaf used the Core Self-Evaluations Scale (Judge et al., 2003) to capture the self-concept dimension and applied an instrument especially adapted for Brazil by the Ayrton Senna Institute, Senna 1.0, to measure self-regulation and openness to experience. Senna 1.0 is a self-reported scale, designed to measure the Big Five personality dimensions and was based on eight previous instruments (Lima, 2016; Primi et al., 2016). Each of the dimensions are measured by three or four scores, and Inaf data also has an aggregation of those scores into measures for each dimension. In our empirical exercises, we use the aggregate measures of the three dimensions as scores for the non-cognitive latent skills, but in the exercises where we detail the role of each of those dimensions, we use the original scores associated to them. Inaf is carried out by Paulo Montenegro Institute and by the Non-Governmental Organization (NGO) Ação Educativa, and non-cognitive abilities tests were carried out by Ayrton Senna Institute.

Our variable of tertiary education is a dummy variable that equals one for any years of college education and zero for all other levels of schooling. We restrict our analysis to individuals aged 22 years or older, as they are more likely to have completed their education in the moment of the interview. All measures of cognitive and non-cognitive skills are standardized, so that the results are in terms of standard deviations. We use other individual's observable characteristics as covariates in our models in equations 1, 3, and 2. They are obtained from the socioeconomic part of Inaf's questionnaire and includes age (dummy variables for age the groups 22-39 years, 40-59 years, and 60 or more years), race (dummy variable for African Brazilians or native people), gender (dummy variable for females) and parents schooling (dummy variables for incomplete middle school, for complete middle school, and complete high school). The total data has about 2,000 observations, but there were approximately 600 observations with no valid information on income, and about 300 with no valid information on any of the covariates we use. Moreover, about 150 observations had age below 22 years, so we also excluded them. Our final sample has 920 observations.

Table 1 shows descriptive statistics of all variables used in the analysis. In general, those with college education have higher levels of income and of scores for cognitive and non-cognitive skills. They also have lower age when started school, are younger, and have more educated parents. Approximately 20% of observations had missing values for the exact levels of individual income, so we input the middle point of the income band they declared to be in.[7] Following Krishnakumar and Nogales (2020), we consider the age when a person entered school as a measure of parents' investment in education that might affect income through schooling choices and we include it as a control in equation 1.

One limitation to our results is that cognitive and non-cognitive abilities are not measured at the time individuals choose whether to start tertiary education. Moreover, as our data is cross-sectional, it is possible that reverse causality issues are present and our model do not address this issue, that is, education might affect cognitive and non-cognitive latent skills. One key assumption to interpret estimates as causal is that education does not affect skills, so that education achievement is a signal of preexistent cognitive and non-cognitive skills Heckman, Stixrud and Urzúa (2006) and Oreopoulos and Petronijevic (2013). Another limitation is that we only have measures of three dimensions of non-cognitive skills, while previous studies usually use at least six of those dimensions (Acosta, Muller and Sarzosa, 2015; Cunningham, Torrado and Sarzosa, 2016), so that there are still aspects of noncognitive skills that are not controlled for in our regressions.

TABLE 1
Summary statistics

	College education			Other		
	Observations	Mean	Standard deviation	Observations	Mean	Standard deviation
Income	202	2,445	2,202	718	1,388	1,336
Literacy proficiency	202	0.63	0.79	718	-0.18	0.98
Numeracy proficiency	202	0.66	0.70	718	-0.19	0.99
Computer ability	202	0.70	0.42	718	-0.20	1.03
Self-Regulation	202	0.20	1.07	718	-0.06	0.97
Openness	202	0.35	0.98	718	-0.10	0.98
Self-Concept	202	0.16	1.12	718	-0.04	0.96
Age when started school	202	7.06	3.37	718	7.63	2.63
22-39 years old	202	0.62	0.49	718	0.49	0.50
40-59 years old	202	0.33	0.47	718	0.41	0.49
60 years old or more	202	0.05	0.22	718	0.10	0.30
Black or indigenous	202	0.50	0.50	718	0.60	0.49
Female	202	0.51	0.50	718	0.43	0.50
Mother did not complete middle school	202	0.39	0.49	718	0.78	0.42
Mother completed middle school	202	0.18	0.39	718	0.11	0.31
Mother completed high school	202	0.43	0.50	718	0.12	0.32

(Continues)

7. Our main qualitative results do not change when we estimate only with observations whose income was not inputted (appendix A, table A.1).

(Continued)

	College education			Other		
	Observations	Mean	Standard deviation	Observations	Mean	Standard deviation
Father did not complete middle school	202	0.44	0.50	718	0.82	0.39
Father completed middle school	202	0.23	0.42	718	0.08	0.28
Father completed high school	202	0.34	0.47	718	0.10	0.30
North region	202	0.05	0.23	718	0.07	0.26
Northeast region	202	0.21	0.41	718	0.25	0.43
Southeast region	202	0.50	0.50	718	0.45	0.50
South region	202	0.17	0.38	718	0.16	0.37
Midwest region	202	0.06	0.25	718	0.07	0.25

Authors' elaboration.
Note: Summary statistics of variable used in the analysis. Sample of 22 to 65 years-old individuals. Income represents individual income and is in Brazilian 2015 R$. Tertiary education is a dummy variable for tertiary education, including those who did not complete it. Literacy, numeracy, and computer ability are scores for cognitive ability, self-regulation, openness, and self-concept are scores for non-cognitive abilities. All scores are standardized.

4 RESULTS

4.1 Importance of latent skills

We first show OLS estimations of log income on observable characteristics and scores for cognitive and non-cognitive abilities in table 2. The main results are that while tertiary education has a large association with income, which is highly sensitive to cognitive ability measures, is not affected when we introduce controls for non-cognitive skills. Moreover, while both cognitive and noncognitive abilities are important determinants of income, not all their dimensions are as relevant.

