Universities, Socioeconomic Standards and Inclusion Policies: Assessing the Effects on the Performance of Brazilian Undergraduates[1]

Ana Maria Morais°
Admir Antonio Betarelli Junior[2♦♦]
Silvinha Pinto Vasconcelos♦♦
Marcel de Toledo Vieira♦♦

Abstract

The higher education system has been a recurring concern of policymakers around the globe. Many researches have analyzed the additional effects on students' academic performance from an economic and social perspective, which includes the provision and quality of universities. Nevertheless, the joint association among these dimensions were briefly explored for higher education in Brazil, especially in a period of expansion of public university operations with inclusion policies of lower social students. This paper is aimed at contributing to this gap by examining the effect of university settings, socioeconomic features and student financing and quota policies on the ability of Brazilian undergraduates between 2013 and 2017. To accomplish this task, our analysis was conducted by Structural Equation Modelling (SEM) with three confirmatory factors. The main findings revealed that the quality and size of university operations stand out increasingly to improve the academic performance of students, especially for those who attend a public university.

Keywords: Educational economics; Undergraduate education; Universities; Socioeconomic patterns; Structural Equation Model.

JEL Code: C30, I26; I28.

1 Introduction

The debate on the academic performance of higher education students has gained importance in the recent economic discussion, especially due to the expressive growth in the number of university vacancies in several countries since the 1990s (Carneiro & Lee, 2011; Oppedisano, 2014). This movement of expansion of higher education occurred in the midst of transformations in the labor

[1] This paper was supported by CNPq and FAPEMIG under the grant "research project", and CAPES under the grant "PhD Scholarship".
[2♦♦] Professors of the Economics Department, Federal University of Juiz de Fora (UFJF); ° PhD in Economics from the UFJF.

market and technological changes promoted by the advanced internationalization of productive processes and markets, which began to demand more and more workers with higher educational level (Athreye & Cantwell, 2007; Flannery & O'Donoghue, 2013; Heinesen, 2018; Lam & Shiu, 2010; Mallidis, Dekker, & Vlachos, 2012; Oppedisano, 2011). In this context, the literature postulates that the students training at universities can improve their professional position, as well as align such training with the requirements of productivity gains and human capital in the long run of an economy (Becker, 1962; Grossman & Helpman, 1991; Jones, 1995; Romer, 1990). In addition to the guidelines given by the markets, tax incentives, subsidies, public expenditures, credit lines and other forms of support have also been instrumental in fostering the supply and demand for university vacancies, as well as in expanding the educational inputs available.

However, a challenge that can be highlighted when it comes to higher education growth is related to combining the instruments of expansionist policies in universities and simultaneously improving, or not worsening, the academic performance of students. Some authors argue that policies to expand higher education create diminishing marginal effects on the students' level of learning (Harris, 2007; Heyneman & Loxley, 1983). For example, regarding social inclusion policies, in which the movement of expanding the supply of education is accompanied by a greater number of enrollments among individuals from less socially and economically favorable families, there is a tendency to reduce the average level of academic performance of students at universities (Méndez, 2019; Oppedisano, 2011, 2014). On the other hand, some authors argue that educational investments can provide better academic performance results (Albernaz, Ferreira, & Franco, 2002; M. E. Ferrão & Fernandes, 2003; Suryadarma, Suryahadi, Sumarto, & Rogers, 2006).

In line with this international trend of higher education expansion, the Brazilian government started to implement institutional and financial mechanisms to encourage the supply and demand for higher education in the country, as well as make it feasible, especially after the 2000s. The common goals of these mechanisms were to internalize the positive effects created by higher levels of education for the economy, as well as to expand the supply of more qualified and expert workers according to the demands of increasingly fragmented and globally integrated production processes[3] (Zoghbi, Rocha, & Mattos, 2013). For example, we saw modernization policies and public tenders to expand the operations and qualities of HEI (Higher Education Institutions), the construction of new public universities, as well as student financing programs, such as "Ensino Universitário para Todos" (ProUni), which awarded scholarships to low-income students. All these measures led to a strong growth in the number of enrollments in higher education courses in public and private universities in the country (Dearden, Fitzsimons, & Wyness, 2014; Flannery & O'Donoghue, 2013;

[3] That originated the concept of global value chains, a productive system organized in sequential stages (Los, Timmer, & Vries, 2015)

Long, 2019; Pinto, 2004; Rojas, Sánchez, & Villena, 2016; Senhoras, Takeuchi, & Takeuchi, 2006; Zoghbi et al., 2013). According to data from the Ministry of Education and Culture (MEC), between 2000 and 2018 the number of higher education institutions more than doubled (there was an increase of 218%, in which the number of public institutions increased by 169%, and the number of private institutions increased by 222%). This growth resulted in an increase of approximately 313% in the number of enrollments (an increase of approximately 234% in enrollments in public institutions and 352% in private institutions).

As the results of education are not directly observable given the heterogeneity of both the characteristics of each educational institution and the students' family socioeconomic situation, it is only possible to obtain an accurate view of the performance of educational systems through external evaluation on a large scale (Castro, 2009b). As a result, developed and developing countries have been applying different tests and monitoring techniques to university students. The main goal of such tests is to identify the role of contributory factors to the academic performance, sometimes after the implementation of public policies. The conclusive results proliferated among several countries and are still widely used in applied studies on educational quality and equity (Blau & Kahn, 2005; Long, 2019). For example, Zoghbi et al. (2013) sought to estimate the efficiency of HEIs in Brazil using scores obtained by students in the National Examination of Undergraduate Students (Enade). In addition, these standardized tests are also often used as indicators of students' cognitive ability (eg. Carini, Kuh e Klein, 2006; Desjardins, Ahlburg e Mccall, 2006).

In Brazil, although the implementation of a set of policies made higher education more accessible in the second half of the 2000s, research on the change in students' academic performance due to changes in the quality of universities in the period is still scarce. More precisely, we have found no empirical studies that assess the role of quality and supply transformations in public and private universities on students' cognitive skills, which, by definition, are not observable. As far as we are concerned, the literature focuses only on the analysis of students' proficiency scores, whose variable is observed (Gideon, 2017; Oliveira & Araujo, 2005).

Given these considerations, the present study is different because it develops a latent variable that, in a broader way, captures students' cognitive skills, instead of taking only grades on standardized tests as indicators of skill. Nevertheless, the concept of academic performance can be determined by several factors and depends both on the socioeconomic characteristics of each student as well as on the quality and supply of the Brazilian colleges where each one studies. Some Brazilian studies have already highlighted the importance of social, economic and demographic characteristics of basic education students in their own performance (Albernaz et al., 2002; Araújo & Siqueira, 2010; César & Soares, 2001; Felício & Fernandes, 2005; M. Ferrão et al., 2001; Machado, Moro, Martins, & Rios, 2008; Menezes-Filho, 2007; Rodrigues, 2009; S. Soares &

Sátyro, 2008). However, perhaps because higher education is not mandatory[4], this type of analysis did include academic performance in public and private colleges of the country.

In a way that has never been published before, this article seeks to fill this gap and is aimed at analyzing the influences of both recent public policies on HEIs and socioeconomic variables on students' unobserved ability. More specifically, our assessment considers the effects of the student financing policy and the quota policy–the instruments that made it possible for poor students to enter various colleges across the country. The question that arises then is whether students enrolled and benefiting from these policies caused a drop in average skills upon entering higher education institutions. For this, we apply a structural equation model (SEM), whose statistical approach allows us to extract latent variables from theoretical hypotheses and concepts, and to associate them with each other in an inferential analysis, whether formed by one or several simultaneous equations.

There are two underlying hypotheses to be tested in our research. Firstly, we assess whether the quality of higher education institutions contributes positively to students' cognitive ability in the Brazilian case. Secondly, we analyze whether individual socioeconomic characteristics are capable of influencing or are resilient to students' ability for the years 2012 and 2017. This study is justified because it can contribute to the ongoing debate on social and income transfer policies in the country, as well as point out whether the growth in the number of students in higher education had as a counterpart the improvement in academic performance. More broadly, the estimated results may also provide some direction or answers for the development or conduction of educational policies in the country. This is particularly relevant in a recent context of greater financial fragility in the public budget and austerity in terms of financial resources for higher education in the country.

Thus, the article is divided into five sections, including this introduction. The second section presents a brief theoretical discussion about the hypotheses to be tested. The third section presents the methodology and observed variables that form the theoretical structure of this article. The fourth section brings the discussions about the estimated results of a theoretical structure elaborated from variables observed in a SEM model. Finally, the fifth section presents the final considerations of the research, highlighting the main conclusive results.

2 Policies and Forms of Evaluation of Academic Performance in Brazil

In Brazil, higher education courses are offered by educational institutions accredited by the Ministry of Education (MEC), which can be colleges, universities, university centers[5], integrated colleges, or institutes. According to data from the National Institute of Educational Studies and

[4] In Brazil, basic education is mandatory for all children, which is not the case for higher education.
[5] Colleges are focused on specific areas of knowledge; universities, in addition to the roles of the college, must have professors with masters and doctor's degree; and university centers are in transition from college to university.

Research Anísio Teixeira (INEP), the majority of vacancies in higher education courses are offered by private institutions, which totaled 1929 colleges in 2018, while the sum of public institutions of any kind represented only 299 of the total HEI.

The Brazilian Federal Constitution determines the guarantee of a quality standard for higher education, defined as minimum quantities of inputs indispensable for the development of the teaching-learning process. In this sense, since the 1990s, educational policies have focused on investments that provide not only a larger number of vacancies, but also contribute to the quality of teaching, such as through the presence of more libraries in HEIs, better teacher training, and adaptation of facilities (Castro, 2009a; Felício & Fernandes, 2005; J. Johnes, 2006; Oliveira & Araujo, 2005; José Francisco Soares, 2007). According to data from the Higher Education Census conducted by INEP, in 2000, only 32.11% of the professors from public institutions had a doctorate, and for private institutions this number was even lower, 11.32%. In 2018, this percentage increased to 63% of professors with doctorates at public universities, and 25.79% at private universities. The interest in improving teaching facilities can be seen in the increase of expenses with maintenance and funding of HEIs, which was of about 9% between 2012 and 2017.

