

Like Father, Like Son? An Analysis of the Effects of Circumstances on Student Performance in PISA 2012

Erik Figueiredo - UFPB and CNPq - Professor and researcher
Lauro Nogueira - UFERSA and UFPB - Professor and PhD student
Email: lauro.nogueira@ufersa.edu.br
Phone: (84) 9682-5261

15 de julho de 2014

Abstract

This study investigates three important issues related to student performance in the Program for International Student Assessment PISA 2012. First, we estimate the intergenerational transmission of education. Second, we verify whether circumstance variables influence effort variables in the PISA. Third, we carry out a counterfactual analysis of the improvement in the socioeconomic background of parents of those students who took the test. Results showed poor parental transmission of education in South American countries. Specifically, in Brazil, the rate is seven times lower than that of the Czech Republic and 20% of the total observed in France and in Japan. In addition, there are significantly positive effects of circumstances on individual effort. Moreover, the average gap for parental education hovers around 8%. About 2% of that is explained by the difference in parental education in the distribution.

Keywords: Intergenerational Transmission of Education, Equal Educational Opportunities, Treatment Effect.

Resumo

Este estudo investiga três importantes questões referentes ao desempenho na avaliação do Programa Internacional de Avaliação de Estudantes PISA 2012. Primeiro, estima-se a transmissão intergeracional da educação. Segundo, verifica-se se variáveis circunstanciais exercem efeitos sobre as variáveis de esforço no desempenho da avaliação PISA. Terceiro, faz-se uma análise contrafactual proveniente de um aumento no nível socioeconômico parental dos estudantes que prestaram o exame. Os resultados mostraram uma baixa transmissão educacional parental nos países da América do Sul. Especificamente, no Brasil, apura-se um valor aproximadamente sete vezes inferior ao encontrado na República Tcheca e 20% do total encontrado na França e no Japo. Adicionalmente, verificam-se efeitos significativamente positivos das circunstâncias sobre o esforço individual despendido. Além disso, o gap médio observado para educação parental em torno de 8%. Onde cerca de 2% explicado por diferença de nível educacional parental da distribuição.

Keywords: Transmissão Intergeracional da Educação, Igualdade de Oportunidades Educacionais, Efeito Tratamento.

JEL-Classification: I20; I21; D63.

1 Introduction

There is a consensus agreement that good-quality education is a strong indicator of welfare. Therefore, unequal education across regions, countries, and continents has been the target and subject of several public policies. Along this line of thought, Daude (2011) advocates that access to good-quality education is a powerful tool for the promotion of individual welfare, but certain conditions must be met for that to occur. For instance, all individuals need to have access to a homogenous good-quality education, regardless of their set of opportunities. Moreover, society has to acknowledge the importance of merit so that talent and individual skills prevail. Without these guarantees, the returns of investments in education are dissatisfactory, mainly for the most vulnerable ones in the society, thereby reducing intergeneration mobility.

In other words, education is deemed to be a key element for the economic and social development of any society. A qualified labor force increases the productivity of economic activities, allowing for large growth of sectors and making the absorption of technology and innovation easier. Besides these aspects, education is also essential for the good exercise of democracy, encouraging people to vigorously enforce their rights and duties as citizens.

Nevertheless, the 2005 report of the United Nations Educational, Scientific, and Cultural Organization (UNESCO) stresses that, despite the growing consensus over the importance of quality of education, the number of actions related to this concept is a lot smaller, actually. According to the report, two principles often make a distinction about the ways used to define quality of education. The first one regards the cognitive development of students as the major explicit goal of any educational system, endorsing their success as an indicative sign of its quality. The second one underscores the role of education in the promotion of shared values and in creative and emotional development. In this case, the achievement of these goals is way more complex to analyze.

Based on these concepts, social scientists such as Niknami (2010), Ferreira and Veloso (2003), Black, Devereux and Salvanes (2005) have recently investigated the intergenerational transmission of education. However, research results are few and far between and also inconclusive. In addition, most studies use educational attainment years of schooling instead of educational performance. Notwithstanding, results suggest a low level of transmission, chiefly in developing countries. These results contribute towards the intergenerational persistence of education and also towards a broader gap in inequality of educational opportunities.

However, numerous factors may be accountable for the low level of educational transmission, especially the assumption of poor quality of education in those countries. Even though schooling has improved in those economies in the past few years, educational attainment does not necessarily mean having equal opportunities, as years of schooling do not reflect the quality of education of a system, due basically to the existing heterogeneity of schools in their respective regions, countries, and continents.¹

That being said, understanding the underlying mechanisms in this relationship intergenerational transmission of education is of utmost importance. Therefore, the aim of the present paper is to investigate three important issues. First, the process of intergenerational transmission of education in the Program for International Student Assessment PISA associated with the theory of equal educational opportunities.² To do that, we use the educational production function proposed by Hanushek (1970,1979), with some peculiarities. Second, the analysis of whether circumstances have an impact on effort variables in PISA performance. In this context, we adopt the same approach developed by Bourguignon, Ferreira and Menendez (2007), adapted here to the context of educational opportunities. Third, we perform a counterfactual analysis based on the socioeconomic level improvement of the parents of those students who took part in PISA 2012. To do that, we used the counterfactual

¹See, for instance, Ferreira and Gignoux, (2008); Anshenfelter and Rouse (1998).

²For example, Ferreira and Gignoux, (2011).

inference method developed by Chernozhukov, Fernandez-Val and Melly (2013).

Note that there are major differences in this study. First, we checked the pattern of intergenerational transmission of education between the economies submitted to PISA 2012. Second, we decomposed the circumstance effects (both direct and indirect) on the performance in the assessment. Third, we simulated the counterfactual effect of possible public educational policies. Finally, the paper is organized as follows. Aside from this introduction, Section 2 provides a brief review on transmission and inequality of educational opportunities, focusing mainly on studies that use PISA data. Section 3 explains the methodology used, including the description and analysis of data. Section 4 describes the results. In the last section we make some remarks about the results.

2 Literature Review

This section introduces some studies on the topic. It initially highlights the study of Black, Devereux and Salvanes (2005). The authors investigate why better-educated parents have better educated children. According to their study, there are several possible explanations. However, two of them stand out. First is the case of pure selection or indirect effects. That is, better-educated parents earn higher salaries and, therefore, some variables contribute substantially to the education of their children. For example, enrolling their children in the best schools, buying them the best books, and investing in mechanisms that help them with learning. Second, the so-called causality or direct effect. In this case, having access to better education makes one a better parent, thus predisposing ones children to better educational outcomes. This direct relationship of causality occurs by means of potentially unobservable factors, such as shared environments and genetic inheritance.

On the other hand, according to Roemer (1998), two concepts of equal opportunities prevail nowadays in western democracies. The first one advocates that every individual with relevant potential in his/her learning period ought to be accepted as a possible candidate in the competition for positions in society. The second one, known as non-discrimination principle, establishes that in the competition for positions in society every individual who has relevant attributes for a given job should be included as an eligible candidate who will only be assessed based on relevant characteristics.

In the same vein, Lefranc, Pistoletti and Trannoy (2009) advocate that equal opportunities constitute a basic principle for the reduction of inequalities between individuals. Results then depend on a set of relevant deterministic and random factors for the economic agents success or failure. To analyze equality of opportunities, it is necessary to identify the variables individuals are and are not accountable for.