Column 1 shows that a tertiary education has a positive association with income, controlling for the selected observable characteristics. In addition, higher income levels are positively associated with middle-aged individuals and those with higher parent's schooling levels, and negatively associated with being a female. In columns 2 and 3, we introduce abilities' scores, one type of ability at a time. In column 2, when we introduce cognitive abilities' scores, we observe that the magnitude of point estimate of tertiary education drops about 20% from the level of column 1. Numeracy and ability with a computer are significantly correlated with income, but the estimate for computer ability is larger. In column 3, we control for non-cognitive abilities' scores and not for measures of cognitive skills and observe that the estimate for college education returns to a magnitude slightly smaller (about 6%) than estimate in column 1. This result is consistent to that Buchmueller (2019) find with data from the United Kingdom, that controlling for non-cognitive skills does not change estimates of returns to schooling. We also observe that point estimates for the three scores are positive, but only that for openness to experience is significant. In column 4, we control for all scores at the same time and results from columns 2 and 3 persist, but with smaller magnitudes. Estimate for college education reduces marginally (about 4%) compared to estimate in column 2. According to estimates in column 4, college education is associated with a level of income 36% higher than lower levels of schooling, and a one standard deviation increase in scores of numeracy, computer score, and openness are associated with, respectively, about 0.3%, 5%, and 7% higher levels of income. These results indicate that skills deeply affect the relationship between schooling and income and are relevant to explain income, even after considering differences in schooling levels.

TABLE 2
OLS estimations: income and latent skills

Independent variables	Dependent variable: log income			
	(1)	(2)	(3)	(4)
Tertiary education	0.474*** (0.056)	0.375*** (0.056)	0.445*** (0.054)	0.361*** (0.055)
Literacy proficiency	-	0.000 (0.001)	-	-0.000 (0.001)
Numeracy proficiency	-	0.003** (0.001)	-	0.003** (0.001)
Computer ability	-	0.058*** (0.014)	-	0.052*** (0.014)
Self-Regulation	-	-	0.031 (0.024)	0.030 (0.023)
Openness	-	-	0.090*** (0.022)	0.074*** (0.022)
Self-concept	-	-	0.019 (0.020)	0.016 (0.020)
Age when started school	-0.006 (0.006)	-0.001 (0.005)	-0.004 (0.005)	0.000 (0.005)
40-59 years old	0.142*** (0.045)	0.184*** (0.044)	0.145*** (0.044)	0.182*** (0.044)
60 years old or more	0.049 (0.066)	0.188*** (0.070)	0.062 (0.065)	0.183*** (0.069)
Black or indigenous	-0.064 (0.045)	-0.038 (0.044)	-0.052 (0.044)	-0.031 (0.043)
Female	-0.380*** (0.040)	-0.374*** (0.040)	-0.387*** (0.041)	-0.381*** (0.040)
Mother completed middle school	0.129** (0.062)	0.072 (0.063)	0.111* (0.061)	0.064 (0.061)
Mother completed high school	0.067 (0.063)	0.024 (0.062)	0.049 (0.062)	0.014 (0.061)

(Continues)

(Continued)

Independent variables	Dependent variable: log income			
	(1)	(2)	(3)	(4)
Father completed middle school	0.231*** (0.064)	0.172*** (0.064)	0.216*** (0.063)	0.165*** (0.063)
Father completed high school	0.107 (0.072)	0.050 (0.072)	0.077 (0.070)	0.032 (0.070)
Constant	7.143*** (0.063)	6.627*** (0.119)	7.129*** (0.061)	6.684*** (0.117)
Observations	920	920	920	920
R-squared	0.256	0.287	0.279	0.303

Authors' elaboration.

Note: 1. OLS estimations of log income on observable characteristics and test scores. Regressions further controlled for regional fixed effects. Sample of 22 to 65 years-old individuals. Literacy, numeracy, and computer ability are scores for cognitive ability, self-regulation, openness, and self-concept are scores for non-cognitive abilities. All scores are standardized.
2. Robust standard errors in parenthesis.
3. Significance: *** $p < 0.01$; ** $p < 0.05$; and * $p < 0.10$.

4.2 Main results

Table 3 shows structural estimates for the association between schooling choice, income and latent factors (equations 1 and 2). We have two main results. First, both cognitive and non-cognitive skills influence the probability of having tertiary education, but cognitive abilities have a larger effect. Second, the direct association between cognitive and non-cognitive abilities and income depends on the schooling level. For the college graduate outcome, both cognitive skills and noncognitive skills estimates are larger than those when schooling level is lower, and the effect of cognitive skills is relatively larger. On the other hand, for the scenario with lower schooling level, non-cognitive skills are more important to determine income.

TABLE 3
Structural estimations: income and latent skills

Independent variables	Dependent variable		
	Terciary education	Log income	
		Tertiary education	Other
	(1)	(2)	(3)
Cognitive	0.807*** (0.109)	0.168* (0.089)	0.099*** (0.032)
Non-cognitive	0.240** (0.112)	0.144* (0.079)	0.137*** (0.049)
Age when started school	-0.017 (0.020)	-	-
40-59 years old	-0.003 (0.119)	0.379*** (0.094)	0.094** (0.047)
60 years old or more	-0.268 (0.233)	0.718*** (0.202)	-0.047 (0.076)
Black or indigenous	-0.261** (0.125)	-0.075 (0.101)	-0.083* (0.049)

(Continues)

(Continued)

Independent variables	Dependent variable		
	Terciary education	Log income	
		Tertiary education	Other
	(1)	(2)	(3)
Female	0.310*** (0.110)	-0.297*** (0.084)	-0.412*** (0.044)
Mother completed middle school	0.509*** (0.166)	-0.084 (0.130)	0.177** (0.077)
Mother completed high school	0.893*** (0.157)	0.070 (0.109)	0.073 (0.082)
Father completed middle school	0.789*** (0.165)	0.290** (0.118)	0.242*** (0.084)
Father completed high school	0.693*** (0.165)	0.084 (0.115)	0.206** (0.087)
Constant	-1.627*** (0.307)	7.275*** (0.226)	6.971*** (0.098)
Observations	920	920	920

Authors' elaboration.

Note: 1. Structural estimations of log income on observable characteristics and latent factors. Regressions further controlled for regional fixed effects. Sample of 22 to 65 years-old individuals. Literacy, numeracy, and computer ability are scores for cognitive ability, self-regulation, openness, and self-concept are scores for non-cognitive abilities. Column 1 shows results for the schooling choice (equation 1) and columns 2 and 3 show estimates for, respectively, individuals that concluded tertiary education and the others.