In 2003, a special committee was created to evaluate higher education in order to suggest changes to the evaluation system and, in 2004, the National Examination of Undergraduate Students' Performance[6] (Enade), coordinated by INEP, linked to the Ministry of Education, started to be applied. With the consolidation of these proficiency exams in Brazil since the mid-1990s, studies on the factors relevant to the performance of Brazilian students became feasible (Andrade & Laros, 2007; Castro, 2009b; Zoghbi et al., 2013). These school performance tests, such as Enade, evaluate education systems and help to identify factors that hinder learning through an exam carried out with graduates in undergraduate courses. But a problem to be considered is the lack of incentives that students receive to commit to taking the tests responsibly.

More precisely, the Enade test analyzes students' knowledge in relation to the syllabus of the curriculum of their courses and students' knowledge about the Brazilian and global reality today. However, in addition to factors involving the institutions, there are also external situations that may be associated with educational opportunities and that are capable of influencing student performance. This is the case, for example, of variables of family socioeconomic situation and school trajectory previously experienced by the individual, which tend to present a greater dispersion in developing countries (Alves & Soares, 2009; Felício & Fernandes, 2005; Luz, 2006).

This influence is corroborated by research on education, whose theoretical basis postulates that the performance obtained by students in tests of academic knowledge can be influenced both by the

[6] The test consists of 40 questions, 10 related to general knowledge and 30 related to the specific knowledge of each course.

socioeconomic situation of individuals (Becker, 1972; Coleman, 1968; Rodrigues, 2009), and by the quality of the study environment in which they are inserted (Brooke & Soares, 2008; Hanushek, 1986).

Although, on the one hand, there is a consensus over the ability of students' socioeconomic conditions to influence their academic results, on the other hand, there is a divergence in relation to the importance of the characteristics of the HEIs. This occurs because not all variables that influence educational performance are under the control of HEIs (Costa, Ramos, de Sousa, Sampaio, & Barbosa, 2015; J. Johnes, 2006). However, due to the scarcity of applied works that relate the quality of teaching and the socioeconomic conditions of students to the academic performance of students in higher education, we also base our discussion on basic school analysis to understand what happens in terms of HEIs. In this regard, some studies have already been carried out in an attempt to explain the relationship between the quality of schools, the socioeconomic conditions of students, and school performance (Brunello & Checchi, 2005; Ermisch & Francesconi, 2001; Oppedisano, 2011; Rios-Neto, César, & Riani, 2002). However, there is no consensus on the effect of these variables on student performance. Some applied studies indicate that educational institutions have only a limited effect on learning and cannot compensate for the influence of family background and socioeconomic differences (Baqueiro, 2015; Felício & Fernandes, 2005; Rodrigues, 2009; Jose Francisco Soares & Andrade, 2006). For example, Rodrigues (2009) sought to identify and measure the main factors linked to the maintenance of low levels of school performance in Brazil. He found out that the alteration of the students' profile due to the increase in the relative participation of students with less favorable resources, in addition to having contributed directly to reducing the average global performance, contributed indirectly to the reduction of the learning of other students who could have better results, through an overflow effect.

Albernaz, Ferreira and Franco (2002), Felício and Fernandes (2005) and Menezes-Filho (2007) used data from the National Basic Education Assessment System (Saeb), implemented by the Ministry of Education (MEC), to assess the characteristics of students, teachers and schools that favor better school performance. Albernaz, Ferreira and Franco (2002) estimated an educational production function for Brazil, applying linear hierarchical models for the year 1999, to investigate the contribution of different school and teacher variables. However, they found that most of the performance variance was a consequence of differences between students' family characteristics, reflecting an important clientele selection effect (Albernaz et al., 2002; Felício & Fernandes, 2005; Menezes-Filho, 2007).

The results found by Albernaz, Ferreira and Franco (2002) were grouped into two categories. The first refers to the individual and family characteristics of the students, who played an important

role in determining school performance. The second refers to school and teacher variables, which also had significant effects on performance for most variables (Albernaz et al., 2002). Menezes-Filho (2007) also concluded that the variables that most explain school performance are the characteristics of the family and the student (Menezes-Filho, 2007). Flanerry and O'Donoghue (2013) explain that a lower family income and lower parental education, in addition to negatively influencing the participation of children in higher education, also result in worse educational outcomes (Flannery & O'Donoghue, 2013). However, in general, according to the authors of this category of applied studies, the conclusions drawn do not prevent the levels of learning from being extended by improving the quality of schools (Albernaz et al., 2002; Felício & Fernandes, 2005; Menezes-Filho, 2007).

Johnes and Johnes (2009) argue that HEIs do not represent a homogeneous group, and they estimate the costs of higher education in a structure that allows institutions to differ in terms of efficiency and cost technology (G. Johnes & Johnes, 2009). Machado et al. (2008) and Sátyro and Soares (2008) sought to identify which school inputs affect performance. Machado et al. (2008) analyzed the determinants of performance in mathematics by students from the state of Minas Gerais using hierarchical models, recognized for separating the effects of various levels. The authors found evidence that the effects of school have little influence compared to student characteristics and family background (Machado et al., 2008; S. Soares & Sátyro, 2008).

However, some authors point out that the effect of school quality is more significant in developing countries, as is the case of Brazil, than in developed countries. Results related to the effectiveness of school inputs may vary according to the outcome variables used, to the social environment in which the school is located, and to the stage of development in which the analyzed country is located, in addition to the temporal variation. In less developed countries, school inputs tend to be significant because of the great variability in the resources available in schools (Wail, Said, & Abdelhak, 2011; Willms & Somers, 2001).

Although many studies point to family factors as the main determinants of academic performance, there are studies that indicate that school factors are also important. So, students from the same socioeconomic context may present different performance because they study in different environments. Several studies have concluded that there is sufficient evidence to consider that the school makes a difference in educational results, and that the explanatory power of school factors is high enough to cause a change in the students' school trajectory (César & Soares, 2001; M. E. Ferrão & Fernandes, 2003; Riani & Rios-Neto, 2008; Rios-Neto et al., 2002; Suryadarma et al., 2006). Ferrão et. al (2001) also found a positive and significant effect of school infrastructure, safety and cleanliness variables on the performance of 4th grade students in 1999 (M. Ferrão et al., 2001). Araújo and Siqueira (2010) sought to highlight the personal and school attributes that are

decisive in the students' performance, based on math proficiency tests for the 4th grade. They found out that both the school context and the school infrastructure help to explain the students' performance (Araújo & Siqueira, 2010). Fernandes and Menezes-Filho (2020) pointed out that schools that receive public funding are able to improve the mechanisms for hiring teachers, making the system more equitable, which helps to improve student performance (Fernandes & Menezes-Filho, 2020). Table 1 summarizes the main applied researches divided by similar approaches for Brazil.

Table 1 - Applied studies to the Brazilian economy

References	Indicators and main highlights	Models	Main results
César and Soares (2001)	Student: gender, race, job, income. School: education network, average and standard deviation of socioeconomic level.	Hierarchical model	The segregation of poor students in schools causes underperformances.
Albernaz, Ferreira and Franco (2002)	Student: income, parent education, goods and services, repetition, race. School: administrative, resources, classrooms, education and teachers' salary, socioeconomic level.	Hierarchical model	Family socioeconomic status affects performance. Better schools raise results.
Rios-Neto, César and Riani (2002)	Student: education, socioeconomic conditions and parent occupation. School: teachers' salary and education, and teacher / student ratio.	Hierarchical model	School inputs compensate for precarious socioeconomic conditions.
Felício e Fernandes (2005)	Student: gender, race, parental education, type of family, number of people in the family and income.	Fixed Effects	Investment in education increases performance.
Soares e Andrade (2006)	Social position: characteristics of students' homes, parents' education, parents' employment, family income, administrative dependency.	Response Theory Methods	Public and private schools make a difference in performance. School quality is associated with inequity.
Menezes-Filho (2007)	Student: mother's education, race, school delay, number of books, computer at home, age. School: computers, principals and students, education, age and salary of teachers, class hours, State.	Econometric models	Family and student characteristics affect performance.
Machado et al. (2008)	Student: proficiency, race, books, newspapers, homework, parental employment, academic shift. School: spending per student, student financing, region, teacher education, garbage collection, municipal GDP.	Hierarchical model	The effect of the school and the municipalities on student achievement is smaller than their characteristics.
Riani e Rios-Neto (2008)	Student: age, gender, race, home situation, parental education, parental occupation. School: class hours; students per class; schooling of teachers, courts, libraries, laboratories.	Hierarchical model	Some factors in the school network can make the system more egalitarian.
Sátyro e Soares (2008)	Schooling of teachers, students per class, hours of instruction, infrastructure, library, location, municipality income, urbanization, population.	Fixed Effects	School supplies increase educational results.
Rodrigues (2009)	Student: proficiency, gender, motivation, lag, race, family and socioeconomic status. School: proficiency; proportion of students with certain characteristics, infrastructure, schooling of teachers and principals, region, administrative.	Non-parametric model	Changing student profiles reduces performance.

In summary, as mentioned earlier, the discussion on the performance of students in primary and higher education in the country has gained momentum over the past 20 years, concurrently with the

growth of the educational system in the country. Some empirical results from different approaches point to a growing role of both socioeconomic characteristics (internal) and educational institutions (external) in students' abilities, although in different ways and to varying degrees. There is an expectation that the combination of higher education growth (HEI) and social policies aimed at the poorest individuals can contribute to a change in academic performance, in addition to admittedly mitigating the distortions generated by socioeconomic inequality in the country. In the following section, we present the empirical strategy used in the analysis of several factors that are capable of influencing academic performance, as postulated in the aforementioned literature.