Nevertheless, there exist a large number of studies on the determinants of educational performance, which use educational success data, but most of them are just descriptive studies. For instance, Aguirreche (2012) looks into how the level of unequal opportunities of a nation affects the average performance of its students. A structural model based on Fleurbaey and Schokkaert (2009) was used. Results point out a high level of unequal opportunities, i.e., more than 30% of unfair inequality. In addition, there is a negative ratio of (-0.69) between unequal opportunities and educational performance.

In another study, which included PISA data for 2006-2009, Gamboa and Waltenberg (2012) assess unequal educational opportunities in six Latin American countries. To do that, they use the nonparametric approach proposed by Checci and Peragine (2010) to decompose an inequality index. The study demonstrates that unequal educational opportunities range from 1% to 25%, which denotes considerable heterogeneity across the six countries.

Likewise, Ferreira and Gignoux (2008) use PISA data for 2000 and calculate the level of inequality observed in educational performance, which is determined by the set of opportunities. In brief, the highest levels of unequal educational opportunities are reported for developing countries Latin America although there is considerable heterogeneity across developed countries. The weight of

circumstance variables in total educational inequality varies between 9% and 30% on math tests and between 14% and 33% on the reading test.

In another study, Ferreira and Gignoux (2011) propose two methods to infer educational inequality. The first one is targeted at educational performance variance or standard deviation and, the second one at educational opportunity percentage of variance that explains the influence of circumstances. Results indicate that 35% of all disparities in the educational performance of 57 countries that participated in PISA 2006 are due to unequal opportunities.

In general, the set of opportunities is decisive for the educational performance of individuals. These aspects are more remarkable in developing countries, especially in South American, Eastern European, and Asian countries. Conversely, the lowest levels of unequal opportunities are mostly observed in North American and Western European countries and in Oceania. And parental education is the circumstance that most affects the results.

3 Empirical Strategy

The traditional theoretical framework dealing with unequal opportunities, suggested by Roemer (1998), considers individual outcome, w_i , to be explained by a set of variables: i) circumstance, C_i , for example, family background, gender, region of birth, etc.; ii) effort, E_i , such as hours worked, time spent on reading, etc. and; iii) brute luck component, u_i . Where $f(\cdot)$ is an unknown function.

$$w_i = f(C_i, E_i, u_i). \quad (1)$$

However, combining these concepts with the theory of intergenerational transmission of education and adapting them to the education production function, henceforth EPF, proposed by Hanushek (1970, 1979), with some peculiarities, yields:

$$\ln(A_{it}) = g(F_i^t, P_i^t, I_i, S_i^t). \quad (2)$$

Where A_i^t is the educational vector of the i -th student at t ; F_i^t is the vector of individual and family characteristics of the i -th student accumulated at t ; I_i is the vector of the student body (peer influence), i.e., socioeconomic variables and family background of other students accumulated at t ; S_i^t is the vector of initial endowments of the i -th individual; and S_i^t is the vector of relevant school inputs for the i -th student accumulated at t . Actually, EPF analyzes how different inputs of the educational process may affect individuals educational outcomes.

By including control variables that denote the circumstances and individual efforts, it becomes identical to (3).

$$\ln(W_{f,i}^*) = \alpha + \beta_1 W_{p,i}^* + \epsilon_i. \quad (3)$$

Where $\ln(W_{f,i}^*)$ is the log of performance of family i 's child over time; $W_{p,i}^*$ is a vector of the child's socioeconomic characteristics, such as the log of parental education, type of school, preschool attendance, etc. and ϵ_i is a term orthogonal to $W_{f,i}^*$. That is:

$$E(\epsilon_i) = 0, E(\epsilon_i, W_{p,i}^*) = 0, E(\epsilon_i^2) = \sigma_\epsilon^2. \quad (4)$$

Where β_1 stands for the level of intergenerational transmission of education. Nonetheless, given some possible weaknesses of this approach, due mainly to omitted variables, and given the possible endogeneity of the vector of covariates, a method based on Bourguignon, Ferreira and Menendez (2007) is used. In this case, a dependence between circumstance and effort variables is assumed. So, the hypothesis assumed in (4) is relaxed. Mathematically, one has:

$$w_i = f(C_i, E_i(C_i, v_i), u_i). \quad (5)$$

After that, a link is established with the intergenerational transmission approach and a log-linear specification as proposed in (3) is assumed, implying:

$$\ln(\hat{w}_i) = \beta_1 W_{p,i}^* + \beta_2 E_{f,i}^* + u_{f,i}. \quad (6)$$

Additionally, one assumes the endogeneity of certain variables that represent individual circumstances, yielding:

$$E_{f,i}^* = \gamma W_{p,i}^* + v_i. \quad (7)$$

Where β_1 and β_2 denote two coefficient vectors. In addition, u_i and v_i are random terms with white noise properties that denote unobservable circumstance and effort variables, as well as the luck factor. And γ is a matrix of coefficients that links specific circumstance variables to effort variables. In brief, it is assumed that coefficients of this matrix influence effort variables.

This structure allows some effort variables to be clearly influenced by circumstances. It is thus possible to observe the direct (β_1 and β_2 indirect effects of circumstance variables on individual educational performance. Now let $M(w)$ denote the marginal distribution of performance obtained from (5). Thereafter, two counterfactuals are constructed: i) in the first one, the overall effects of circumstances are canceled out, i.e., $w_i = f(\bar{C}_i, E_i(\bar{C}_i, v_i), u_i)$; and ii) in the second one, only the direct effects of circumstances are canceled out, $w_i = f(\bar{C}_i, E_i(C_i, v_i), u_i)$. Therefore, if expressions (6) and (7) can be estimated sensibly, the level of inequality persistence in individual performance is given by:

$$\Theta_I = \frac{I(M(w)) - I(M(\tilde{w}))}{I(M(w))}. \quad (8)$$

In other words, the resulting vector of performance in (8) contains the overall inequality arising from circumstance variables. On the other hand, equation (9) shows the level of inequality persistence arising from direct circumstance effects.

$$\Theta_I^d = \frac{I(M(w)) - I(M(w^d))}{I(M(w))}. \quad (9)$$

So, indirect circumstance is simply determined by:

$$\Theta_I^I = \Theta_I - \Theta_I^d. \quad (10)$$

Fortunately, for the decompositions proposed in (8) and (9), it is not necessary to compute structural systems (6) - (7). Then, substituting (7) into (6), one gets the reduced form:

$$\ln(\hat{w}_i) = (\beta_1 + \beta_2 \gamma) W_{p,i} + v_i \beta_2 + u_{f,i}. \quad (11)$$

However, as underlined by the authors, the omission of relevant variables in the determination of the outcome and the possible endogeneity of circumstances cause the parameters obtained in (5) and (11) to be biased and, consequently, counterfactuals are biased as well. That being said, and in the absence of an appropriate set of instrumental variables, the authors suggest a Monte Carlo simulation to create intervals for the estimated coefficients. In order to circumvent such problems, this study utilizes the post-estimation method bootstrapping replicating random samples one thousand (1,000) times from the original sample and thus constructing a confidence interval for the estimated parameters.³ Second, an attempt was made to eliminate possible biases from the construction of counterfactuals using a quantile approach, i.e., not based upon estimations of average effects alone.

³By using command `sqreg` in STATA 12, one estimates simultaneous quantile regressions for certain quantiles. It produces the same coefficients as the quantile regression for each quantile. The reported standard deviations are similar; however, the variance-covariance matrix of estimators (VCE) is obtained via bootstrapping. VCE includes blocks between the quantiles. Thus, it is possible to test and construct confidence intervals, comparing coefficients that describe different quantiles.