2. Standard errors in parenthesis.

3. Significance: *** $p < 0.01$; ** $p < 0.05$; and * $p < 0.10$.

Column 1 shows estimates for the schooling choice equation 1. Cognitive and non-cognitive skills are both positively associated with choosing tertiary education. One standard deviation increases in the levels of cognitive and non-cognitive skills are associated with, respectively, increases of 80 p.p. and 24 p.p. in the probability of having tertiary education. These results indicate that cognitive abilities matter more to have college education in Brazil, which is consistent with previous results in the literature (Nybom, 2017). Moreover, estimates on other observable characteristics indicate that females and people with more educated parents have higher probabilities of getting college education, while black or Indigenous people have lower probabilities. The age that individuals entered school and dummies for current age groups are all negative, but are not significant.

Columns 2 and 3 show estimates of the relationships between the latent skills and personal income in the counterfactual scenarios of having or not tertiary education. Estimates for cognitive and non-cognitive skills are positive and significantly different from zero in both scenarios. Increments of one standard deviation in cognitive and non-cognitive skills are respectively associated with increases of 17% and 14% in income. Moreover, in the scenario without college education, one standard deviation raises in cognitive and non-cognitive abilities are associated with increments of 10% and 14% in income. Despite differences in point estimates, it is difficult to extract any conclusion about their relative sizes, as estimates are noisy. However, these results follow a similar pattern to those Heckman, Stixrud and Urzúa (2006) found for impacts of cognitive and non-cognitive skills on hourly wages in the United States, especially among male workers.

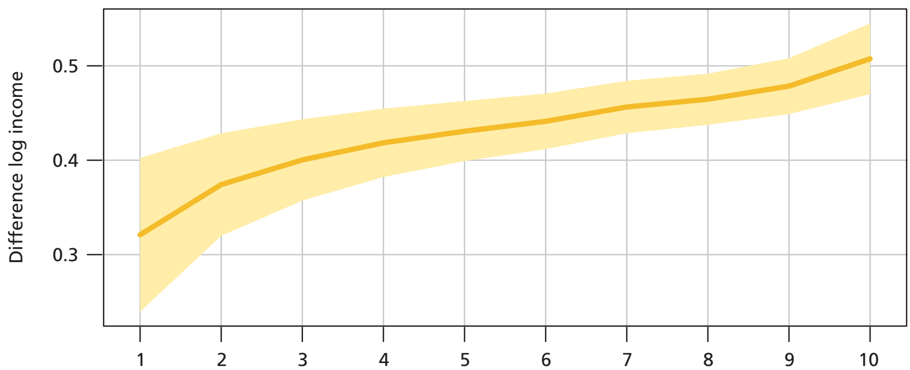
As for the observable characteristics, estimates for females are negative in both scenarios, which shows that there is still a gender gap, even after controlling for abilities. These results are consistent with previous estimations of negative correlations of the female gender with earnings, probabilities of having a job or being in a job with good characteristics (formal, wage worker, or white collar) in developing countries, even after controlling for latent skills (Acosta, Muller and Sarzosa, 2015; Cunningham, Torrado and Sarzosa, 2016). In addition, in both scenarios, people with more educated parents have higher income. More educated parents might affect income in a variety of ways. As education is highly correlated with income, children of more educated parents might borrow or receive more resources from their parents to start their own business or to take riskier career decisions. They might also inherit parents' business or their contacts, which facilitate transitions to better jobs. Moreover, estimates for black or Indigenous are negative, but is only significant in the scenario with no college education. This result is consistent with evidence that racial earnings gap remains even after controlling for abilities in the United States (Urzúa, 2008).

We assess the impact of tertiary education of individual income by simulating the counterfactual scenarios, based on previously estimated parameters, for each decile of the distributions of cognitive and non-cognitive skills. The overall impact of having tertiary education on income is a 43% increase – with a 95% confidence interval (CI) of [40%,46%]. This effect is larger than that of OLS estimates in table 2 and means that once we exclude measurement errors in skills' scores and control for the endogeneity of schooling choice, the impact of higher education increases. To assess the goodness of fit of our model, appendix A, table A.2 compares income means by schooling levels, for original data and the simulated data and shows that the predicted values in simulations are close to the observed for both levels of schooling.

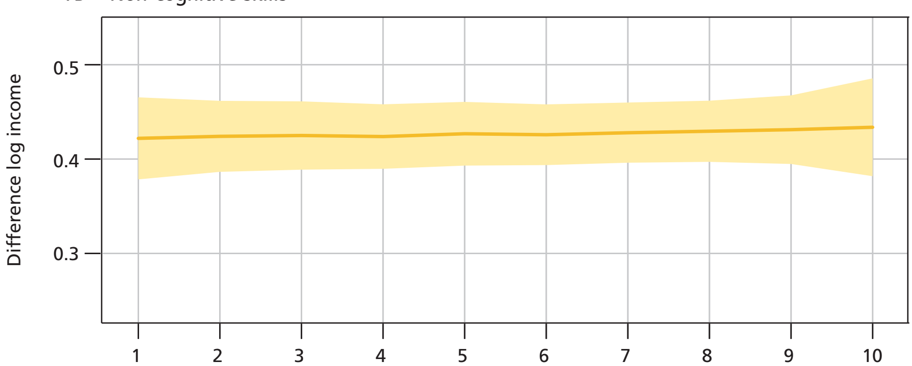
Figure 1 shows that when we evaluate the heterogeneity of the college education ATE separately over the distribution of each latent skill, the effect increases with cognitive skills, but remain stable over the distribution of non-cognitive abilities. Panel 1A shows that the magnitude of the impact over cognitive skills distribution varies from 32% in the lowest decile to 51% in the highest decile. In contrast, panel 1B shows that the variation of the ATE over the distribution of non-cognitive skill remain within the 95% CI. In the appendix A, figure A.1, we plot averages of the predicted log income by tenth of the distribution of cognitive and non-cognitive skills and those results help explain patterns observed in figure 1.

As average predicted income grows faster over the distribution of cognitive skills for college education than for other education levels, the ATE of college education also increases over that distribution. In the case of non-cognitive skills, the trajectory of average predicted income is almost the same in both levels of education, so that the ATE of college education does not vary as much.

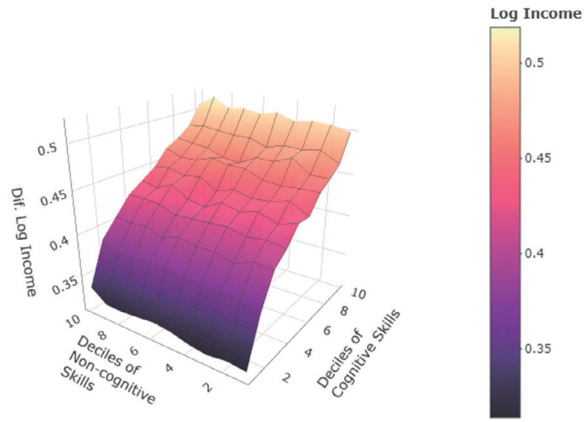
FIGURE 1
Effects of tertiary education on log income by levels of cognitive and non-cognitive skills
1A – Cognitive skills



1B – Non-cognitive skills



1C – Cognitive and non-cognitive skills



Authors' elaboration.