3 Empirical Strategy

The empirical strategy adopted in this study is capable of simultaneously evaluating, from a theoretical framework, the various theoretical relationships and multiple dependencies that influence students' cognitive skills. The analysis was based on two successive steps. At first, our study used Spearman's (1904) exploratory factor analysis (EFA)[7] to identify the latent factors derived from the theoretical relationships of all observed (random) variables, which naturally emerge from the data structure. In an unrestricted way, this step allows to extract the latent factors that reproduce the most representative common relations of subset of observed variables. Therefore, it is possible to verify if the latent factors that arise naturally from the random vector are consistent with the latent factors idealized for a given restricted set of observed variables. So, we use EFA as an intermediate step to validate our latent constructs, idealized from the hypotheses of the theoretical literature (Rencher & William, 2012).

In the second stage of the empirical strategy, we applied Confirmatory Factor Analysis (CFA) to obtain three latent constructs, validated by the EFA and derived from certain relationships between observed and latent variables, assuming a predetermined pattern. In this statistical technique, the number of latent factors must be previously known and can be used to test a hypothesis arising from a theory. We estimate the relationships between the latent constructs by means of regression coefficients in a system of simultaneous equations, also known as structural equation model (SEMs). Therefore, SEM is developed from the EFA to study the relationships between latent variables. There are, therefore, two models for testing hypotheses arising from theories: a measurement model, represented by the EFA; and a structural model (SEM) (Bartholomew, Steele, Galbraith, & Moustaki, 2008). SEM has the advantage of examining concepts that are not directly observable, in addition to allowing the simultaneous analysis of several multiple dependency relationships with statistical efficiency. Another advantage of SEM is

[7] As the variables are categorical, we used the polychoric correlation matrix.

that it incorporates errors in the estimation of multiple dependency relationships caused by imperfect measurement of latent variables (Hair Jr., Black, Babin, & Anderson, 2014).

The technique is also capable of handling a large number of endogenous and exogenous variables, as well as latent variables specified as linear combinations of observed variables (Golob, 2003; Yang, 2018). However, as this statistical method requires prior knowledge about the variables, a two-step approach was chosen, according to the procedure adopted by Yang (2018). That is, in the first step, a confirmatory factor analysis was performed to validate the measurement model and then the structural model was applied[8]. The estimation for binary variables proceeded from a logit model of Bernoulli function, whereas for ordered categorical variables we use the ordered logit[9].

In the measurement model, three latent variables were considered, which represent the objective of this study. The first represents the individual characteristics of the students, the second corresponds to the characteristics of the institutions and the third reproduces the ability of the students, which will be explained later. The mathematical equations[10] of the measurement model that define the relationship between the observed variables and the exogenous latent variables (Social and Quality) are:

$$x_i = \alpha_{i1} Individual + \delta_i \quad (1)$$

$$x_j = \alpha_{j2} Institution + \delta_j \quad (2),$$

where the vector x_i represents the vector of i observed variables related to the first latent (Individual), x_j corresponds to the vector of j observed variables related to the second latent (Institution), α refers to the factor loads and δ represents the disturbances of the explanatory variables related to each exogenous latent. In turn, the endogenous latent variable (Ability) is defined by:

$$y_k = \alpha_{k3} Ability + \varepsilon_K \quad (3)$$

where y_k represents the vector of k observable variables related to the endogenous latent variable, and the vector ε represents its disturbances. The factorial load $\alpha_{13} = 1$, this constraint being used to fix the scale of the latent variable. This restriction ensures that the latent *Ability* has the same scale measure as the observed variable y_1 (Bartholomew et al., 2008).

In the structural model, we assume that the latent variable Ability is endogenous and determined by other latent (exogenous) factors, that is:

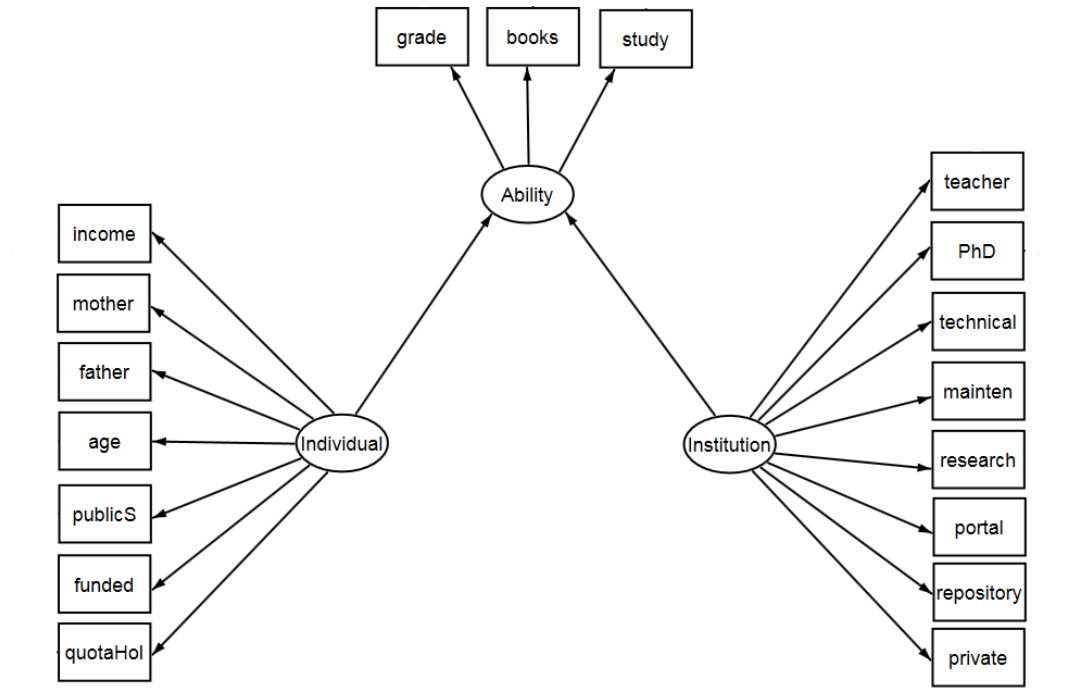
$$Ability = \gamma_1 Individual + \gamma_2 Institution + \vartheta \quad (4)$$

[8] All analyses were performed on Stata15.
[9] The model was also replicated in probit and the results were remarkably similar. We chose to expose in this work the results by logit.
[10] The mathematical representation follows the notation of the book of Bartholomew et al. (2008).

where ϑ is the disturbance of the structural model.

The equations (1) to (4) establish our theoretical model, whose relationships between the observed variables and the latent constructs are illustrated by the trajectory diagram in Figure 1. In this diagram, we can also observe the relationships between the latent constructs, as defined by equation (4). Each variable used in the theoretical model is described in the following section.

Figure 1 - Theoretical model



To scale the latent variables in the CFA so that they present the same scale, we commonly choose a reference variable to have a unit factor load. In this procedure, we choose the observed variable that best represents the latent variable and apply the unit factor load restriction (Bartholomew et al., 2008). The models were estimated with different methods of interaction, and after each model, tests were performed reporting the Akaike (AIC) and Bayesian (BIC)[11] criteria that serve to compare the quality of the models. Smaller values for these tests indicate better adjustment, which was obtained by the Gauss-Hermite interaction method. In addition, the model was also tested with the students' perception variables for teachers, facilities and events, but according to the AIC and BIC tests, the model is better specified without considering these variables. The Wald test was also applied to determine whether more restrictions should be added. This test was applied to each coefficient and each constant, and in all cases, it was possible to reject the hypothesis at a significance level of 5%, so that such values would be equal to each other or equal to zero. Therefore, the model is well specified, without the need to include any further

[11] These tests do not judge the adjustment in absolute values, but they serve to compare different models.

restrictions. The coefficients were significant at a 99% level of significance, always with values within the confidence interval.

3.1 Database

The data used in this study to evaluate our theoretical model with a focus on the ability of higher education students corresponds to two sources of information: National Student Performance Exam (Enade) and the Higher Education Census. The Enade test, which is applied by INEP and is part of the National Higher Education Assessment System (Sinaes), evaluates the academic performance of graduates of undergraduate courses in relation to the syllabus contained in the course curricula. It also applies a "Student Questionnaire", which collects the students' socioeconomic and training process information. The data provided by Enade are basically categorical, given that they are answers to a questionnaire. On the other hand, data from the Higher Education Census, also made available by INEP, present information from educational institutions in the country.

We extracted variables from these two main data sources, with information covering the years 2012 and 2017. We chose 2012 because in that period there was a strong expansion of resources associated with higher education financing programs (Chaves & Amaral, 2016). In addition, the annual databases that precede 2012 are irregular and do not fully address the variables observed in the 2017 database. Thus, if we chose a year prior to 2012, the latent constructs would have less variables and would be more incomplete, especially when compared with some of the indicators of the studies applied in Brazil (Table 1). In turn, 2017 was chosen as the most recent year for Enade data. Thus, by preserving the regularities and correspondence between the bases of 2012 and 2017, it is possible to carry out a comparative analysis of the estimated results and even a critical assessment regarding the impacts of educational policies on society. Only face-to-face courses were analyzed, since online courses have a different spending logic. In addition, we consider only the students present in the Enade test[12], and in both years considered (2012 and 2017), the test was applied only to the students completing their courses. Altogether there are 135,139 students for the year 2012 and 323,482 for the year 2017[13]. It is worth noting that there are several variables with the potential to describe the study questions in this article. But only those that are available in the Enade and *Censo Superior* databases, released by INEP, will be used, whose information is

[12] It was necessary to merge the INEP and Superior Census databases. This was possible due to an identification variable for each HEI.