After that, in order to determine to what extent an increment in social background circumstances affects the performance on the test, as these factors are decisive for individual outcomes,⁴ one uses the inference on counterfactual distributions method developed by Chernozhukov, Fernandez-Val and Melly (2013), henceforth ICD. The use of ICD is justified by the fact that this method is based on different major approaches to estimate conditional quantile functions and conditional distribution functions. Another advantage of this method is that it can be used to analyze the effect of simple interventions (single change in a characteristic) and in complex changes (general changes in the distribution of characteristics).

ICD is especially employed in cases in which an intervention policy gives rise to a change by modifying part of the distribution of the set of explanatory variables X - covariates - which determine the response in the variable of interest Y . In other words, ICD consists in estimating the effect on the distribution of Y given a change in the distribution of X .

The results observed are obtained from the sample before the policy intervention and, therefore, observable, whereas counterfactual outcomes stem from the sample after policy intervention and, therefore, unobservable. Then it is admitted that covariates are observable before and after policy intervention. That means that the observed outcomes are used to establish the relationship between the variable of interest and covariates, which, alongside the counterfactual distribution of covariates, determine the distribution of the outcome after the intervention under some conditions.

According to the authors, in order to specify a model that allows checking how the counterfactual outcome is produced, it is important to analyze the relationship between the observed result and covariates using a conditional quantile representation. For example, let W^0 denote the observed result, and X^0 be $(P \times 1)$ vector of covariates with distribution function F_X^0 before the intervention policy. Where $Q_Y(u|X)$ denotes the u quantile conditional on W^0 given X^0 . Thus, the result W^0 can be linked to the conditional quantile function via the Skorohod representation, i.e.:

$$W^0 = Q_W(U^0|X^0), \longrightarrow U^0 \sim U(1, 0) \text{ independently from } X^0 \sim F_X^0. \quad (12)$$

Where (12) highlights that the result is a function of covariates and of error term U^0 . In classic regression models, the error term is separated from independent variables as in point regression models, but in general, it does not have to be. This method encompasses both cases.

In fact, the counterfactual inference process consists in designing the vector of covariates for a different distribution, i.e., $X^C \sim F_X^C$, where F_X^C is a known distribution function of covariates after the intervention policy. Hence, under the assumption that the conditional quantile function is not altered by the policy, the counterfactual outcome Y^C is obtained from:

$$W^C = Q_W(U^C|X^C) \longrightarrow U^C \sim U(1, 0) \text{ independently from } X^C \sim F_X^C. \quad (13)$$

Additionally, ICD assumes that the quantile function $Q_Y(u|x)$ can be assessed at each point of x in the base of distribution of the covariates of F_X^C . This assumption requires that the base of F_X^C be a subset of the base of F_X^0 , i.e., that the quantile function be properly extrapolated. These assumptions are formulated as follows.

- S1 the conditional distribution of the result based on the covariates is the same before and after the intervention policy;
- S2 the conditional model holds for all $x \in X$, where X is a compact-sized subset of \mathbb{R}^P that contains bases of F_X^0 and F_X^C .

Furthermore, ICD contemplates two different types of changes in the distribution of covariates. First, the covariates are designed for a different subsample before and after the intervention. This

⁴See, for instance, Barros (2009).

subsample can correspond to different demographic groups types time periods or geographical locations. For instance, workers characteristics in different years, parental socioeconomic distribution for white or black individuals, or occasionally, for distribution of covariates in treatment groups versus a control group. Second, the intervention can be used as a known transformation of the distribution of observed covariates. In summary, $X^C = g(X^0)$, where $g(\cdot)$ is a known function. For example, unit changes in the location of one of the covariates, $X^C = X + e_j$, where e_j is a unit vector ($P \times 1$) with one at position j ; or it can preserve the redistribution of covariates implemented as $X^C = (1 - \alpha)E[X^0 + \alpha X^0]$.

This type of intervention can be used, for instance, to estimate the effect on food expenses brought about by a change in income tax; the influence on real estate prices as a result of the removal of hazardous waste in the surrounding area, and so forth. Note that both cases above describe different experimental scenarios. The main difference is that the second case corresponds to an almost perfectly controlled experience, which provides additional information for the identification of more characteristics of the joint distribution of results before and after the intervention. However, to make inferences about the overall effect on the outcome, caused by the intervention, it is necessary to identify the distribution and quantile functions of the outcome before and after the policy. The conditional distribution function associated with the quantile function $Q_W(u|x)$ is expressed by:

$$F_W(w|x) = \int_0^1 1 \{Q_W(u|x) \leq w\} du. \quad (14)$$

Given the assumptions about how the counterfactual outcome is produced, the marginal distribution is expressed by:

$$F_{W^j} = Pr\{W^j \leq w\} = \int_x F_W(w|x) dF_x^j(x). \quad (15)$$

With corresponding u quantil marginal functions.

$$Q_{W^j}(u) = inf\{W : F_{W^j}(w) \leq u\}. \quad (16)$$

Where j corresponds to the status before or after the intervention, $J \in (0, C)$. The u quantil treatment effect of the intervention policy is determined by:

$$QTE|W(u) = Q_{W^C}(u) - Q_{W^0}(u). \quad (17)$$

Analogously, the effect on u distribution of the policy is expressed by:

$$DE_W(w) = F_{W^C}(w) - F_{W^0}(w). \quad (18)$$

ICD allows determining other functions of interest, for example, the Lorenz curve of the observed counterfactual outcome.

$$L_W^j(w) = \frac{\int_{-\infty}^w t dF_W^j(t)}{\int_{-\infty}^{\infty} t dF_W^j(t)}. \quad (19)$$

Provided that integrals exist and $\int_{-\infty}^{\infty} t dF_W^j(t) \neq 0, J \in (0, C)$. Nevertheless, it is often more interesting to estimate the marginal distribution functions of the outcome before and after the intervention.

3.1 Data Description and Analysis

The results of this paper will be available in PISA microdata for 2012. The program is developed and coordinated by the Organization for Economic Cooperation and Development (OECD). There is a national coordination body in each participating country. In Brazil, the assessment is coordinated by the National Institute of Educational Studies and Research Ansio Teixeira (Inep). This database is chosen for the wealth of information it provides. PISA is an international survey carried out every three years to assess educational systems all over the world. It seeks to test the skills and knowledge of 15-year-olds. The 2012 version is the fifth edition of the program and assesses the competences of students in reading, math, and sciences, giving emphasis to math.

An interesting fact is that around 510,000 students from 65 countries participated in the 2012 edition, accounting for about 28 million teenagers worldwide. Moreover, it is considered to be the only survey that assesses activities that are not totally related to the school curriculum. As a matter of fact, the tests are designed to assess to what extent students at the end of elementary school are able to apply their knowledge to real-life, everyday situations and how able they are to participate fully in society. In addition, the information collected by questionnaires provides ancillary data for the interpretation of results.

Tabela 1: Average Performance - Test Scores - PISA 2012

Country/Region	Mathematics	Reading	Science	Average
South America	394.39	410.73	408.51	404.54
Argentina	396.47	403.99	410.58	403.68
Brazil	383.42	400.98	395.9	393.43
Chile	445.88	461.01	465.61	457.50
Colombia	384.98	413.28	407.99	402.08
Peru	367.19	383.70	372.82	374.57
Uruguay	411.96	414.53	419.14	415.58
OECD	488.30	490.81	494.86	491.33
Germany	515.48	509.28	535.37	516.71
Canada	509.29	510.95	514.36	511.53
United States	481.03	497.62	497.57	492.07
France	498.42	509.27	502.53	503.41
United Kingdom	499.18	498.09	509.71	499.18
Spain	495.82	494.45	504.83	498.37
Japan	534.99	536.67	545.77	539.14
Mexico	418.44	428.84	419.59	422.29
Asian	540.20	515.12	525.58	526.97
Shanghai	612.03	569.36	580.33	587.24
Singapore	566.89	536.55	545.66	549.70
Hong Kong	559.83	543.35	554.28	552.49
Chinese Taipei	557.84	521.22	521.78	533.61
South Korea	552.61	535.61	537.14	541.78

Note: Own elaboration from data from PISA 2012.