Panel 1C shows how the effect of tertiary education varies over the tenths of the joint distributions of cognitive and non-cognitive skills. The effect of college education has the lowest value (31%) in the lowest deciles of the distributions of both cognitive and non-cognitive skills. On the opposite position in both distributions, the highest tenths the impact of college education has its highest value (51%). Figure A.2 in the appendix A shows that, for each level of non-cognitive skills, average predicted income increases faster over the distribution of cognitive skills for college education. On the other hand, for each level of cognitive skills, average predicted income increases at similar paces in both education levels.

These results show that college education brings benefits in terms of income for the majority of individuals, including those with low levels of cognitive skills and non-cognitive skills. These results are similar to those of previous studies (Heckman, Stixrud and Urzúa, 2006; Nybom, 2017) and indicate that there is complementarity between cognitive skills and education in the Brazil when it comes to income. However, differently from the results of Nybom (2017) for Sweden, it seems that in Brazil, higher education would benefit individuals in all levels of skills.

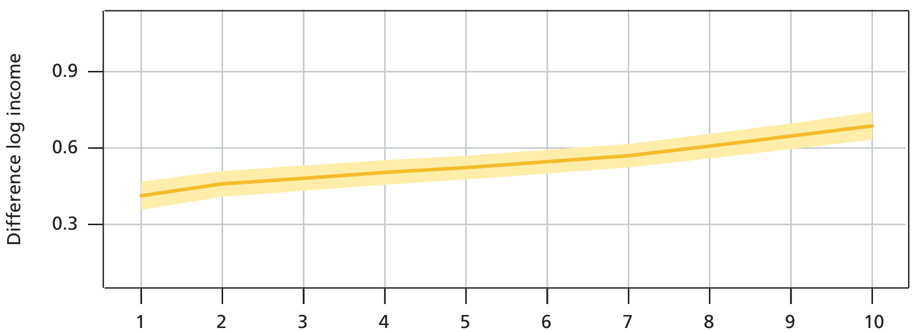
4.3 Three dimensions of non-cognitive skills

In this section, we examine that heterogeneity of higher education impacts on income along dimensions of non-cognitive skills. We use models of three latent factors, in which we always use cognitive skills as one factor and, for each estimation, include two of the three dimensions of the non-cognitive skills our data capture (self-concept, self-regulation, and openness to experience).

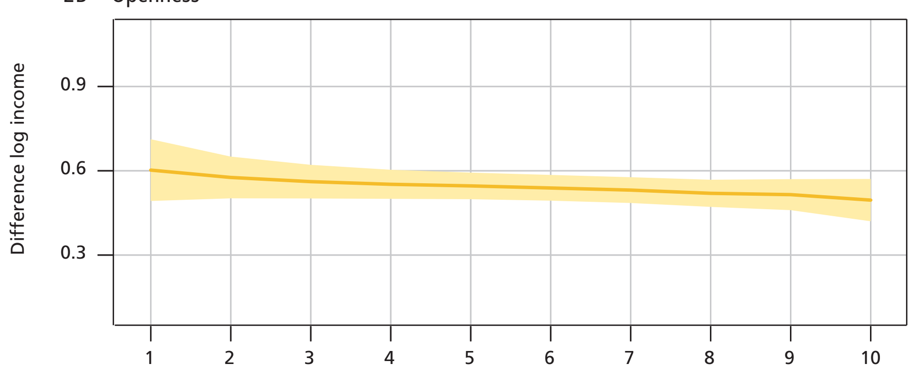
Figures 2, 3, and 4 show that self-concept affect the impact of higher education. Results for self-regulation and openness are not consistent over specifications. In the models in figures 2 and 3, college impacts on income are, respectively, 54% (with a 95% CI of [50%,59%]) and 42% (with a 95% CI of [39%,45%]). In those models, college impact increases from the lowest to the highest deciles of the distribution of self-concept, respectively, about 28 p.p. (from 41% to 69%) and 38 p.p. (from 24% to 62%). These results derive from the fact that self-concept is more important to increase income in the scenario with college education. One possible interpretation for these results is a direct effect, by which occupations that require college education – as managing and leadership positions – also might pay more for aspects of self-concept, as emotional stability, confidence, and self-esteem.

FIGURE 2
Effects of tertiary education on log income by levels of cognitive skills, self-concept, and openness