[13] The analyses were performed both for the database with missing and without missing values, but as the results were similar, it was decided to follow the analysis without considering the missing values.

sufficient to characterize well the higher education in Brazil. The more detailed description of each variable considered is presented in the following section.

3.1.1 Description of Variables

Ability is the main latent construct in the structural model, that is, it is not directly observed, although it can be inferred from other observed variables. The main observed variable used in the literature that can compose students' cognitive ability is the grades obtained in standardized tests of knowledge. For that, we use the grade variable in Enade. However, this work innovates when it also considers other variables observed as possible determinants of the skill: number of hours dedicated to the study and quantity of books that the students read in the year, except those indicated in the bibliography of the courses. This information is categorical and ordered and was informed by the students when they answered a questionnaire with five options (from “none / none” to “more than twelve”).

The observed skill of the students must be related to some socioeconomic characteristics of the students, as well as to some characteristics of the institutions in which they study. The variables that characterize students' family background and socioeconomic conditions accompany the applied literature that deal with family income, parental education and school trajectory of each student (Dearden, Fitzsimons, & Wyness, 2014; Flannery & O’Donoghue, 2013; Heinesen, 2018; Oppedisano, 2011, 2014; Rojas et al., 2016). In general, family income has a positive relationship with the parents' education and the fact that the student studied in private schools during high school. These variables are expected to be quite representative for the students' socioeconomic situation, which should influence academic performance. We include the age variable, because, in general, students from higher-income families have the opportunity to enter at a younger age at higher education courses as, in addition to other possible factors, they tend to need less to enter the labor market directly and are usually academically best prepared. In addition, in order to assess whether certain policies are related to students' ability and socioeconomic characteristics, we consider a dichotomous variable that denotes whether the student entered the course through affirmative action or social inclusion policies (quota), and another that identifies if the student has any scholarship or funding to pay for the course.

Like Heinesen (2018), we related the quality of educational institutions to the proportion of resources per student. As main factors that may represent the quality of the faculties, we chose the teacher/student relationship, the level of education of teachers (represented by ratio of doctoral professors per the total number of professors at the institution), expenditures per student, which includes expenses with maintaining the HEI and with research and development, and aspects of

school infrastructure, following some studies applied in Brazil (César & Soares, 2001; M. Ferrão et al., 2001; Machado et al., 2008; Rios-Neto et al., 2002). In this subset of variables, we add a dichotomous variable for accessing the CAPES portal of journals, another dichotomous variable that identifies the presence of a scientific production repository in the HEI, and a dichotomous variable that distinguishes between public and private HEIs. These indicators help to characterize the learning opportunities that students have within each HEI. There is also information obtained from the Enade questionnaire regarding the assessment made by students on infrastructure and teaching (e.g. teachers in the discipline, the adequacy of the HEI facilities, and the support for events). However, because it is a subjective assessment of students in a questionnaire, this information tends to be biased. What is observed is that most students make a maximum assessment of the items considered and many may not be sincere, careful or, in fact, do not know how to answer these questions. For these reasons, the data from these skewed assessments by the students were disregarded in our theoretical model.

Table 2 provides some descriptive statistics for continuous variables, while Table 3 shows the nominal variables. As shown in Table 2, the average grades in Enade did not change significantly between the years 2012 and 2017. In other words, apparently there were no changes within this period capable of causing big changes in the students' academic performance. However, on average, students who attended higher education in 2012 were younger compared to the students in 2017. The expansion of the supply of higher education can be seen in the comparison between the two years of the variables of teachers per student, the number of doctorate degree teachers, the number of administrative technicians and maintenance and research expenses, all with higher values for the year 2017, mainly the expenditure of HEIs on research and development, which has increased significantly by more than 40%.

In general, Brazilian students study few hours and do not read much (Table 3). In a comparison of students' socioeconomic and family conditions between 2012 and 2017, it is possible to see that in 2017 higher education became more accessible to people with lower incomes and who attended public schools during high school[14]. This trend is also perceived in the variable of parents' education, since in 2017 the proportion of children of poorly educated parents attending higher education courses was higher. The expansion of student financing was quite evident in the comparison between these two years, with an increase of approximately 60% in the number of students who declared to receive some type of financing for their studies. There is also an expansion in affirmative action and social inclusion policies. In 2017, almost 70% of the students analyzed

[14] Students may have moved between public and private schools during high school, but the type of school they attended most was considered.

studied in public schools during high school, about 62% of students received some form of funding for their studies, and 23% of students entered the course through the quota system.

Table 2 - Description of Continuous Variables

Continuous Variable	Description	2012		2017	
		Mean	stand. Error	Mean	stand. Error
grade*	Gross grade in the ENADE test.	46.91	14.50	44.43	13.85
age*	Age.	26.66	6.39	27.20	6.86
teacherº	Teacher to student ratio.	0.05	0.04	0.06	0.04
PhDº	Proportion of teachers with PhD in relation to the total number of teachers.	0.35	0.24	0.39	0.25
technicalº	Ratio of the number of administrative technical per student.	0.07	0.11	0.06	0.07
maintenº	Expenses for maintenance of HEI or maintainer by total students + teachers + technicians per year.	4473.74	52670.93	4875.22	121956.50
research°	Expenses with Research and development of the IES or the maintainer by the total number of teachers per year.	2919.96	8334.95	4167.28	9756.11

Source: *Enade/INEP; º Source: Censo Superior/INEP

As some of the variables used are categorical variables obtained through responses to a questionnaire and due to the complexity of the model used, it was only possible to execute the model after transforming all variables into categorical ones. Therefore, the continuous variables exposed in Table 2 were transformed into variables with 10 categories for the application of the model. Figure 2 illustrates the average and maximum[15] grades according to groups of categories of some characteristics of educational institutions and students, for 2012 and 2017. There is a positive trend, as expected, between more hours of study and grade. A higher ratio of professors to students and professors with a doctorate among teachers are also associated with higher grades. Among the characteristics of students, we can see that the grades are higher among students from families with higher incomes, children of parents with a higher level of education[16] and who studied in private schools during high school.

[15] The minimum score was not represented because it is always zero in all cases.
[16] The organization of this graph considered the maximum level of the parent's educational background.

Table 3 - Description of Categorical Variables

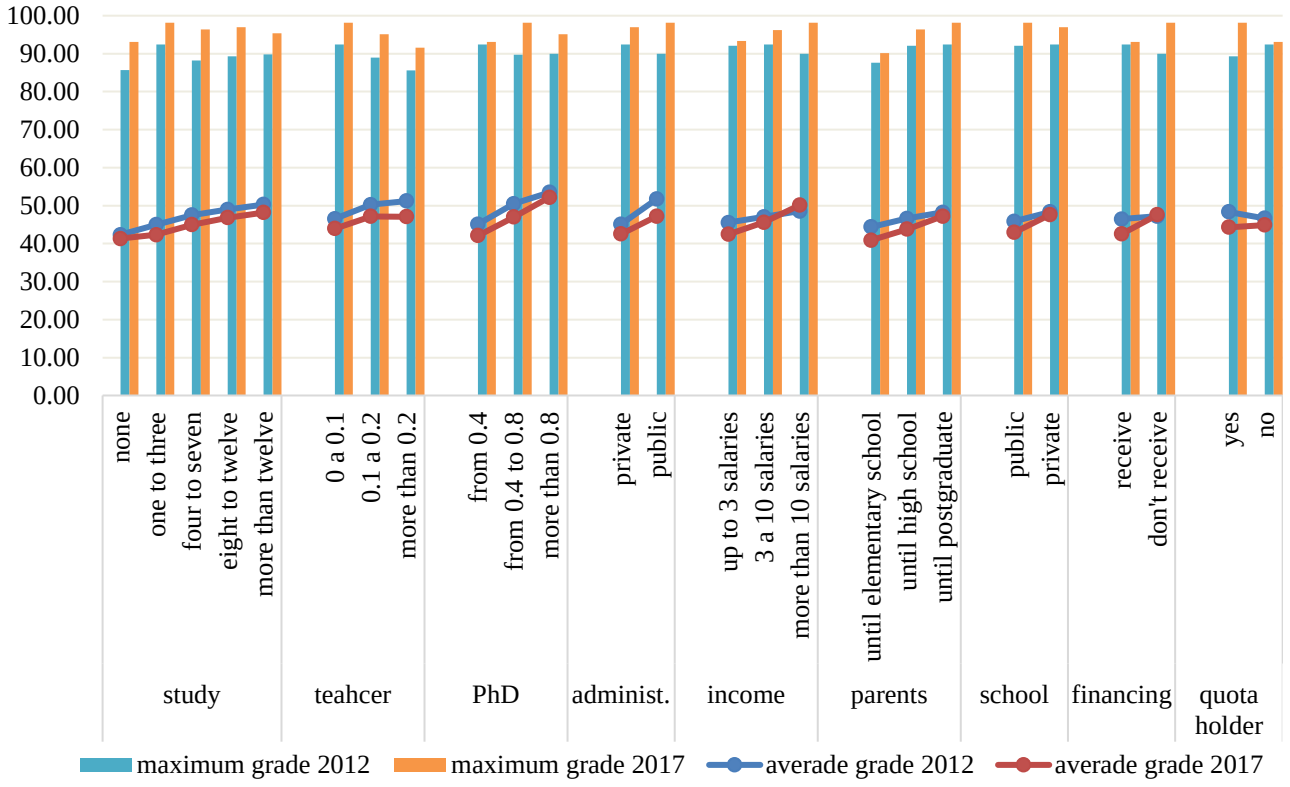
Variables (Description)	Categories	Percentage 2012	Percentage 2017	Variables (Description)	Categories	Percentage 2012	Percentage 2017
Academic performance							
study*		2012	2017	**books***			
Hours per week devoted to studies, excluding class hours.	none	2.65	3.82	Number of books read in the year, except those indicated in the course bibliography.	none	10.26	13.05
	one to three	40.16	39.71		one or two	36.07	37.01
	four to seven	30.38	30.23		three to five	31.33	30.28
	eight to twelve	13.84	13.31		six to eight	10.13	8.96
	more than twelve	12.97	12.93		more than eight	12.20	10.7
Socioeconomic characteristics							
income*				**quotaHol***			
Total household income, in minimum wage (mw)	until 1,5 mw	13.10	21.11	Participation in quota system.	no	84.76	76.45
	from 1,5 to 3 mw	25.18	27.51		yes	15.24	23.55
	from 3 to 4,5 mw	19.10	20.6	**funded***			
	from 4,5 to 6 mw	13.35	11.14	Scholarship or funding to comer all or most of the tuition.	no	60.98	37.25
	from 6 to 10 mw	14.06	11.59		yes	39.02	62.75
	from 10 to 30 mw	12.56	6.99	**publicS***			
	over 30 mw	2.65	1.07	Type of high school.	Public	58.16	68.46
					Private	41.84	31.54
mother*				**father***			
Mother's education.	none	2.98	3.88	Father's education.	none	5.91	6.16
	school: 1st to 5th year	19.19	21.56		school: 1st to 5th year	32.73	26
	school: 6th to 9th year	12.90	14.49		school: 6th to 9th year	19.07	15.01
	high school	33.25	33.73		high school	43.62	32.26
	higher education	20.11	17.03		higher education	24.62	15.13
	Postgraduate	11.57	9.3		Postgraduate	9.19	5.43
Facilities and evaluations of institutions							
private*				**learning***			
HEI is public or private.	public	27.39	40.71	Student assessment of the teacher's mastery of the subject.	very low	0.62	0.7
	private	72.61	59.29		low	1.63	1.44
portal°					reasonable	4.80	4.63
IES library has access to the CAPES journal portal.	yes	24.90	26.13		good	10.98	12.79
	no	75.10	73.87		high	23.16	30.13
repositary°					very high	57.92	50.3
HEI has an online database with scientific production from the HEI.	no		40.99	**events***			
	yes		59.01	Student assessment of HEI's support for students to attend academic events.	very low	2.90	2.72
installations*					low	4.29	3.83
Student assessment of installation infrastructure.	very bad	6.38	3.74		reasonable	6.98	7.73
	bad	6.93	4.6		good	10.72	14.26
	reasonable	10.62	8.44		strong	17.08	22.69
	good	15.56	15.24		very strong	56.21	48.77
	very good	24.47	24.58				
	great	70.15	43.41				