As to the preliminary analyses of the data, reported in Table 1, note that there is a high level of inequality in the average performance in all fields of knowledge, especially in South American countries. For example, Brazil, in spite of being one of the 10 largest economies on the planet, according to the International Monetary Fund (IMF), had an extremely modest result. However, recall that the results do not change significantly in relation to the method and reports used by the program. According to PISA, Brazil ranks in the 58th position among the 65 surveyed countries.

On the other hand, the performance of most Asian countries was excellent, with higher averages than those obtained by OECD countries. As demonstrated by the data, the five best outcomes belong to Asian countries. Additionally, there is a remarkable difference in the math test performance, which is the focus of PISA 2012. Specifically, Shanghai scored approximately 90 points compared to Japan,

an OECD member. Asian countries outperformed South American ones by nearly 30%. For example, in comparison with Chile which had the best performance among South American countries Singapore showed a positive difference around 28% in favor of Asian countries, despite a larger variance in average scores in the latter. This occurs due to the poor performance notably obtained by Malaysia and Indonesia.

That being said, the microdata used are split into: i) variables related to the students; ii) variables related to the parents (characteristics of individuals and their families) iii) variables related to the schools (specific characteristics). The combination of these databases allows compiling information about circumstance and effort variables and individual outcomes.

Nonetheless, four specification tests were performed to choose the best set of covariates.⁵ The following tests were carried out: *ovtest*, *backward stepwise*, *forward stepwise*, and *hierarchical*. *Ovtest* aimed to detect problems with omitted variables. The other three tests were used to identify the inclusion or exclusion of variables.⁶ Besides the observation of test results, classic variables used in the literature were also chosen.⁷ Based upon this, in the first approach, the log of the individual average on PISA was regressed as a function of the circumstance variables described in what follows. The second approach included variables that express individual effort. Specifically, the variables used in this paper are summarized in Chart 1.

Chart 1 - Description of Variables

VARIABLES	
Individual results	log in arithmetic mean of points obtained tests of Languages; Mathematics and Science.
Variables for Circumstances	
Parental education	log greater degree of parental education - years of education - the father or the mother.
Educational difference	difference in years of education of father and mother.
School Type	assumes value 0 for public, 1 for private school.
Gender	assumes value 0 for male and 1 for female.
Preschool	assumes value 0 preschool for 1 met otherwise.
School Location	assumes value 0 for rural and 1 for urban areas.
Family structure	assumes value 0 to biparental 1 for Two and One Parent Families.
Siblings - Brothers	assumes value 0 if the individual has siblings living at home and 1 otherwise.
Variables Individual Effort	
Repetition of Year	assumes value 0 if the student has already repeated one year and 0 if you have never repeated.
Migrant	assumes value 0 for migrant e 1 otherwise.
Perseverance	based on the responses of students about their willingness to work on problems that are difficult.
Real effort	index of self-reported effort.
Effort Potential	response rate of the test

Note: Own elaboration from data from PISA 2012.

As pointed out earlier, the first approach uses only the vector of circumstance variables, providing information about 336.286 students out of 422.413 from 58 countries. On the other hand, in the second approach, the vector of variables that denote effort is included in the estimations. At this stage, data on 225.629 students from 58 countries are used. The reason why the data neither reflect the available total nor the failure to include all countries had to do with missing information. However, the restriction concerning the nature of the data used is a consensus, given that, in assessments in which there are no direct consequences for students, teachers, or schools, students skills are likely to be underestimated.⁸

After such remark, we now provide a brief account of the PISA sample. Table 2 summarizes some characteristics of the database. In the analyzed data, nearly 50% of those who took the test in 2012

⁵Although *ovtest* has indicated the existence of omitted variables supposedly skill and motivation - it also regards the log-linearized model as the best one.

⁶For further details, see Chatterjee and Hadi (2013).

⁷For example, Firmo and Soares (2008); Bauer and Riphahn (2007).

⁸See, for instance, Quintano, Castellano and Longobardi (2009).

are female; 80% of the students attend a public school; approximately 86% of the candidates attended preschool for at least one year; 61.5% belong to OECD countries, 8.9% of which are from South America. Furthermore, about 37% of parents have at least a college degree. And only around 4% of students were born to individuals who did not go beyond the first cycle of elementary education.

Tabela 2: Descriptive Analysis

Students for Variables Not Responsibility				
Gender		Members - Excetos other		
Male	Female	OECD	South America	Asian
49,42%	50,48%	61,5	8,90%	9,12%
Education Level Father - Mother				
Primria		Bsica	Secundria	Superior
3,42%		21,28%	32,66%	35,94%
3,89%		21,42%	33,48%	37,41%
School Type			Preschool	
Public		Private	Attended	Not attended
80,58%		19,42%	0,869	13,10%

Note: Own elaboration from data from PISA 2012.

4 Results

This section presents the major results of the paper. Table 3 reports the estimations of equation (3)⁹ for the general sample. As expected, results suggest strong influence of circumstances on individual outcomes. The set of opportunities accounts for approximately 35% of all disparities observed on the test.

Tabela 3: Determinants of Individual Result - Average

Variables	OLS	QREG	SQREG
Parental education	0.1449*** (0.0010)	0.1610*** (0.0013)	0.1610*** (0.0015)
Dif. education	-0.0024*** (0.0002)	-0.0021*** (0.0003)	-0.0021*** (0.0003)
School Type	0.0468*** (0.0008)	0.0477*** (0.0011)	0.0477*** (0.0010)
Gender	-0.0172*** (0.0006)	-0.0121*** (0.0008)	-0.0121*** (0.0008)
Preschool	0.1150*** (0.0010)	0.1181*** (0.0013)	0.1181*** (0.0014)
Local School	0.0440*** (0.0012)	0.0416*** (0.0015)	0.0416*** (0.0014)
Family structure	0.0543*** (0.0009)	0.0551*** (0.0012)	0.0551*** (0.0009)
Siblings	0.0226*** (0.0009)	0.0261*** (0.0012)	0.0261*** (0.0011)
Intercept	5.6000*** (0.0028)	5.5704*** (0.0037)	5.5704*** (0.0040)
Observations	336286		
Adj. R2	0.138		

Note: Standard deviations in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

