rbe
REVISTA BRASILEIRA DE ECONOMIA

DOI 10.5935/0034-7140.20200023
ISSN 1806-9134 (online)

FGV
FUNDAÇÃO GETULIO VARGAS

A note on performance differences between urban and rural schools in Brazil*

Luciana de Oliveira Rodrigues†
Edward Martins Costa‡
Vitor Hugo Miro Couto Silva§
Francisca Zilania Mariano¶
Jaime de Jesus Filho⊠

Contents

Keywords
school performance, education inequality, urban-rural differential

JEL Codes
I21, I24, I25

Abstract · **Resumo**

This note analyzes which factors contribute to the performance differential of students attending schools in rural and urban areas in Brazil. Our results show that, in both subjects (Math and Portuguese), students from schools located in urban areas perform better than students from rural area schools. The decomposition exercise shows that the characteristic-effect explain more the urban-rural differential than the return-effect (or structural-effect). Also, the characteristics of the school attended by the students are the major drivers of the difference in grades mainly in the upper quantiles and especially in Math.

*This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES). All ideas, omissions and errors are ours.

†Instituto de Pesquisa e Estratégia Econômica do Ceará. Av. General Afonso Albuquerque Lima, s/n, Cambeba, Fortaleza, CE, CEP 60822-325, Brasil. 0000-0003-1811-405X

‡Universidade Federal do Ceará, Departamento de Economia Agrícola, Programa de Pós-Graduação em Economia Rural (UFC/DEA/PPGER). Avenida Mister Hull 2977, Bloco 826, Campus do Pici, Fortaleza, CE, CEP 60356-000, Brasil. 0000-0002-9187-8534

§Universidade Federal do Ceará, Departamento de Economia Agrícola, Programa de Pós-Graduação em Economia Rural (UFC/DEA/PPGER). Avenida Mister Hull 2977, Bloco 826, Campus do Pici, Fortaleza, CE, CEP 60356-000, Brasil. 0000-0002-5392-8764

¶Universidade Federal do Ceara, (UFC/Sobral). Rua Coronel Estanislau Frota, 563, Bloco I, Centro, Campus de Sobral, Mucambinho, Sobral, CE, CEP 62010-560, Brasil. 0000-0002-1282-8812

⊠Cientista chefe do Tribunal de Contas do Estado do Ceará (TCE-CE). Rua Sena Madureira 1047, Centro, Fortaleza, CE, CEP 60055-080, Brasil. 0000-0003-3646-6585

lrodrigues_s@hotmail.com edwardcost@gmail.com vitormiro@gmail.com zilania@ufc.br jjaimefilho@gmail.com

Received on 17 de maio de 2019
Final Acceptance on 08 de abril de 2020

1. Introduction

According to the literature, there are many factors that can influence students' school performance. They and can be classified into three groups: individual characteristics, family background and school characteristics (Nieto & Ramos, 2014).

Regarding the location of the schools, there are significant differences between urban and rural schools in the indicators of failure rate and dropout rate. This causes expressive variation in school cycles and in the rate of distortion age-grade. The inequalities between rural and urban schools are also significant regarding the students' performance. The students attending urban schools have better results than those from rural schools.

The literature on school effectiveness and factors that contribute to the student learning is extensive, however, there are only few studies that investigates the quality of education considering the location of the school and the student's residence, that is, rural and urban areas. This rules out geographical and structural characteristics of each zone, ignoring, in turn, the problems related to the low performance of the students that are exclusively related to the social context where they live.

As confirmed by S. Soares, Razo, and Farinas (2006), children living in rural areas in Brazil besides of having poor family structure compared to urban children, they also study in poorly equipped schools with low qualifications teachers. This explains an important part of the differences in performance between rural and urban schools.

Analyzing the quality of education in the state of Ceará, Lavor and Arraes (2014) focused on the differences observed between rural and urban schools, especially regarding the availability of various school resources, such as internet access and the library. For the authors, students in rural areas, in addition to registering a higher incidence of child labor, they also attend schools with poor infrastructure, insufficient didactic resources and less skilled teachers.

Menezes-Filho (2007) examines the factors associated with student performance in Brazil. Among the results, as variables that most explain performance, such as family and student resources. On the other hand, the effects of school characteristics are very small. These results reinforce the paper of T. M. Soares (2005).