Source: * Enade/INEP; ° Source: Censo Superior/INEP

In addition, Enade grades were, on average, higher for students from public universities and who did not receive funding (these variables must be related, since funding must occur in the case of students from private institutions). Especially in 2017, the grades of financed students were lower than those of other students, possibly due to the strong expansion that occurred in student financing between the two periods analyzed. It is likely that by benefiting a significantly larger number of students, funding programs will be less selective and demanding. Regarding the entry into the courses through quota policies, while in 2012 the students who used quotas presented a better average performance in the Enade tests, in 2017 the performance of this type of student was slightly

worse. This may also be a consequence of the increase in the number of quotas available. Another observed fact is that the average grade was, in general, higher in 2012, but the maximum grades were higher in 2017, perhaps due to the greater number of students who took the test in 2017. The detailed data of this graph are reported in Annex A.

Figure 2 – Grades according to Category Groups - 2013 and 2017



4 Results

After applying the exploratory factor analysis (EFA) by principal components for the validation of the theoretical model, it is possible to detect significant patterns between the original variables and extract the main factors. The results of the EFA produced three latent factors according to the Kaiser criterion (1958) with eigenvalues greater than one (Table 4). Both in 2012 and 2017, the first factor reveals a greater relationship with variables that describe the inputs of the HEIs and, therefore, can be called "Characteristics of the institutions" reflecting a little of the quality/offer of the institutions[17]. There is a positive interpellation between more professors per student, more professors with doctorates, more administrative technicians per student, greater expenditure on research and maintenance and the fact that the HEI is public and not private. This information must be associated with a higher quality offered by HEIs.

[17] Only the *funded* variable that presented a high factor load for factor 1 refers to the student, and not to the HEI (although in fact it is expected that there is such a relationship between this variable and IES inputs, since the financing is linked to the fact that the institution is private).

Table 4 - Exploratory Factor Analysis

Variables	Description	2012			2017		
		F1	F2	F3	F1	F2	F3
grade	Gross grade in the ENADE test.	0.28	-0.11	0.32	0.34	0.17	0.38
study	Hours per week devoted to studies.	0.21	0.01	0.64	0.20	0.03	0.73
books	Number of books read in the year.	-0.02	0.24	0.58	0.05	0.01	0.73
income	Total household income, in mw.	0.46	0.56	0.04	0.23	0.67	0.04
age	Age.	-0.49	-0.18	-0.15	-0.27	-0.39	-0.05
mother	Mother's education.	0.52	0.55	0.14	0.31	0.71	0.00
father	Father's education.	0.53	0.57	0.12	0.30	0.73	0.01
publicS	Type of high school.	-0.62	-0.56	0.08	-0.40	-0.76	0.09
funded	Scholarship or funding.	-0.58	-0.07	0.46	-0.89	0.30	0.06
quotaHol	Participation in quota system.	-0.27	-0.52	0.45	0.02	-0.58	0.22
teacher	Teacher to student ratio.	0.60	-0.33	0.08	0.73	-0.23	-0.01
PhD	Proportion of teachers with PhD.	0.75	-0.34	-0.05	0.86	-0.09	-0.04
technical	Number of Administrative technical staff per student.	0.65	-0.31	0.10	0.75	-0.14	-0.02
mainten	Expenses for maintenance of HEI.	0.55	-0.35	0.08	0.71	-0.08	-0.05
research	Expenses with Research and development.	0.22	-0.19	0.12	0.42	-0.08	-0.03
repository	HEI has an online database with scientific production.				0.78	-0.09	-0.10
portal	HEI library has access to the CAPES journal portal.	0.52	-0.33	-0.06	0.52	-0.01	-0.06
private	HEI is public or private.	-0.82	0.44	0.15	-0.88	0.32	0.05
eigenvalue		4.54	2.41	1.40	5.57	2.88	1.30
Cumulative Variance		0.28	0.43	0.52	0.32	0.49	0.57

Source: Own Preparation from INEP data, 2019.

In turn, the second factor reproduces the family background and the school trajectory of students who, in both years, presented higher factor loads for the variables of socioeconomic condition. In this second factor, students from richer families, who studied in private schools during high school, children of parents with higher levels of education, who did not participate in quota programs and, who are younger (age variable with negative sign) stand out. Finally, the third factor stands out as “Ability” for relating mainly the highest grades, more hours of study and more books read by students. Thus, students with higher grades, more hours of study and more books read have a higher score on this factor, that is, greater skills. The only variable that came out of the expected pattern of individual characteristics of students *versus* characteristics of institutions, is the variable that identifies students who receive funding. However, as the funding exists only for private college students, this information makes sense, since it is directly related to whether an institution is public or private.

From the exploratory factor analysis of the data, it was possible to observe the expected pattern of the variables, in such a way that it makes sense to continue with the confirmatory factor analysis to test the hypotheses of this study. We note that the EFA technique has shown the existence of three latent factors of the data structure that group in an unrestricted way the main variables observed in three subsets: “Skill”, “Institutions” and “Individual”. These factors mainly reproduce

the common relationships of a subset of observed variables consistent with the applied literature (e.g., Albernaz, Ferreira and Franco (2002); Machado et al. (2008); Soares and Sátyro (2008)). Therefore, we can restrict the latent factors according to the main factor loads of the observed variables and, thus, apply our theoretical model established by Figure 1. This restriction allows us to confirm the latent constructs by the measurement model in the CFA and, afterwards, to evaluate the relationship of the latent variables using the structural model.

Table 5 provides the results of the measurement model for both years (2012 and 2017). The path coefficients are the factor loads that represent the effect of the latent variable on each explanatory variable. The variable “hours of study” was chosen as a reference because it is the main determinant in the students' observed ability, with a very high load. It is worth mentioning that the number of books read by students is also an indication of higher skill. This means that, in addition to the grade’s variable in proficiency tests, which is widely used in the literature, the variable “hours of study” is of great importance for the analysis of the determinants of student ability.

In relation to individual characteristics, in convergence with the works by Baqueiro (2015), Felício and Fernandes (2005) and Rodrigues (2009), what is observed is that better socioeconomic conditions of the evaluated students are related to higher incomes, higher levels of the parents’ education, opportunity to enter higher education at a younger age and attendance to private schools during high school. The “parents' education” variables are the most important for the definition of the socioeconomic conditions of the students. In addition, it was also possible to observe that better socioeconomic conditions are related to students who do not receive funding and have not entered higher education through quotas. In other words, this work advances when verifying that the financing and quota policies do in fact serve lower income students.

The latent variable of Institutions, on the other hand, is related to a higher proportion of teacher by students, PhD teachers among the teachers, and administrative technicians serving the academic community, which signals a higher quality/supply of institutions, as well as in the works of Rios-Neto, César and Riani (2002), Riani and Rios-Neto (2008), César and Soares (2001); Ferrão and Cristiano (2003). In addition, access to the CAPES portal of journals, and the fact that universities are public and with higher expenses for maintenance and research, are related to a higher quality of educational institutions. In 2012, a higher proportion of doctoral professors and the fact of the HEIs being public are the main determinants for greater quality, as mentioned earlier. This shows that in Brazil, public investment in higher education tends to be stronger than private investment.