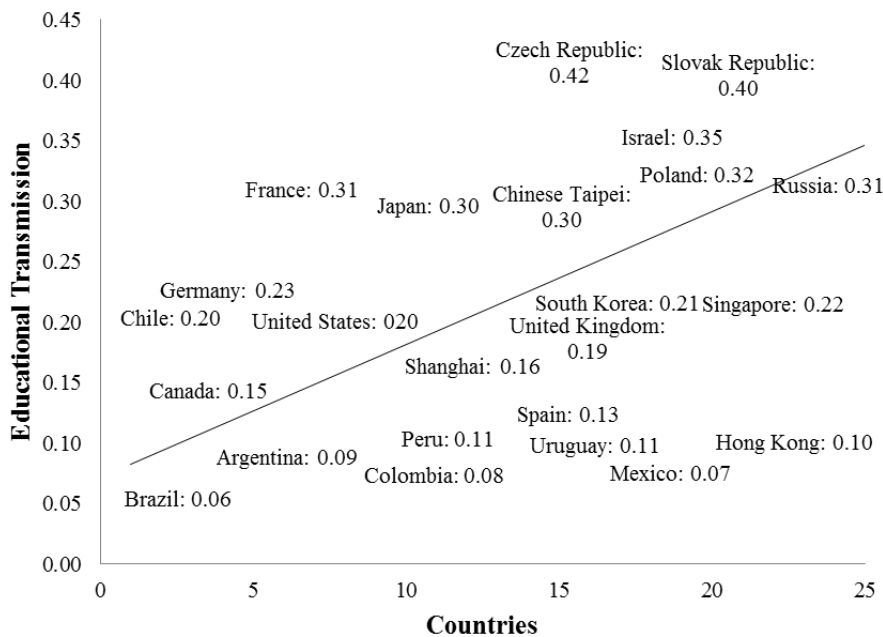
Following this same line of thought, it should be noted that the characteristics that most influenced the results of PISA 2012 were, respectively, parental education, preschool attendance, type of school,

⁹(2) was estimated by OLS, quantile regression, and quantile regression dealing with heteroskedasticity.

location of the school, educational difference between parents, and gender.¹⁰ In sum, as demonstrated by other studies (e.g., Ferreira and Gignoux (2011)), social background circumstances has a significant impact on outcomes. Another interesting aspect can be observed as to the negative effect of parental educational difference on the test. The negative influence of this variable informs that the larger the asymmetry educational disparity between fathers and mothers the lower the score. This finding is consistent with the literature (e.g., Bourguignon, Ferreira and Menendez (2007)), although it is still quite incipient. In addition, in upper quantiles, the sex variable sometimes is not statistically different from zero.

Additionally, the influence of parental education on test results is more striking in developed countries. All in all, intergenerational educational transmission is clearly stronger in these countries. The results shown in graph 1 illustrate this situation. Note that, for example, Argentina and Brazil together have parental educational transmission less than 50% of the effect observed for the Japanese, French, and Chinese, and almost three times lower than that found for the Czech and the Slovak. Specifically, Argentina and Brazil occupy, respectively, the 51st and 56th position in the educational transmission ranking among 58 countries. According to general estimates, Brazil only has slightly better results than Macao and Qatar.

Figure 1 - Parental Influence Educational PISA test - Average



Note: Data Research

Could these results suggest, for example, in the theory of educational mobility, that Brazil and Argentina have low educational persistence and, therefore, high mobility? And that the Czech Republic, Slovakia, Japan, and France have high persistence and, therefore, low mobility? The answer is no, because this study employs an educational assessment system PISA 2012 rather than income or educational attainment as most studies in this area. Several studies that indicate low mobility in developing countries confirm this. In these countries, individuals who attend college or have a college degree are often children of individuals who have at least a college degree.¹¹ In developed countries, this relation is not so predominant, as there are more opportunities and better quality of education.

¹⁰The results remained unchanged for the average score in sciences, languages, and math. See tables in the appendix.

¹¹See, for instance, Daude (2011) and Hertz et al. (2007).

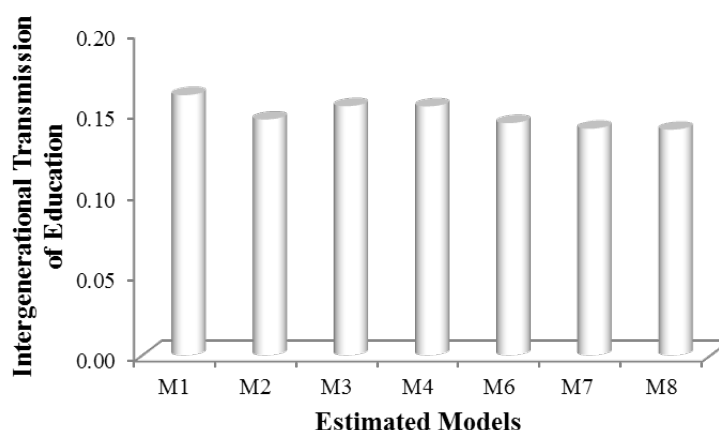
According to Black and Devereux (2010), it can be conceptually admitted that the educational choice of young people is associated with some factors, especially with educational costs, educational returns, and family income. The latter is related mainly to cases with credit restrictions. On the other hand, there appears to be some agreement that educational returns are larger for more skillful individuals and also for those with better-educated parents. These hypotheses imply that individuals with better educated parents tend to pursue a higher level of education because of: i) direct effects better educated parents also interpreted as causal channel, and, ii) indirect effects, i.e., being more skillful inheritability evinced by the intergenerational transmission of education.

Moreover, according to the authors, it is possible that underlying mechanisms bring direct effects of parental education on the performance of individuals. In general, the better the parental education, the higher the family income and, consequently, the more positive the impact on school performance. Second, this characteristic may expand the time devoted to developing remedial activities with their children. Also, it allows increasing the bargaining power of families, as more educated mothers can be more susceptible to spending on investments and activities targeted at teenagers or children in their families.

Another important aspect is the presence of qualified intergenerational educational mobility. In brief, developed countries usually have a higher level of education and, therefore, will have better educated children. Specifically, in Brazil and in developing countries, this might not hold. On average, parental education years of schooling of parents in these countries is lower, as reported in Table A.4 in the Appendix.

Looking at Table A.4, another interesting pattern concerns the value obtained for variables such as: type of school, preschool attendance, and location of the school, which plays a crucial role in developing countries, suggesting heterogeneity in quality of education, and differs from what is observed for developed countries, where these variables do not have a strong influence on school performance. Particularly in Brazil and Argentina, the simple fact of attending a private school affects the PISA score (17% and 14%, respectively). However, in countries such as Germany and Japan, these effects are incipient.

Figure 2 - Parental Influence Educational



Note: Own elaboration from data from PISA 2012.

Some variables related to individual effort were added in this line of research, although, in principle, the hypothesis postulated in equation (4) was assumed. However, before discussing the results shown in Table 4, graph 2 shows eight different specifications that take into account continental dum-

mies. The estimated results are clearly similar.¹² Notwithstanding, the specification tests inform that the general model has a better fit without the continental dummies.

In other words, the results described in Table 4 do not differ significantly from those reported in Table 3.¹³ Conspicuously, the estimated coefficients effects of the covariates albeit smaller, are quite similar, thereby suggesting good fit of the model. Additionally, the afore-mentioned identification tests corroborate these results. Nevertheless, two facts deserve special attention among the covariates

Tabela 4: Determinants of Individual Result - Average

Variables	OLS	QREG	SQREG
Parental education	0.1361*** (0.0012)	0.1489*** (0.0015)	0.1489*** (0.0012)
Dif. education	-0.0012*** (0.0002)	-0.0012*** (0.0003)	-0.0012*** (0.0004)
School Type	0.0503*** (0.0009)	0.0515*** (0.0012)	0.0515*** (0.0010)
Gender	-0.0056*** (0.0007)	-0.0019** (0.0009)	-0.0019*** (0.0007)
Preschool	0.0996*** (0.0012)	0.1029*** (0.0016)	0.1029*** (0.0013)
School location	0.0430*** (0.0013)	0.0391*** (0.0017)	0.0391*** (0.0020)
Migrant	0.0204*** (0.0011)	0.0175*** (0.0015)	0.0175*** (0.0015)
Repeat student	0.1124*** (0.0010)	0.1192*** (0.0013)	0.1192*** (0.0011)
Perseverance	0.0160*** (0.0004)	0.0145*** (0.0005)	0.0145*** (0.0004)
Real effort	0.0038*** (0.0002)	0.0047*** (0.0003)	0.0047*** (0.0003)
Effort Potential	-0.0009*** (0.0001)	-0.0010*** (0.0001)	-0.0010*** (0.0001)
Intercept	5.5540*** (0.0038)	5.5228*** (0.0049)	5.5228*** (0.0039)
Observations	225629		
Adj. R2	0.190		

Note: Standard deviations in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

that represent individual effort in the results. First, the strong influence of a student not having repeated a grade in school. That is, this influences, on average, 11.2% the test performance. Second, the small effect attributed to actual effort and potential effort, with an emphasis on the sign of the latter, indicating a high rate of non-response. This characteristic shows the limitation of these data. In other words, as the test theoretically does not have any influence on the life of students, teachers, and schools, there is a high rate of non-responses, especially regarding more complex questions. In the same vein, the coefficient of the perseverance variable seemingly supports this body of evidence.