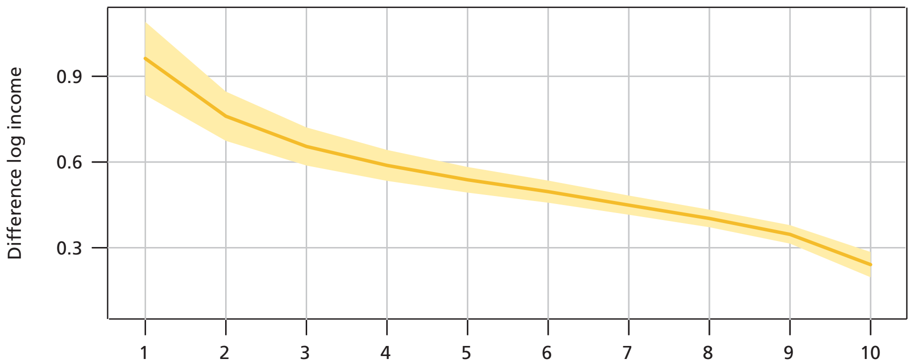
2A – Self-concept



2B – Openness



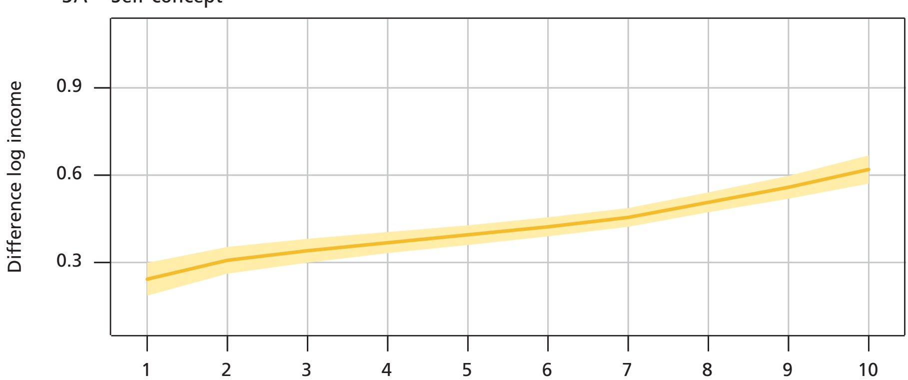
2C – Cognitive skills



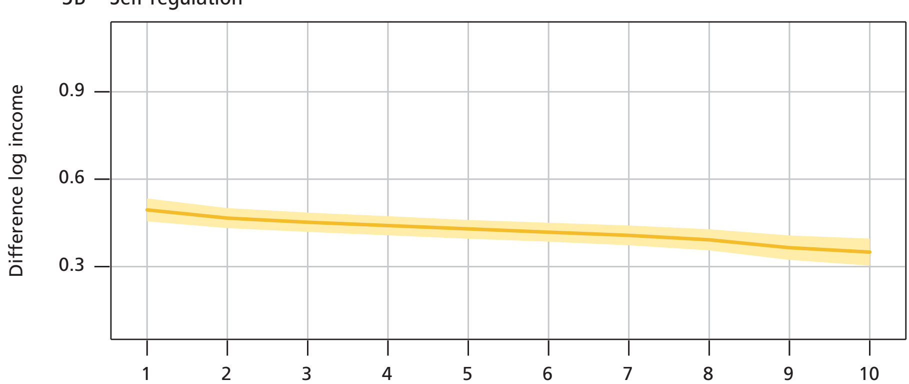
Authors' elaboration.

FIGURE 3
Effects of tertiary education on log income by levels of cognitive skills, self-concept, and self-regulation

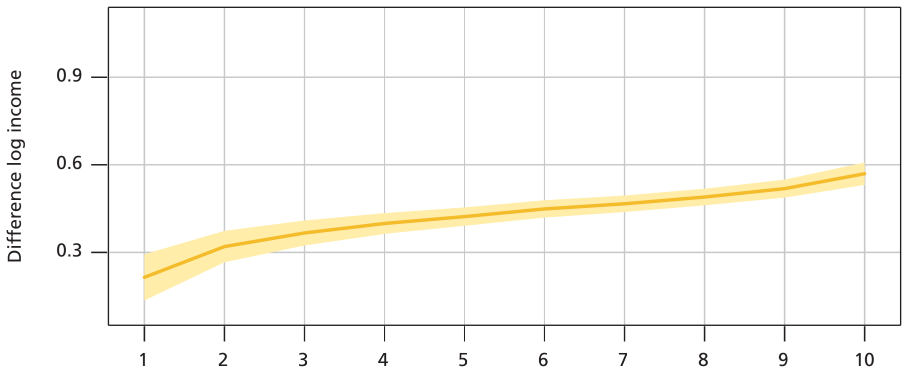
3A – Self-concept



3B – Self-regulation



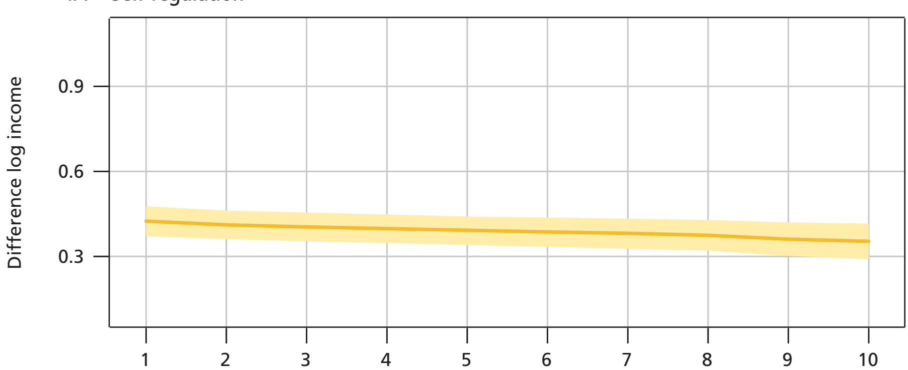
3C – Cognitive skills



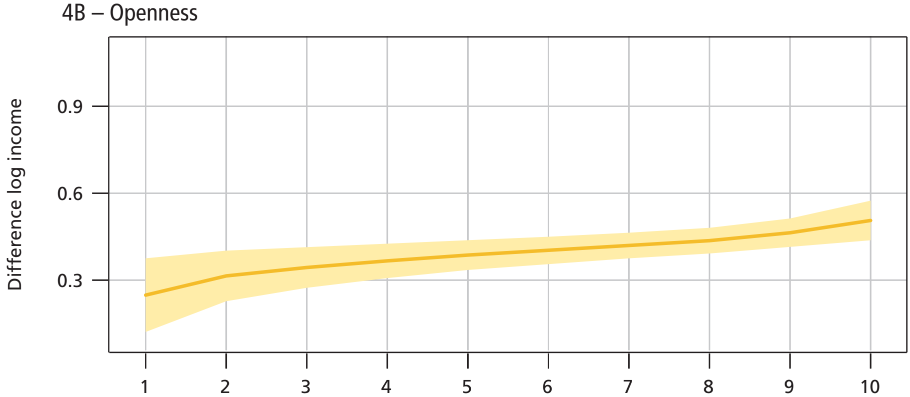
Authors' elaboration.

FIGURE 4
Effects of tertiary education on log income by levels of cognitive skills, self-regulation, and openness

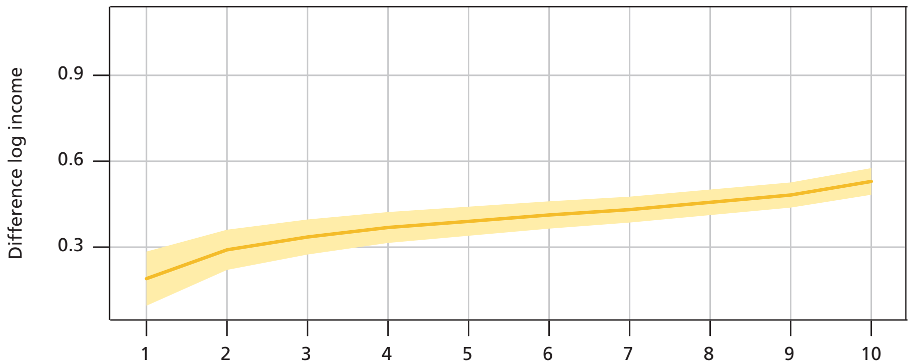
4A – Self-regulation



4B – Openness



4C – Cognitive skills



Authors' elaboration.