The Structural Model that relates the characteristics of Institutions and Individuals with the Ability of students is shown in the end of Table 5. For the year 2012, it was possible to observe an influence of variables that represent individual characteristics of students and characteristics of institutions in the performance of students, in line with the works of Brunello and Checchi (2005);

Ermisch and Francesconi (2001); Oppedisano (2011); Rios-Neto et al. (2002). Both the socioeconomic characteristics of students and the quality of educational institutions were able to positively influence students' ability, albeit in a small proportion. The correlation observed between the latent “Individuals” and “Institutions” was 0.33, which means that, in general, students with better social conditions attended courses with better quality indicators. This indicates that in 2012 there was still a strong inequality in access to higher education of quality to be corrected.

Table 5 - Results of the theoretical model[18]

Latent	Variable	2012			2017		
		coeff.	stand. Error	z-stat	coeff.	stand. Error	z-stat
Ability	grade	1.00	restricted		1.00	restricted	
	study	7.78*	0.32	23.97	0.54*	0.01	64.46
	books	1.77*	0.04	45.15	0.29*	0.01	52.18
Individual	income	1.29*	0.01	161.28	1.27*	0.00	246.16
	mother	1.98*	0.01	169.42	2.13*	0.01	260.42
	father	2.21*	0.01	166.59	2.29*	0.01	256.15
	age	-1.06*	0.01	-107.79	-0.94*	0.01	-159.40
	publicS	-1.97*	0.01	-127.28	-2.01*	0.01	-191.45
	funded	-0.63*	0.01	-84.90	-0.24*	0.00	-56.92
	quotaHol	-0.73*	0.01	-72.90	-0.61*	0.00	-114.06
Institution	teacher	1.50*	0.01	165.25	1.74*	0.01	281.95
	PhD	2.45*	0.01	161.82	2.53*	0.01	281.29
	technical	1.72*	0.00	293.87	2.06*	0.01	293.87
	mainten	1.30*	0.01	170.35	1.68*	0.00	310.47
	research	0.35*	0.00	64.15	0.64*	0.00	177.94
	repositpry	1.29*	0.01	109.08	2.32*	0.01	177.16
	private	-3.97*	0.04	-97.41	-3.60*	0.02	-198.98
	repository	-	-	-	0.89*	0.00	174.40
Ability	Individual	0.04*	0.00	21.55	0.33*	0.00	68.26
	Institutuion	0.05*	0.00	27.18	0.45*	0.00	97.02
Corr (Ind.,Inst.)		0.33			0.25		

Source: Own Preparation from INEP data, 2019.

The same analysis was carried out for the year 2017. Again, Ability is positively related to hours of study and to the number of books read, while the grade variable was restricted to one, although now students’ skill seems to depend less on the number of hours they study. In 2017, both the individual characteristics of the students and the characteristics of the HEIs showed a behavior remarkably similar to that observed in 2012. Once again, we can see that there is a relationship between students with higher incomes, more educated parents, younger students who studied in public schools during basic education, and who do not receive funding or use quotas. In addition,

[18] Restricting latent exogenous variables to have unitary variation as an identification constraint may be desirable (StataCorp, 2013), so the variances of the Background and Quality latent variables were restricted to 1.

parents' education remains the main determinant of students' better socioeconomic conditions. The relationship between school inputs that determine the quality of institutions was also the same as that observed in 2012. In other words, it is possible to say that a greater ratio of professors to students is associated with the presence of more professors with doctoral degrees, more technicians per student, higher expenses with research and maintenance, with the presence of academic repository of access to the Capes portal of journals and, mainly, with the fact that HEIs are of public administration. All these observations characterize higher quality educational institutions.

In the Structural model, in turn, social and the institution's characteristics contributed positively to students' ability in 2017. However, differently from what was observed in 2012, there is a greater importance of these exogenous latent variables in determining the ability of students, represented by higher loads, mainly for the quality of the HEIs. This means that investments in educational inputs– known for the infrastructure, materials and facilities offered by the institutions–are of great importance in student learning and are able to influence academic performance more than their individual social characteristics. This greater relative importance of the characteristics of the institutions in the performance of the students means that the quality factors of the human resources and infrastructure of the schools are able to improve the efficiency indicators of the education system. Thus, they help to minimize the importance of the student's social origin, in order to make the educational system more egalitarian, according to the results obtained by Rios-Neto, César and Riani (2002), Riani and Rios-Neto (2008) and Ferrão and Fernandes (2003).

This new situation of significant importance in the characteristics of educational institutions, may be the result of policies to expand access to higher education promoted by the federal government through subsidies and facilities for private education, especially in the form of student financing (such as cases of FIES and PROUNI), as well as through the expansion and restructure of federal universities in the public sphere. On the one hand, such expansion provided greater democratization of educational opportunities, allowing the participation of students who previously did not have access to higher education. Consequently, the characteristics of the students became more heterogeneous in the institutions [as pointed out by Mancebo, Vale and Martins (2015)]. On the other hand, the increase in the supply of higher education also brought reflections on the conditions necessary for the provision of quality education. Thus, the explanatory variables of inputs gained strong attention from the government, based on ideas of efficiency, which was reflected in their greater importance in the model.

The correlation measure between HEI Quality and Social Characteristics, in turn, was lower in 2017 (0.25) compared to 2012 (0.33). That is, although favorable social conditions are still positively related to attendance at higher quality institutions, this relation is lower for 2017. Students classified in lower social groups have greater access to quality education compared to

2012. In this way, the efforts of educational policies that expanded access to higher education and investments in research were able to reduce the disparities in access to quality education that exist among students from families with different socioeconomic conditions.

The results of the Structural Model point to the importance of characteristics that are external to the institutions (socioeconomic status and families of the students) as well as characteristics internal to the institutions in the performance of the students. This result coincides with that advocated by the works of Brunello and Checchi (2005); Albernaz et. al (2002); Felício and Fernandes (2005); Ferrão and Fernandes (2003) and Araújo and Siqueira (2010), which show that, although a good part of the variance of the average performance between schools is due to differences in the socioeconomic composition and in the family history of the students, school quality can constitute itself as an instrument to counter such determinism. What is observed is that, in fact, the inputs used in the HEIs matter to the students' ability. In addition, the analysis exposed here went beyond the discussions already presented by other works, and has shown the importance of educational policies and public investment to guarantee a more equal access for members of the population to a higher quality education that provides advances in learning.

5 Final remarks

Since the early 1990s, Brazil has witnessed a rapid expansion in the participation of students in higher education courses, reflecting the investment, financing and inclusion policies implemented in the period. However, the main concern is whether this expansion in Brazilian education operations has improved or influenced students' academic performance and skills. There is a consensus in the applied literature that these effects are marginally decreasing and even negative if evaluated at each development stage of an economy. In developing economies, the amount of human capital and education infrastructure is characteristically lower. Consequently, the effects of policies to expand higher education operations on student performance in developing economies tend to be relatively greater than in developed countries, where there are high amount of human capital and educational infrastructure. In particular, we tested this hypothesis of the marginal utility of school supplies, facilities and educational infrastructure for teaching in Brazil, whose provision is still on an upward trend within countries with this profile. Likewise, the students' academic performance is influenced by the economic and social structure in which they are inserted, as some applied studies postulate. There is no consensus whether students from poor families, when entering a certain college due to the support received by the financing and/or quota policy, would increase the average school performance. Part of the literature points out that this effect, in fact, must be negative. As in the Brazilian economy there is an expressive portion of young people outside higher

education, our hypothesis is that the inclusion policies in higher education start to have contributed positively to the average performance.

In this way, our article has contributed to this debate about the relationship between the expansion of higher education and school performance by testing the two main hypotheses mentioned above. Unlike the studies applied in the Brazilian literature, we propose a methodological strategy that allows the analysis of several multiple dependency relationships simultaneously. This allows us to assess the ability of Brazilian students more broadly in relation to the characteristics of higher education institutions, complementary inclusion policies and socioeconomic structures of the individual in Brazil for 2012 and 2017. More specifically, this assessment involved individual characteristics that students already bring with them, such as family income and parents' education, as well as what the institutions offer in terms of teaching and facilities. The information covered all higher education institutions in the country and involved more than 135,000 students in 2012 and 323,000 in 2017.

The main conclusion of the study is that the quality of HEIs contributes positively to the ability of students, which confirms the hypothesis that the effect of quality may be more significant in developing countries, such as Brazil. The expansion of higher education facilities and infrastructure in the country was important to make students more skilled and with a higher academic performance, which in the long run tends to increase the stock of human capital and reduce the distortions generated by social and income inequality itself in the country. Thus, according to the Brazilian experience in the last 20 years, any policy instrument against the expansive trend of educational inputs could restrict the access of individuals with lower social and income structures to higher education, even reducing the average skill of individuals and reducing the amount of human capital per job, the level of which is typically that of a developing country. In short, school inputs matter to the ability of students, the result of which is of interest for the conduct of educational policies in the country.

On the other hand, the individual characteristics of students also have a relationship with the performance they obtain in standardized tests of knowledge, as in the case with the Enade test. This means that educational institutions present in the country are not yet able to erase the disadvantages of social inequality. Another issue that could be observed was the effectiveness of the policies to expand access to higher education taken since 2012 to include students from lower social strata. The increase in the supply of financing and also the quota policy reached students with worse social conditions. Thus, they contributed to a more democratic higher education network, although in educational institutions with higher quality indexes, students with better social conditions still prevail. In this sense, it should also be noted that we observed that being a high-quality institution is strongly related to being a public institution. The fact that the observed ability of the students

maintains a relationship both with individual factors related to the socioeconomic conditions of the students, and with the characteristics of the HEI, shows the importance of maintaining policies that enable the access of poorer students to quality courses.