As to school-related variables such as type of school, location of school, and preschool, they have been suggested to account for nearly 20% of the differences between the performances. Of this variable, preschool attendance is the most important. This variable is even more remarkable when we take into consideration the results of Asian countries. That is, in these economies, the influence of preschool attendance is, on average, four times stronger than that observed for OECD and South American countries, and approximately three times stronger than that of Nordic countries. Nonetheless,

¹²M1 no dummy for continents; M2 dummy for OECD; M3 is equal to M2 plus a dummy (Asian countries); M4 is equal to M3 with inclusion of a dummy (Nordic countries); M5 general model including the vector of individual efforts. For the other models, continental dummies are added one at a time.

¹³The results remained unchanged for the average score in sciences, languages, and math. See tables in the Appendix.

although the general results for the type of school public or private demonstrate a smaller influence on test results, when we assess the results of South American countries, this influence is nearly five times more effective than in OECD countries, about seven times the effect observed in Asian countries, and infinitely larger than that of Nordic countries.

Notwithstanding, there is empirical evidence of possible endogeneity between the covariates circumstance and effort variables and, therefore, it produces biases in the estimated coefficients. Additionally, despite the inclusion of variables that represent individual effort, not all variables omission that influence the result are available. In this respect, in an attempt to check and correct these problems, the method based on Bourguignon, Ferreira and Menndez (2007) is used. In sum, it is assumed that there is some dependence between circumstance and effort variables. That is, the assumption made in (4) is refuted.

Taking this scenario into consideration, we first calculate the level of inequality for the factual distribution of the overall average of results. After that, we estimate the reduced-form coefficients equation 10 for the respective quantiles (0.25; 0.50 and 0.75). In turn, using the mean coefficients estimated in (10), as well as the lower and upper extremes, we simulated the counterfactual distribution $M(w)$ originating from expression: $\tilde{w} = \exp[(\beta_1 + \gamma\beta_2)\overline{W}_{p,i}^* + v_i\beta_2 + u_{f,i}] \rightarrow \tilde{w} = \exp[\beta_i\overline{W}_{p,i}^* + e_i]$, onde $\beta_1 + \gamma\beta_2 = \beta_i$ e $\epsilon_i = v_i\beta_2 + u_{f,i}$.

In brief, the level of inequalities for factual distributions w and \tilde{w} is calculated. After that, following equation (7), one obtains the total participation of the set of opportunities in the inequality of scores. Note that the standard deviation of logs was used as level of inequality. Similar results were obtained with the coefficient of variation. However, other indices are broadly used in the literature, e.g. Ferreira and Gignoux (2011); and Lefranc, Pistolesi and Trannoy (2009).

Likewise, the partial or direct effect of the participation of opportunities observed in total inequality is estimated. First, (5) is estimated directly, and then, by using the estimated coefficients, counterfactual distribution $M(w_i^d)$ is built, derived from: $w_i^d = \exp[\beta_4\overline{W}_{p,i}^* + E_i] + u_i$. It should be noted that, in the first counterfactual $M(\tilde{w})$, the effect of circumstances is totally canceled out, while in the second one, $M(w_i^d)$, only the partial or direct effect is canceled out. Factual and counterfactual outcomes and the decompositions proposed in (7) and (8) are reported in Table 5.

It is suggested that the inequality in PISA 2012 performance results is reduced by approximately 21% when one equals the set of individual opportunities. In turn, the participation of the set of opportunities denoted by eight variables in test performance varies in a stable fashion across quantiles. In other words, unequal opportunities do not seem to change the pattern across the quantiles of the observed distribution.

As to direct effects, they represent around 18%. That is, by canceling out only the direct effects of circumstances, the total inequality of the distribution of scores decreases, on average, 18.7%. These results are noteworthy as the indirect effect of circumstances is given by: $\Theta_I^i = \Theta_I - \Theta_I^d$. That is, there is a positive effect of favorable circumstances on individual effort around 2.3% to 2.8%. Based on these results, two inferences can be made: i) circumstances have an effect on individual effort; and ii) the magnitude of this effect is around 15% of the total of unfair inequality observed on the test.

In other words, the individual effort put in by different sets of opportunities is influenced by the circumstances that the individual is under. Interestingly, the observed effects are stronger in the lower quantile. That is, the worse the performance of the student on the test, the stronger the influence of circumstances on effort.

Nonetheless, the effects of circumstances on effort variables grade retention, perseverance, student migration, actual and potential efforts are likely to be stronger. Unobserved circumstances are likely to account for a larger variance in random residuals v_i in equation (6). In addition, it is reasonable to assume that the decision of some variables denoted here as effort is a parents decision, especially because the students who took the test are all 15 year old. In sum, let us suppose that student migration and retention are decided by the parents and that the other three variables are strongly

Tabela 5: Decomposition of the Joint Participation Opportunities

Inequality Observed Total - Standard deviation of the log			
			0,21457
quantile	25%	50%	75%
Panel A: Total Effect of Circumstances			
Upper Limit	0.16991	0.16978	0.16996
Mean estimate 2a	0.16991	0.16978	0.16996
Lower Limit	0.16991	0.16978	0.16996
% Inequality in Circumstances Test			
Upper Limit	0.20811	0.20874	0.20789
Mean share((1)-(2b))/1)	0.20811	0.20874	0.20789
Lower Limit	0.20811	0.20874	0.20789
Panel B: Direct Effect of Circumstances			
Upper Limit	0.17584	0.17571	0.17595
Mean estimate 2a	0.17531	0.17526	0.17552
Lower Limit	0.17483	0.17484	0.17513
% Direct of Inequality in Circumstances Test			
Upper Limit	0.18048	0.18108	0.17999
Mean share((1)-(2b))/1)	0.18294	0.18319	0.18196
Lower Limit	0.18519	0.18515	0.18380
Panel C: Treating the Condition as Observed Effort			
Upper Limit	0.16991	0.16978	0.16996
Mean estimate 2a	0.16991	0.16978	0.16996
Lower Limit	0.16991	0.16978	0.16996
% Effort and Circumstncias watched in Inequality Test			
Upper Limit	0.20811	0.20874	0.20789
Mean estimate 2a	0.20811	0.20874	0.20789
Lower Limit	0.20811	0.20874	0.20789

Note: Own elaboration from data from PISA 2012. Significant at 95%

influenced by parents instructions. So, these hypotheses sometimes regarded as extreme consider all our variables as observable circumstances.