5 CONCLUSION

This paper estimates the heterogeneous income returns to higher education in Brazil, using a structural factor model that accounts for unobserved heterogeneity of cognitive and non-cognitive latent skills. We use Brazilian cross-sectional data that includes scores for literacy and numeracy, and measures for three dimensions of non-cognitive skills: self-concept, self-regulation, and openness to experience. We deal directly with endogeneity of schooling decision by modeling it as a function of observable characteristics and latent abilities. Based on the model estimates, we simulate the counterfactual scenarios with and without tertiary education to estimate the ATE of college education on individual income.

Our main results indicate that college education impact in income is 43%. To address the question of whether college education brings benefits to people with lower levels of skills, we also estimate the ATE for deciles of the distribution of cognitive and non-cognitive skills. We find that the ATE varies more over the distribution of cognitive skills than of non-cognitive skills, but higher education college generates benefits even to individuals in the lowest levels of both types of skills. Among the dimensions of non-cognitive skills, ATE varies consistently only over the distribution of self-concept and indicates that managing and leadership occupations, followed by college educated workers, might value this skill relatively more.

We also find weak evidence that the occupations taken up by college educated workers value cognitive skills relatively more, while non-cognitive skills are more relevant to workers with no college education. Finally, we confirm that latent skills highly influence schooling decisions, especially cognitive skills.

In Brazil, in contrast with other Latin-American countries, higher education brings high returns in terms of income, that seem to benefit even those with low levels of skills. However, these results do not directly translate in recommendations to expand the access to higher education even more, as they are relative to the average effect and not the marginal effect. It is possible that those that might enter college with eventual expansions would benefit less. Moreover, our study compares those with college education with all other schooling attainment, so that the marginal gains from high school degree to college education are possibly lower.

REFERENCES

ABAY, K. A.; BLALOCK, G.; BERHANE, G. Locus of control and technology adoption in developing country agriculture: evidence from Ethiopia. **Journal of Economic Behavior & Organization**, v. 143, p. 98-115, 2017.

ACOSTA, P.; MULLER, N.; SARZOSA, M. **Beyond qualifications**: returns to cognitive and socio-emotional skills in Colombia. Washington: World Bank, Sept. 2015. (Policy Research Working Paper, n. 7430).

BANKS, J.; O'DEA, C.; OLDFIELD, Z. Cognitive function, numeracy and retirement saving trajectories. **The Economic Journal**, v. 120, n. 548, 381-410, Nov. 2011.

BINELLI, C.; MENEZES-FILHO, N. A. Why Brazil fell behind in college education? **Economics of Education Review**, v. 72, p. 80-106, 2019.

BUCHMUELLER, G. **Non-cognitive skills and the returns to education**: do non-cognitive skills explain part of the returns to education? Bonn: IZA, Jan. 2019.

CARD, D. Estimating the return to schooling: progress on some persistent econometric problems. **Econometrica**, v. 69, n. 5, p. 1127-1160, Sept. 2001.

CARNEIRO, P.; HANSEN, K. T.; HECKMAN, J. J. **Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college**. Cambridge, United States: NBER, 2003. (Working Paper Series, n. 9546).

CARNEIRO, P.; HECKMAN, J. J.; VYTLACIL, E. J. Estimating marginal returns to education. **American Economic Review**, v. 101, n. 6, p. 2754-2781, Oct. 2011.

CARRANZA, J. E.; FERREYRA, M. M. Increasing higher education access: supply, sorting, and outcomes in Colombia. **Journal of Human Capital**, v. 13, n. 1, p. 95-136, 2019.

CUNHA, F.; HECKMAN, J. The technology of skill formation. **American Economic Review**, v. 97, n. 2, p. 31-47, May 2007.

CUNNINGHAM, W.; TORRADO, M. P.; SARZOSA, M. **Cognitive and non-cognitive skills for the Peruvian labor market**: addressing measurement error through latent skills estimations. Washington: World Bank, Feb. 2016. (Policy Research Working Paper, n. 7550).

EDIN, P.-A. et al. The rising return to noncognitive skill. **American Economic Journal: Applied Economics**, v. 14, n. 2, p. 78-100, Apr. 2022.

GONG, X.; ZHU, R. Cognitive abilities, non-cognitive skills, and gambling behaviors. **Journal of Economic Behavior & Organization**, v. 165, p. 51-69, Sept. 2019.

GONZÁLEZ-VELOSA, C. et al. **Returns to higher education in Chile and Colombia**. [s.l.]: IDB, Mar. 2015. (Working Paper Series, n. IDB-WP-587).

HECKMAN, J. J. et al. **Education, health and wages**. Cambridge, United States: NBER, Mar. 2014. (Working Paper, n. 19971).

HECKMAN, J. J.; HUMPHRIES, J. E.; VERAMENDI, G. Returns to education: the causal effects of education on earnings, health, and smoking. **Journal of Political Economy**, v. 126, p. 197-246, Oct. 2018.

HECKMAN, J. J.; STIXRUD, J.; URZÚA, S. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. **Journal of Labor Economics**, v. 24, n. 3, p. 411-482, 2006. Retrieved from: https://doi.org/10.1086/504455.

JUDGE, T. A. et al. The core self-evaluations scale: development of a measure. **Personnel Psychology**, v. 56, n. 2, 303-331, 2003.

KEJRIWAL, M.; LI, X.; TOTTY, E. The efficacy of ability proxies for estimating the returns to schooling: a factor model-based evaluation. **Journal of Applied Economics**, v. 39, n. 1, p. 3-21, Jan. 2024.

KRISHNAKUMAR, J.; NOGALES, R. Education, skills and a good job: a multidimensional econometric analysis. **World Development**, v. 128, p. 1-56, 2020.

LIMA, A. L. D. Estudo especial sobre analfabetismo e competências socioemocionais na população adulta brasileira. In: INSTITUTO PAULO MONTENEGRO; INSTITUTO AYRTON SENNA (coord.). **Estudo especial sobre alfabetismo e competências socioemocionais na população adulta brasileira**. São Paulo: eduLab21, Dec. 2016. p. 8-29.