School inputs can accentuate or mitigate the effects generated by individual student characteristics. In this sense, universities can promote greater equity in the individual development of their students, offering access to resources capable of alleviating the inequalities generated by individual socioeconomic characteristics. This is a way to create opportunities for these students to be able to overcome the problems generated by their unfavorable social conditions and obtain greater school performance. This may be reflected in higher future wages, which should break the trend of social inequalities in the country, conditioned to an individual's initial living standards. Nevertheless, a limitation to be considered in this study is the use of Enade grades as the main indicators of students' ability, as there are no real incentives for students to commit to obtaining good results in this type of test. Even though it is difficult to apply, an alternative would be to look for other more real indicators of student performance, such as academic performance indexes within each institution, or monitoring students after they graduate.

References

Albernaz, A., Ferreira, F., & Franco, C. (2002). Qualidade e Eqüidade no Ensino Fundamental Brasileiro. *Pesquisa e Planejamento Econômico*, *32*(3), 453–476. Retrieved from http://ppe.ipea.gov.br/index.php/ppe/article/viewFile/139/74

Alves, M. T. G., & Soares, J. F. (2009). Medidas de nível socioeconômico em pesquisas sociais: Uma aplicação aos dados de uma pesquisa educacional. *Opiniao Publica*, *15*(1), 1–30. https://doi.org/10.1590/S0104-62762009000100001

Andrade, J. M. de, & Laros, J. A. (2007). Fatores associados ao desempenho escolar: estudo multinível com dados do SAEB/2001. *Psicologia: Teoria e Pesquisa*, *23*(1), 33–41. https://doi.org/10.1590/S0102-37722007000100005

Araújo, F. R. A. de, & Siqueira, L. B. O. de. (2010). Determinantes Do Desempenho Escolar Dos Alunos Da 4ª. Série Do Ensino Fundamental No Brasil. *Revista Economia e Desenvolvimento*, *9*(1), 70–103.

Athreye, S., & Cantwell, J. (2007). Creating competition? *Research Policy*, *36*(2), 209–226. https://doi.org/10.1016/j.respol.2006.11.002

Baqueiro, D. F. de A. (2015). *Equidade e Eficácia na Educação: Contribuições da Política de Assistência Estudantil na Permanência e Desempenho Discente*. Retrieved from https://repositorio.ufba.br/ri/bitstream/ri/19896/1/DISSSERTAÇÃO EQUIDADE E EFICÁCIA NA EDUCAÇÃO.pdf

Bartholomew, D. J., Steele, F., Galbraith, J., & Moustaki, I. (2008). Analysis of multivariate social science data, second edition. In *Analysis of Multivariate Social Science Data, Second Edition*. https://doi.org/10.18637/jss.v029.b13

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*. https://doi.org/10.1086/258724

Becker, G. S. (1972). Schooling and inequality from generation to generation: comment. *Journal of Political Economy*, *80*(3).

Blau, F. D., & Kahn, L. M. (2005). Do cognitive test scores explain higher U.S. wage inequality? *Review of Economics and Statistics*, *87*(1), 184–193. https://doi.org/10.1162/0034653053327649

Brooke, N., & Soares, J. F. (2008). *Pesquisa em eficácia escolar: origem e trajetórias* (UFMG, Ed.). Belo Horizonte: Editora UFMG Belo Horizonte.

Brunello, G., & Checchi, D. (2005). School quality and family background in Italy. *Economics of Education Review*, *24*(5), 563–577. https://doi.org/10.1016/j.econedurev.2004.09.001

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education*, *47*(1), 1–32. https://doi.org/10.1007/s11162-005-8150-9

Carneiro, P., & Lee, S. (2011). American Economic Association Trends in Quality-Adjusted Skill Premia in the United States , 1960 — 2000. *The American Economic Review*, *101*(6), 2309–2349. https://doi.org/10.2307/23045644

Castro, M. H. G. de. (2009a). A Consolidação da Política de Avaliação da Educação Básica no Brasil. *Meta: Avaliacao*.

Castro, M. H. G. de. (2009b). Sistemas de avaliação da educação no Brasil: Avanços e novos desafios. *Sao Paulo Em Perspectiva*, *23*(1), 5–18.

César, C. C., & Soares, J. F. (2001). Desigualdades acadêmicas induzidas pelo contexto escolar. *Revista Brasileira de Estudos de População*, *18*(1/2), 97–110. Retrieved from http://www.abep.nepo.unicamp.br/docs/rev_inf/vol18_n1e2_2001/vol18_n1e2_2001_6artigo_97_110.pdf

Chaves, V. L. J., & Amaral, N. C. (2016). Política De Expansão Da Educação Superior No Brasil - O Prouni E O Fies Como Financiadores Do Setor Privado. *Educação Em Revista*, *32*(4), 49–72. https://doi.org/10.1590/0102-4698162030

Coleman, J. S. (1968). Equality of Educational Opportunity. *Equity and Excellence in Education*, *6*(5), 19–28. https://doi.org/10.1080/0020486680060504

Costa, E. M., Ramos, F. de S., de Sousa, H. R., Sampaio, L. M. B., & Barbosa, R. B. (2015). Dinâmica da eficiência produtiva das instituições faderais de ensino superior. *Planejamento e Políticas Públicas*, (44), 51–84.

Dearden, L., Fitzsimons, E., & Wyness, G. (2014). Money for nothing: Estimating the impact of student aid on participation in higher education. *Economics of Education Review*, *43*, 66–78. https://doi.org/10.1016/j.econedurev.2014.09.005

Desjardins, S. L., Ahlburg, D. A., & Mccall, B. P. (2006). *The effects of interrupted enrollment on graduation from college : Racial , income , and ability differences*. *25*, 575–590. https://doi.org/10.1016/j.econedurev.2005.06.002

Ermisch, J., & Francesconi, M. (2001). Family matters: Impacts of family background on educational attainments. *Economica*, *68*, 137–156. https://doi.org/10.1111/1468-0335.00239

Felício, F., & Fernandes, R. (2005). O efeito da qualidade da escola sobre o desempenho escolar: Uma avaliação do ensino fundamental do Estado de São Paulo. *Anais Do XXXIII Encontro Nacional de Economia*, 20. Retrieved from http://www.anpec.org.br/encontro2005/artigos/A05A157.pdf

Fernandes, R., & Menezes-Filho, N. (2020). Charter schools, equity and efficiency in public education. *Education Economics*, 1–16. https://doi.org/10.1080/09645292.2020.1725959

Ferrão, M., Beltrão, K., Fernandes, C., Santos, D., Suárez, M., & Andrade, A. (2001). O SAEB– Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. *Revista Brasileira de Estudos de População*, *18*(1/2), 111–130. Retrieved from https://rebep.org.br/revista/article/view/347

Ferrão, M. E., & Fernandes, C. (2003). o efeito-escola e a mudança - dá para mudar? Evidências da investigação brasileira. *Revista Electrónica Iberoamericana Sobre Calidad, Eficacia y Cambio En Educación*, *1*(1), 5. Retrieved from https://dialnet.unirioja.es/servlet/articulo?codigo=638717

Flannery, D., & O'Donoghue, C. (2013). The demand for higher education: A static structural approach accounting for individual heterogeneity and nesting patterns. *Economics of Education Review*, *34*, 243–257. https://doi.org/10.1016/j.econedurev.2012.12.001

Gideon, A. (2017). Introduction: An Interdisciplinary Analysis of the Mission of European Higher Education Institutions and Potential External Constraints. In Springer (Ed.), *Higher Education Institutions in the EU: Between Competition and Public Service* (pp. 1–19). https://doi.org/10.1007/978-94-6265-168-5_1

Golob, T. F. (2003). Structural equation modeling for travel behavior research. *Transportation Research Part B: Methodological*, *37*(1), 1–25. https://doi.org/10.1016/S0191-2615(01)00046-7

Grossman, G. M., & Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European Economic Review*, *35*(2–3), 517–526. https://doi.org/10.1016/0014-2921(91)90153-A

Hair Jr., J., Black, W., Babin, B., & Anderson, R. (2014). Multivariate Data Analysis (MVDA). In *Pharmaceutical Quality by Design: A Practical Approach* (17th ed.). https://doi.org/10.1002/9781118895238.ch8

Hanushek, E. A. (1986). *The Economics of Schooling: Production and Efficiency in Public Schools*. *24*(3), 1141–1177.

Harris, D. N. (2007). Diminishing marginal returns and the production of education: An international analysis. *Education Economics*, *15*(1), 31–53. https://doi.org/10.1080/09645290601133894

Heinesen, E. (2018). Admission to higher education programmes and student educational outcomes and earnings–Evidence from Denmark. *Economics of Education Review*, *63*(November 2016), 1–19. https://doi.org/10.1016/j.econedurev.2018.01.002

Heyneman, S. P., & Loxley, W. A. (1983). The Effect of Primary-School Quality on Academic Achievement Across Twenty-nine High- and Low-Income Countries. *American Journal of Sociology*, *88*(6), 1162–1194. https://doi.org/10.1086/227799

Johnes, G., & Johnes, J. (2009). Higher education institutions' costs and efficiency: Taking the decomposition a further step. *Economics of Education Review*, *28*(1), 107–113. https://doi.org/10.1016/j.econedurev.2008.02.001

Johnes, J. (2006). Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review*, *25*(3), 273–288. https://doi.org/10.1016/j.econedurev.2005.02.005

Jones, C. I. (1995). R & D-based models of economic growth. *Journal of Political Economy*, *103*(4), 759–784. https://doi.org/10.1086/262002

Lam, P. L., & Shiu, A. (2010). Economic growth, telecommunications development and productivity growth of the telecommunications sector: Evidence around the world. *Telecommunications Policy*, *34*(4), 185–199. https://doi.org/10.1016/j.telpol.2009.12.001

Long, N. Van. (2019). Financing higher education in an imperfect world. *Economics of Education Review*, *71*, 23–31. https://doi.org/10.1016/j.econedurev.2018.06.004

Los, B., Timmer, M. P., & Vries, G. J. de. (2015). How global are global value chains? A new approach to measure international fragmentation. *Journal of Regional Science*, *55*(1), 66–92. https://doi.org/10.1111/jors.12121

Luz, L. (2006). Os determinantes do desempenho escolar: a estratificação educacional e o efeito valor adicionado. *XV Encontro Nacional de Estudos Populacionais.*, 1–20.