After considering this, we constructed the counterfactual $w = \exp[\beta_i \overline{W_{p,i}^*} + \overline{E_i} \beta_i + u_i]$. The results reported in panel (D) of Table 5 indicate a total participation rate of 21% in the overall inequality in the score. So, if all variables circumstances and effort were equated, overall inequality would be around 16.9%. Note that the values observed are exactly the same as those for total participation of circumstances in test results.

In another exercise, shown in Table 6, we sought to check the percentage of participation of circumstance variables separately in overall inequality. To do that, we simulated the following reduced-form counterfactual: $w_i^j = \exp[\beta_i \overline{W_{p,i,j}} + W_{p,i,-j} \beta_{-i} + u_i]$, where $M(w_i^j)$ is the counterfactual of the conditional distribution of scores when one of the circumstance variables was kept constant and the others were allowed to vary.

Tabela 6: Equalization of Opportunities

Equalization	Quantil		
	0.25	0.5	0.75
Inequality Total Observed			0.21457
Parental education	0.17468	0.17453	0.17464
Parental occupation	0.18067	0.18071	0.18090
School Type	0.17932	0.17950	0.17994
School Location	0.18035	0.18124	0.17786
Preschool	0.17748	0.17756	0.18062
Gender	0.18125	0.18043	0.18149

Note: Own elaboration from data from PISA 2012. Significant at 95%

Results are consistent with those found in the literature (see, for instance, Ferreira and Veloso (2006). Note that education, type of school, and preschool attendance are the three major determinants of unfair inequality.¹⁴ Specifically, if we equated parental education, type of school, and preschool attendance, the overall inequality in test performance would decline by 17% to 19%. On the other hand, sex is the least relevant circumstance. It is statistically insignificant in many scenarios. In short, sex is not a determinant of inequality in test performance.

In addition, Table 7 summarizes the decomposition results for the influence of circumstances in terms of countries and continents. Note that the indirect effects on individual effort are more significant in Asian and Nordic countries, respectively, and less remarkable in OECD and South American countries, but the magnitude of inequality in test performance is larger in the latter group. In turn, it is necessary to understand which factors are involved in this indirect effect. Aspects such as skill and motivation through intergenerational transmission of parental education could be a decisive factor. Would the fact that these countries historically have a better quality of education explain such difference? According to the literature, the answer is yes, as it is commonly known that three are the factors that can determine intergenerational educational transmission: genetics, parents behavior, and environmental factors.¹⁵ However, the results are inconclusive because, although France has an indirect effect of 0.17, OECD and some of its key members Japan, for example yielded smaller values than those found in South American countries.

Tabela 7: Decomposing the Effect of Circumstances for Countries in PISA

Countries/Region	Inequality	Total effect	Direct effect	% Total effect	% Direct effect	% Indirect effect
South America	0.20151	0.14126	0.15244	0.29897	0.24350	0.05547
Argentina	0.22500	0.13650	0.14835	0.39331	0.34066	0.05265
Brazil	0.18623	0.13574	0.14749	0.27111	0.20804	0.06307
Chile	0.17562	0.12993	0.14094	0.26013	0.19746	0.06267
OECD	0.19225	0.15734	0.16316	0.18158	0.15131	0.03026
Germany	0.18367	0.13295	0.14427	0.27615	0.21453	0.06162
France	0.20298	0.12776	0.16310	0.37060	0.19648	0.17411
Japan	0.18142	0.15335	0.15653	0.15471	0.13720	0.01752
United States	0.17963	0.14907	0.15987	0.17012	0.10997	0.06016
Asian	0.19408	0.14794	0.16909	0.23770	0.12876	0.10894
Singapore	0.18949	0.15955	0.16894	0.15803	0.10846	0.04956
Shanghai	0.14753	0.12284	0.12943	0.16735	0.12267	0.04468
Chinese Taipei	0.18598	0.14631	0.15764	0.21326	0.15238	0.06089
Hong Kong	0.16114	0.13883	0.14718	0.13846	0.08660	0.05186
Nordic	0.19411	0.14483	0.16438	0.25384	0.15316	0.10069
Sweden	0.20620	0.15056	0.16629	0.26982	0.19353	0.07629
Denmark	0.18417	0.13493	0.15264	0.26739	0.17122	0.09617
Finland	0.18858	0.13187	0.15399	0.30071	0.18344	0.11727

Note: Own elaboration from data from PISA 2012. Significant at 95%

It is believed that the indirect effect of circumstances on South American countries is less strong due to several factors, such as quality of parental education and quality of schools, since this characteristic outperforms the effects of parental education, except for Chile. To have an idea, in Brazil and in Argentina, the effect of school on test performance is respectively 30% to 60% stronger than the effect of education. These findings, which correspond to 73.5% of adolescents from these countries who took the PISA 2012, indicate a possible solution to the problem, i.e., it is clearly necessary to ameliorate the quality of schools, especially of public ones.

Moreover, this problem deteriorates when preschool attendance and location of the school are taken into account, as shown in Table A.4. In general, indirect effects are more remarkable in de-

¹⁴See, for instance, Lefranc, Pistolesse and Tranoy (2009).

¹⁵See, for instance, Bjrkklund, Lindahl and Plug (2004).

veloped countries, mainly in France and Finland. On the other hand, where this phenomenon is not observed, as in Japan and South Korea, factors like type of schools, location of schools, and preschool attendance are not statistically significant. That is, these factors are homogenous and, therefore, their indirect effect can be observed on effort.

Following the objectives of this paper, we now verify the effects of an increment in individual socioeconomic set through the ICD method. First, we look at the results of the log estimation of the overall average of individual scores as a function of a dummy for the fathers and mothers education, as well as for family structure, type of school, preschool attendance, and location of school.

The quantile average treatment effect QATE estimated for all circumstance variables except for sex and living with siblings is shown in Table 8. Results were statistically significant at 1%. However, the average gap observed for parental education was around 8%, of which 2% is explained by differences in the level of parental education of the distribution and 6% is due to differences in the average coefficients between children of parents with at least a college degree and children of parents with at most high school education. As to the type of school either public or private the mean difference in the test performance is around 6%, of which 1.5% is related to education in a private school. Note that the share explained by residuals is negligible for all the proposed treatments.

Tabela 8: Quantile Average Treatment Effect Circumstances - QATE

Individual Performance - the overall average log - quantile								
0.10	0.50	0.90	0.10	0.50	0.90	0.10	0.50	0.90
Mother Education			Father Education			Family structure		
0.0251	0.0202	0.0128	0.0247	0.0196	0.0120	0.0240	0.0144	0.0068
0.0651	0.0633	0.0614	0.0592	0.0575	0.0566	0.0367	0.0360	0.0362
0.0807	0.0857	0.0642	0.0709	0.0797	0.0626	0.0603	0.0505	0.0307
School Type			Preschool			School Location		
0.0197	0.0139	0.0091	0.0282	0.0290	0.0233	0.0314	0.0271	0.0170
0.0578	0.0495	0.0451	0.1040	0.1069	0.1103	0.0467	0.0397	0.0365
0.0842	0.0654	0.0324	0.1383	0.1381	0.0996	0.0832	0.0661	0.0568

Note: Own elaboration from data from PISA 2012. Significant at 95%

The average treatment referring to preschool attendance was approximately 14%. However, only 21% of this effect has to do with preschool attendance. In the analyzed database, there is no information about the age at which preschool attendance began, which limits our analysis. According to Spinath et al. (2003), early childhood 0 to 6 years of age plays a pivotal role in the development of the overall cognitive skill.¹⁶ With respect to the other variables family structure and location of schools respectively, the QATE ranged from 5% to 6%. In this case, having a single-parent family and going to a school in the urban area affects the test performance by around 1.5% to 2%, respectively.