LUIK, M.-A.; STEINHARDT, M. F. Immigrant-native differences in stockholding: the role of cognitive and non-cognitive skills. **Journal of Empirical Finance**, v. 38, p. 103-119, 2016.

MONTOYA, A. M.; NOTON, C.; SOLIS, A. **Returns to higher education**: vocational education vs college. Santiago: CEA/University of Chile, 2017. (Documentos de Trabajo, n. 334).

NYBOM, M. The distribution of lifetime earnings returns to college. **Journal of Labor Economics**, v. 35, n. 4, p. 903-952, 2017.

OECD – ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Education at a glance 2022**: OECD indicators. Paris: OECD Publishing, Oct. 2022. Retrieved from: https://doi.org/10.1787/3197152b-en.

OECD – ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Education at a glance 2023**: OECD indicators. Paris: OECD Publishing, Sept. 2023. Retrieved from: https://doi.org/10.1787/e13bef63-en.

OREOPOULOS, P.; PETRONIJEVIC, U. **Making college worth it**: a review of research on the returns to higher education. Cambridge, United States: NBER, May 2013. (Working Paper, n. 19053).

PRIMI, R. et al. Development of an inventory assessing social and emotional skills in Brazilian youth. **European Journal of Psychological Assessment**, v. 32, n. 1, p. 5-16, 2016.

PSACHAROPOULOS, G.; PATRINOS, H. A. Returns to investment in education: a decennial review of the global literature. **Education Economics**, v. 26, n. 5, p. 445-458, 2018.

RODRÍGUEZ, J.; URZÚA, S.; REYES, L. Heterogeneous economic returns to postsecondary degrees: evidence from Chile. **Journal of Human Resources**, v. 51, n. 2 p. 416-460, 2016.

ROY, A. D. Some thoughts on the distribution of earnings. **Oxford Economic Papers**, v. 3, n. 2, p. 135-146, June 1951.

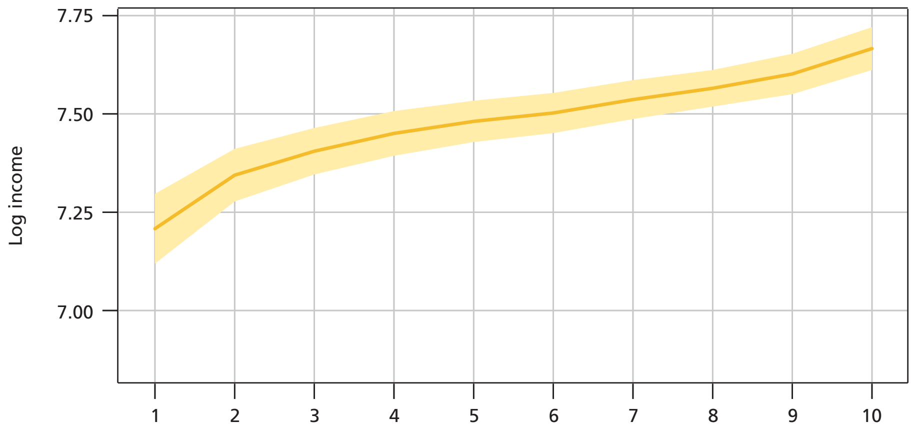
SARZOSA, M.; URZÚA, S. Implementing factor models for unobserved heterogeneity in Stata. **The Stata Journal**, v. 16, n. 1, p. 197-228, Mar. 2016.

STRÖMBÄCK, C. et al. Does self-control predict financial behavior and financial well-being? **Journal of Behavioral and Experimental Finance**, v. 14, p. 30-38, June 2017.

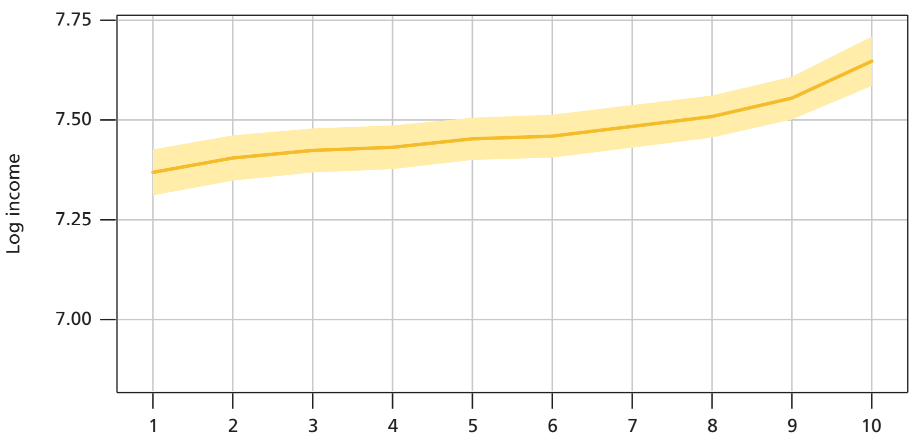
URZÚA, S. Racial labor market gaps the role of abilities and schooling choices. **Journal of Human Resources**, v. 43, n. 4, p. 919-971, 2008.

APPENDIX A

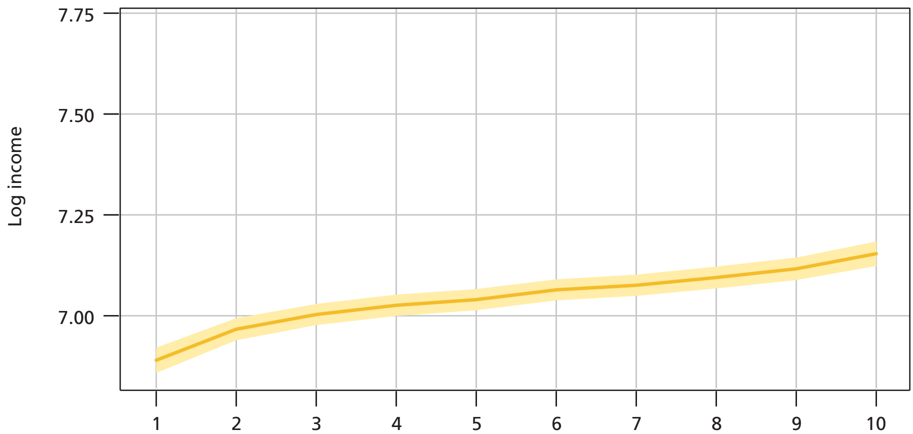
FIGURE A.1
Predicted log income by levels of cognitive and non-cognitive skills
A.1A – Tertiary education: cognitive skills