Machado, A., Moro, S., Martins, L., & Rios, J. (2008). Qualidade do Ensino em Matemática: Determinantes do Desempenho de Alunos em Escolas Publicas Estaduais Mineiras. *Revista Da Anpec*, *9*(1), 23–45. Retrieved from http://www.anpec.org.br/revista/vol9/vol9n1p23_45.pdf

Mallidis, I., Dekker, R., & Vlachos, D. (2012). The impact of greening on supply chain design and cost: A case for a developing region. *Journal of Transport Geography*, *22*, 118–128. https://doi.org/10.1016/j.jtrangeo.2011.12.007

Méndez, L. (2019). University supply expansion and inequality of opportunity of access: the case of Uruguay. *Education Economics*, 1–21. https://doi.org/10.1080/09645292.2019.1684448

Menezes-Filho, N. A. (2007). *Os determinantes do desempenho escolar do Brasil*. Retrieved from http://www.cepe.ecn.br/seminarioiv/download/menezes_filho.pdf

Oliveira, R. P. de, & Araujo, G. C. de. (2005). Qualidade do ensino: uma nova dimensão da luta pelo direito à educação. *Revista Brasileira de Educação*, (28), 5–23.

https://doi.org/10.1590/s1413-24782005000100002
Oppedisano, V. (2011). The (adverse) effects of expanding higher education: Evidence from Italy. *Economics of Education Review*, *30*(5), 997–1008. https://doi.org/10.1016/j.econedurev.2011.04.010
Oppedisano, V. (2014). Higher education expansion and unskilled labour market outcomes. *Economics of Education Review*, *40*, 205. https://doi.org/10.1016/j.econedurev.2014.03.005
Pinto, J. M. de R. (2004). O acesso à educação superior no Brasil. *Educação & Sociedade*, *25*(88), 727–756. https://doi.org/10.1590/s0101-73302004000300005
Rencher, A. C., & William, F. C. (2012). Methods of multivariate analysis: Third edition. In *Methods of Multivariate Analysis: Third Edition*. https://doi.org/10.1002/9781118391686
Riani, J. de L. R., & Rios-Neto, E. L. G. (2008). Background familiar versus perfil escolar do município: Qual possui maior impacto no resultado educacional dos alunos brasileiros? *Revista Brasileira de Estudos de Populacao*, *25*(2), 251–269. https://doi.org/10.1590/s0102-30982008000200004
Rios-Neto, E., César, C., & Riani, J. (2002). Estratificação educacional e progressão escolar por série no Brasil. *Pesquisa e Planejamento Econômico (Rio de Janeiro)*, *32*(3), 395–415.
Rodrigues, C. G. (2009). *A relação entre a expansão do acesso ao ensino e o desempenho escolar no Brasil: evidências com base no SAEB para o período de 1997 a 2005*.
Rojas, E., Sánchez, R., & Villena, M. G. (2016). Credit constraints in higher education in a context of unobserved heterogeneity. *Economics of Education Review*, *52*, 225–250. https://doi.org/10.1016/j.econedurev.2016.03.004
Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5), S71–S102. https://doi.org/10.3386/w3210
Senhoras, E. M., Takeuchi, K. P., & Takeuchi, K. P. (2006). A análise estrutural do ensino superior privado sob perspectiva. *Simpósio de Excelência Em Gestão e Tecnologia, III SEGeT*. Retrieved from https://www.aedb.br/seget/arquivos/artigos06/418_EnsinoSuperiorPrivado.pdf
Soares, José Francisco. (2007). Melhoria do desempenho cognitivo dos alunos do ensino fundamental. *Cadernos de Pesquisa*, *37*(130), 135–160. https://doi.org/10.1590/S0100-15742007000100007
Soares, Jose Francisco, & Andrade, R. J. de. (2006). Nível socioeconômico, qualidade e eqüidade das escolas de Belo Horizonte. *Ensaio: Avaliação e Políticas Públicas Em Educação*, *14*(50), 107–125. https://doi.org/10.1590/s0104-40362006000100008
Soares, S., & Sátyro, N. (2008). *O impacto da infra-estrutura escolar na taxa de distorção idade-série das escolas brasileiras de ensino fundamental*. Retrieved from http://repositorio.ipea.gov.br/handle/11058/1504
Suryadarma, D., Suryahadi, A., Sumarto, S., & Rogers, F. H. (2006). Improving student performance in public primary schools in developing countries: Evidence from Indonesia. *Education Economics*, *14*(4), 401–429. https://doi.org/10.1080/09645290600854110
Wail, B., Said, H., & Abdelhak, K. (2011). *A New Data Set of Educational Inequality in the World , 1950 – 2010*. 1950–2010. Retrieved from http://www.amse.ma/doc/WP_2012_09.pdf
Willms, D. J., & Somers, M. A. (2001). Family, Classroom, and School Effects on Children's Educational Outcomes in Latin America. *School Effectiveness and School Improvement*, *12*(4), 409–445. https://doi.org/10.1076/sesi.12.4.409.3445
Yang, C. S. (2018). An analysis of institutional pressures, green supply chain management, and green performance in the container shipping context. *Transportation Research Part D: Transport and Environment*, *61*, 246–260. https://doi.org/10.1016/j.trd.2017.07.005
Zoghbi, A. C., Rocha, F., & Mattos, E. (2013). Education production efficiency: Evidence from Brazilian universities. *Economic Modelling*, *31*(1), 94–103. https://doi.org/10.1016/j.econmod.2012.11.018

Annex A - Notes according to category groups - 2013 to 2017

Variables	Description	average grade		maximum grade*		Observations	
		2012	2017	2012	2017	2012	2017
grade		46.91	44.43	92.40	98.10	135139	323482
study	none	42.34	41.30	85.70	93.10	3585	12350
	one to three	44.96	42.30	92.40	98.10	54274	128459
	four to seven	47.50	44.97	88.20	96.40	41054	97796
	eight to twelve	48.99	46.83	89.30	97.00	18703	43047
	more than twelve	50.29	48.18	89.80	95.40	17523	4183
teacher	from 0 to 0.1	46.52	43.97	92.40	98.10	121279	27692
	from 0.1 to 0.2	50.25	47.17	89.00	95.10	12858	45451
	more than 0.2	51.22	47.08	85.60	91.60	1002	1107
PhD	until 0.4	45.10	42.10	92.40	93.10	94994	203153
	from 0.4 to 0.8	50.48	47.01	89.70	98.10	30668	88612
	more than 0.8	53.48	52.17	90.00	95.10	9477	31713
administration	private	45.08	42.53	92.40	97.00	98131	131695
	public	51.75	47.20	90.00	98.10	37008	191783
income	up to 3 salaries	45.52	42.44	92.10	93.30	51734	157255
	from 3 to 10 salaries	47.05	45.60	92.40	96.20	96877	140165
	more than 10 salaries	48.54	50.18	90.00	98.10	20558	26058
parents	until elementary school	44.41	40.88	87.60	90.10	21628	59563
	until high school	46.61	43.81	92.10	96.40	58764	15296
	until postgraduate	48.22	47.20	92.40	98.10	54747	110955
school	public	45.88	42.96	92.10	98.10	78597	221443
	private	48.34	47.63	92.40	97.00	56542	102035
financing	receive	46.45	42.55	92.40	93.10	52737	120494
	don't receive	47.21	47.60	90.00	98.10	82402	202984
quotaholder	yes	48.35	44.29	89.30	98.10	20590	247284
	no	46.65	44.91	92.40	93.10	114549	76194

Source: Own Preparation from INEP data, 2019.
Note: * The minimum scores were zero for all cases.

Annex B - Exploratory Factor Analysis with students' perceptions - 2012 and 2017

Variables	2012			2017		
	F1	F2	F3	F1	F2	F3
grade	0.26	0.00	0.24	0.32	0.18	0.19
study	0.16	0.16	0.35	0.18	0.03	0.29
books	-0.07	0.32	0.20	0.03	0.01	0.31
income	0.40	0.59	-0.13	0.22	0.68	0.03
age	-0.47	-0.22	0.05	-0.27	-0.39	0.02
mother	0.47	0.55	-0.20	0.31	0.72	-0.01
father	0.48	0.58	-0.19	0.30	0.74	0.00
publicS	-0.56	-0.58	0.21	-0.39	-0.76	-0.01
funded	-0.58	-0.10	0.12	-0.89	0.29	0.01
quotaHol	-0.22	-0.49	0.28	0.01	-0.57	0.07
teacher	0.59	-0.15	0.32	0.72	-0.22	0.12
PhD	0.75	-0.17	0.28	0.85	-0.08	0.09
technical	0.64	-0.13	0.32	0.73	-0.13	0.18
mainten	0.56	-0.21	0.27	0.69	-0.07	0.15
research	0.22	-0.10	0.19	0.42	-0.08	0.10
portal	0.51	-0.16	0.30	0.77	-0.08	0.09
repository	-	-	-	0.51	0.00	0.14
private	-0.87	0.34	-0.15	-0.89	0.32	0.01
learning	-0.43	0.45	0.56	-0.31	-0.04	0.78
installations	-0.45	0.49	0.51	-0.35	0.10	0.73
events	-0.34	0.42	0.62	-0.25	0.01	0.79
eigenvalue	4.82	2.64	1.93	5.76	2.89	2.10
Cumulative Variance	0.25	0.39	0.49	0.32	0.46	0.57

Source: Own Preparation from INEP data, 2019.