Nevertheless, in order to understand the behavior of the pattern of these effects across countries, we estimated the same treatment separately for each country. The results reported in Table A.9 of the Appendix indicate a large variation in the treatment effects across countries. For example, regarding parental education, although there is a similar mean difference across continents between 9% and 12%, the treatment effect of having better educated parents in South America is 7.5 times stronger than that observed in OECD countries, and respectively 3 to 4.5 times that observed for Asian and Nordic countries.

Conversely, in South American countries, the mean difference between treatment and control groups private and public schools is infinitely larger than in Asian countries, and respectively 2 to 3 times greater than that observed in OECD and in Nordic countries. Specifically, this effect is seven times larger on Brazil than on France and 20 times larger than that noticed in the USA. However,

¹⁶According to Carroll (1997), there is a general intelligence factor that makes it easier to deal with information and problems of a certain class or syllabus.

only about one third is explained by the type of school students attend. Furthermore, these effects are quite similar to those found in some OECD, Asian, and Nordic countries. These findings apparently suggest, once again, that the quality of education is poor.

Nonetheless, the treatment effect of preschool attendance is way more significant in countries that had the best performance on PISA. For example, in France and in Japan, the effect of this characteristic is approximately 10%, compared to 4% observed in Argentina and in Brazil. In general, all results indicate that the poor performance, especially of South American countries compared to other countries, is essentially determined by the set of opportunities. Roughly analyzing Table A.4 in the Appendix, it is clear that the observed inequality explained by the set of individual characteristics sum of individual effects amounts to nearly 0.35 compared to 0.11 in OECD countries, 0.15 in Asian countries, and 0.13 in Nordic countries. That is, they correspond to 91% of the total observed in other three continents together.

In this respect, given that the treatment effect is way more significant in these countries, why is the elasticity intergenerational transmission estimated for them lower? Results indicate that in performance variables, instead of educational ones, such as years of schooling, the quality of education is determined. That is, having parents with a college degree is not enough; it is necessary to have educated parents who transfer knowledge and skills to their children. In addition, the role of school in these countries seems to be essential, according to the results. That is, we must do more than educate individuals; we must prepare them to develop and apply their knowledge.

5 Some Remarks

In this paper, we sought to shed some light upon the mechanisms that underlie educational performance. Three were our major goals. First, to verify the level of transmission and intergenerational correlation of education. Second, to decompose the effects that determine intergenerational transmission set of opportunities into direct and indirect. Third, to make a counterfactual analysis by equalizing the socioeconomic circumstances of the students who took PISA 2012. At least three inferences can be made from the results: i) there is low parental education transmission in developing countries, especially in South America; ii) there are indirect effects of circumstances on effort variables, and these effects vary among individuals, countries, and continents. Finally, in line with the low educational transmission by the parents, the treatment effect is substantially stronger in those countries with these characteristics.

Additionally, looking at the results reported in Table A.4, the factors associated with school, except preschool, are significant only in developing countries, specifically in South American countries. On the other hand, except for Chile, the treatment effect of preschool attendance in these economies is usually weaker than that observed in countries with the best performance scores. These findings suggest that preschool attendance does matter; however, we highlight quality (see, for instance, Foguel and Veloso, 2012). If we take a look at Chile, South American country with the best test performance, we perceive that schooling associated with individual socioeconomic background is most important. For other countries of this continent often with the worst performance scores this factor is not that important. In other words, unlike most of the countries in which the best test scores are observed in preschool, in these countries, preschool attendance has a small influence. This further underscores the doubts about the quality of education in these countries.

Referências

- [1] Aguirreche, A. L. (2012). Inequality of opportunity in education. dissertation, Universidade Del Pais Vasco, Biscay, Espanha. Empirical Applications and Policy.

- [2] Ashenfelter, O., Rouse, C. (1997). Income, schooling, and ability: Evidence from a new sample of identical twins (No. w6106). National Bureau of Economic Research.
- [3] de Barros, R. P. (2009). Measuring inequality of opportunities in Latin America and the Caribbean. World Bank Publications.
- [4] Bauer, P., Riphahn, R. T. (2007). Heterogeneity in the intergenerational transmission of educational attainment: evidence from Switzerland on natives and second-generation immigrants. *Journal of Population Economics*, 20(1), 121-148.
- [5] Bourguignon, F., Ferreira, F. H., Menendez, M. (2007). Inequality of opportunity in Brazil. *Review of Income and Wealth*, 53(4), 585-618.
- [6] Bjrklund, A., Lindahl, M., Plug, E. (2006). The origins of intergenerational associations: Lessons from Swedish adoption data. *The Quarterly Journal of Economics*, 999-1028.
- [7] Carroll, J. B. (1997). Psychometrics, intelligence, and public perception. *Intelligence*, 24(1), 25-52.
- [8] Chatterjee, S., Hadi, A. S. (2013). Regression analysis by example. John Wiley Sons.
- [9] Chernozhukov, V., FernandezVal, I., Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, 81(6), 2205-2268.
- [10] Daude, C. (2011). Ascendance by Descendants?: On Intergenerational Education Mobility in Latin America (No. 297). OECD Publishing.
- [11] Ferreira, F., Gignoux, J. (2008). The measurement of inequality of opportunity: theory and an application to Latin America. World Bank Policy Research Working Paper Series, Vol.
- [12] Ferreira, F. H., Gignoux, J. (2011). The measurement of inequality of opportunity: Theory and an application to Latin America. *Review of Income and Wealth*, 57(4), 622-657.
- [13] Ferreira, S., Veloso, F. A. (2003). Intergenerational mobility of education in Brazil. *Research and Economic Planning*, 33(3), 481-513.
- [14] Firmo, M. G., Soares, R. (2008). An analysis of the intergenerational transmission of human capital in Brazil.
- [15] Foguel, M. N., Veloso, F. A. (2014). Inequality of opportunity in daycare and preschool services in Brazil. *The Journal of Economic Inequality*, 12(2), 191-220.
- [16] Gamboa, L. F., Waltenberg, F. D. (2012). Inequality of opportunity for educational achievement in Latin America: Evidence from PISA 2006-2009. *Economics of Education Review*, 31(5), 694-708.
- [17] Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., Verashchagina, A. (2007). The inheritance of educational inequality: International comparisons and fifty-year trends. *The BE Journal of Economic Analysis Policy*, 7(2).
- [18] Lefranc, A., Pistoiesi, N., Trannoy, A. (2009). Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France. *Journal of Public Economics*, 93(11), 1189-1207.
- [19] OECD. PISA. Access 2014, available at OECD: [acer.edu.au / downloads.php](http://acer.edu.au/downloads.php).

- [20] Quintano, C., Castellano, R., Longobardi, S. (2009). A Fuzzy Clustering Approach to Improve the Accuracy of Italian Student Data: An Experimental Procedure to Correct the Impact of Outliers on Assessment Test Scores. Vita e pensiero.
- [21] Roemer, J. E. (1998). Equality of opportunity. Harvard University Press.
- [22] Spinath, F. M., Ronald, A., Harlaar, N., Price, T. S., Plomin, R. (2003). Phenotypic g early in life: On the etiology of general cognitive ability in a large population sample of twin children aged 24 years. *Intelligence*, 31(2), 195-210.
- [23] UNESCO. Report of the United Nations Educational, Scientific and Cultural Organization. 2005.