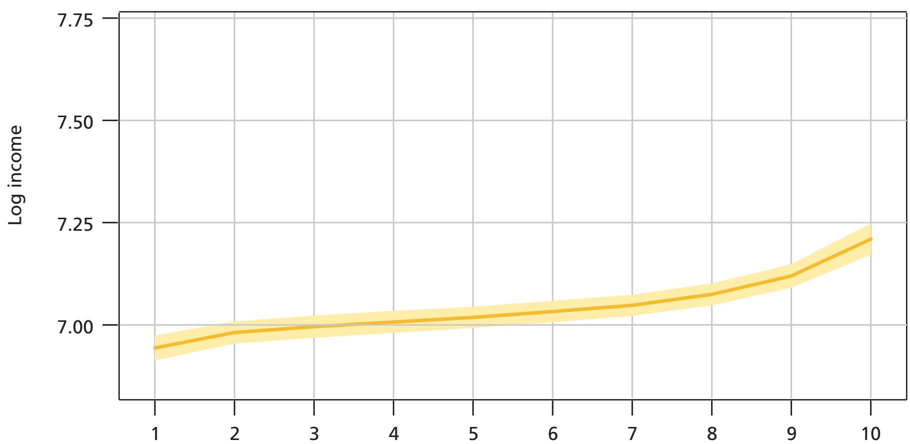
A.1B – Tertiary education: non-cognitive skills



A.1C – Other: cognitive skills

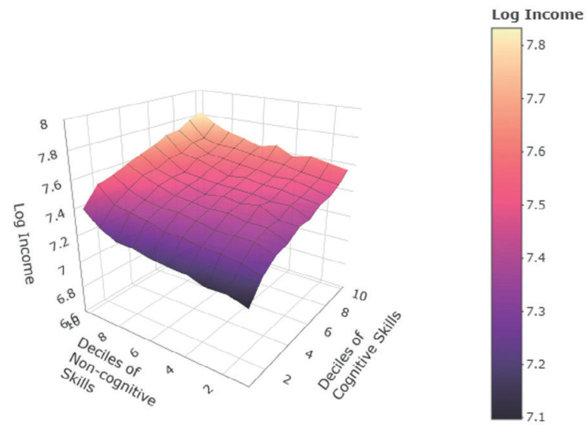


A.1D – Other: non-cognitive skills

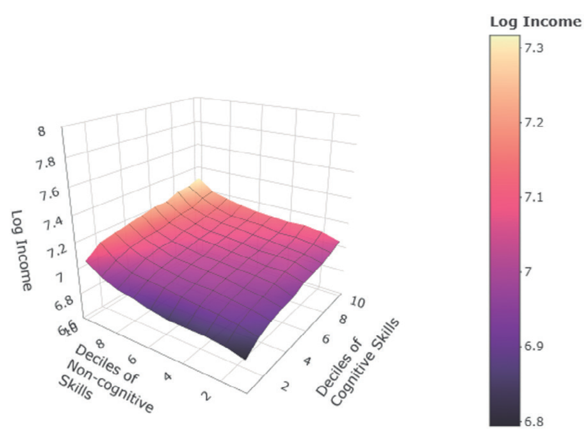


Authors' elaboration.

FIGURE A.2
Predicted log income by levels of the joint distribution of cognitive and non-cognitive skills
A.2A – Tertiary education



A.2B – Other



Authors' elaboration.
Obs.: Figure whose layout and texts could not be formatted due to the technical characteristics of the original files (Publisher's note).

Table A.1 shows estimates of the structural model using only observations whose income was not originally missing. We observe that estimates for skills direct impacts on income have similar magnitudes as those in table 3, but estimate for non-cognitive skills in the scenario with higher education loses significance. However, the impact of non-cognitive skills on higher education choice reduces, which indicates that income missing values are non-randomly distributed over skills' levels.

TABLE A.1
Structural estimations: income and latent skills: no income imputation

Independent variables	Dependent variable		
	Terciary education	Log income	
		Tertiary education	Other
	(1)	(2)	(3)
Cognitive	0.812*** (0.113)	0.176* (0.100)	0.127*** (0.033)
Non-cognitive	0.068 (0.147)	0.160 (0.118)	0.110* (0.058)
Age when started school	-0.018 (0.024)	-	-
40-59 years old	-0.073 (0.138)	0.419*** (0.118)	0.056 (0.052)
60 years old or more	-0.347 (0.295)	0.744*** (0.272)	-0.095 (0.087)
Black or indigenous	-0.284** (0.145)	-0.072 (0.126)	-0.080 (0.054)
Female	0.286** (0.128)	-0.259** (0.105)	-0.461*** (0.049)
Mother completed middle school	0.430** (0.195)	-0.131 (0.162)	0.182** (0.086)
Mother completed high school	0.671*** (0.184)	0.117 (0.137)	0.077 (0.091)
Father completed middle school	0.905*** (0.196)	0.174 (0.156)	0.225** (0.095)
Father completed high school	0.924*** (0.196)	0.045 (0.151)	0.139 (0.100)
Constant	-1.645*** (0.353)	7.296*** (0.279)	6.989*** (0.112)
Observations	740	740	740

Authors' elaboration.
Note: 1. Structural estimations of log income on observable characteristics and latent factors. Regressions further controlled for regional fixed effects. Sample of 22 to 65 years-old individuals. Literacy, numeracy, and computer ability are scores for cognitive ability, self-regulation, openness, and self-concept are scores for non-cognitive abilities. Column 1 shows results for the schooling choice (equation 1) and columns 2 and 3 show estimates for, respectively, individuals that concluded tertiary education and the others.
2. Standard errors in parenthesis.
3. Significance: *** $p < 0.01$; ** $p < 0.05$; and* $p < 0.10$.

TABLE A.2
Conditional means by levels of schooling

	Data	Model
$E(Y\_0 \| D = 0)$	7.007 (0.024)	7.043 (0.001)
$E(Y\_1 \| D = 1)$	7.548 (0.047)	7.474 (0.002)

Authors' elaboration.

Originais submetidos em: ago. 2019.
Última versão recebida em: fev. 2025.
Aprovada em: fev. 2025.