J. J. Soares Neto, Jesus, Karino, and Andrade (2013) analyzed the infrastructure of schools in Brazil. They classified it into four categories: Elementary, Basic, Adequate and Advanced. Rural schools offered very precarious infrastructure compared to urban ones. While more than 85.2% of urban schools were in the Elementary category, only 18.3% of rural schools were on that category.

On an international perspective, Lounkaew (2013), noted that a large part of the differences between urban and rural students is explained by non-measurable characteristics of schools which vary throughout the percentile of students' performance. Analyzing the performance differential between rural and urban schools in Russia, Amini and Nivorozhkin (2015) found that students' performance varied substantially according to a school location in all subjects, with students from urban areas having the higher scores. Moreover, the study revealed that the individual's and family's characteristics of the students were the factors with the major contribution to the educational gap between urban and rural areas.

In this sense, this note intends to shed some light on the major drives of the differences in performance of students in urban and rural schools in Brazil. The main questions we tried to answer were: What are the determinants of performance differential? Are individual characteristics more important than the school structure? Does the teaches quality matter? What is the role of the students' family background? Trying to answer these questions, we use the technical approach proposed by Firpo, Fortin, and Lemieux (2007). We use data on students' performance in Math and Portuguese from the exam *Prova Brasil* 2015. Only students in 5th grade of the PS are considered in our sample.

From the results, we observe that there is no well-defined pattern across the distribution of the scores in both exams. Students in rural schools have a bigger return on their characteristics but have a lower level of the same characteristics. The same happens in the math exam in the lowest quantile. Furthermore, in the lower half of the distribution, the marginal contribution of characteristics of families and schools has a greater weight in explaining the differences in performance. Lastly, in general, at the top of the distribution, the differences between students from rural and urban schools are less explained by the returns of the characteristics considered.

This note is structured as follows: the next section presents the empirical strategy with a description of the data and the treatment of the variables used in the analysis. The fourth section presents the results. Finally, in the last section, we have the final considerations of the analysis.

2. Dataset and variables

The information used in this study was obtained from the data provided by INEP. The performance in the standardized tests on Portuguese and Math, and socioeconomic information of the students and their families were obtained from the database of 2015 Prova Brasil exam[1]. Additional data on school infrastructure, teachers and number of enrollments were obtained from the 2015 School Census and Educational indicators[2] extracted directly from INEP's website.

[1] Available at http://portal.inep.gov.br/basica-levantamentos-acessar

[2] More information at http://portal.inep.gov.br/indicadores-educacionais

For the purposes here, we use information only of students in the 5th grade of the Primary School attending public schools (state or county).[3] One of the main justifications for this choice is that this a group age in which children are still very dependent on the parents or caregivers and the quality of the education offered to them can be decisive in their schooling path. Table 1 summarizes the variables included in our model.

Based on studies consolidated in the literature, differences in school performance among those attending rural and urban schools are estimated by an Educational Production Function (EPF) that uses several inputs, including observable and unobservable characteristics of students, their families and the school features they attend.[4]

Given that the distribution of students' scores is not uniform the best strategy would be to use a statistic different from the average differential score to perform de decomposition exercise. We can obtain information using the entire differential performance distribution and assessing differences by quantiles. Thus, to decompose the differential into its determinants, we adopted the approach proposed by Firpo et al. (2007); Firpo, Fortin, and Lemieux (2009), which estimates unconditional quantile regressions based on the concept of recent influence function (RIF)[5] and generalizes the decomposition of Oaxaca (1973) applied the quantiles.

3. Results

3.1 Data descriptive analysis

Descriptive statistics[6] show that our sample consists of 784,120 students in the 5th grade of Primary School from public schools. From the total, 710,680 students were attending urban schools and 73,440 attending schools in rural areas.

Looking at the students' scores, in the urban area the average grade in Portuguese is 207.84 points and in Math it is 220.72 points. In rural areas, the average is much lower: the average grade in Portuguese is 183.77 points, a difference of more than 24 points in favor of schools located in urban areas. In Math, the average score is 198.74 points, a difference of approximately 22 points lower than the average of urban schools' students. Figure 1 displays the estimated density of the students' scores at different locations.

---

[3] Federal and private schools represent less than 1% of schools at rural areas. Thus, they were left out of our sample.

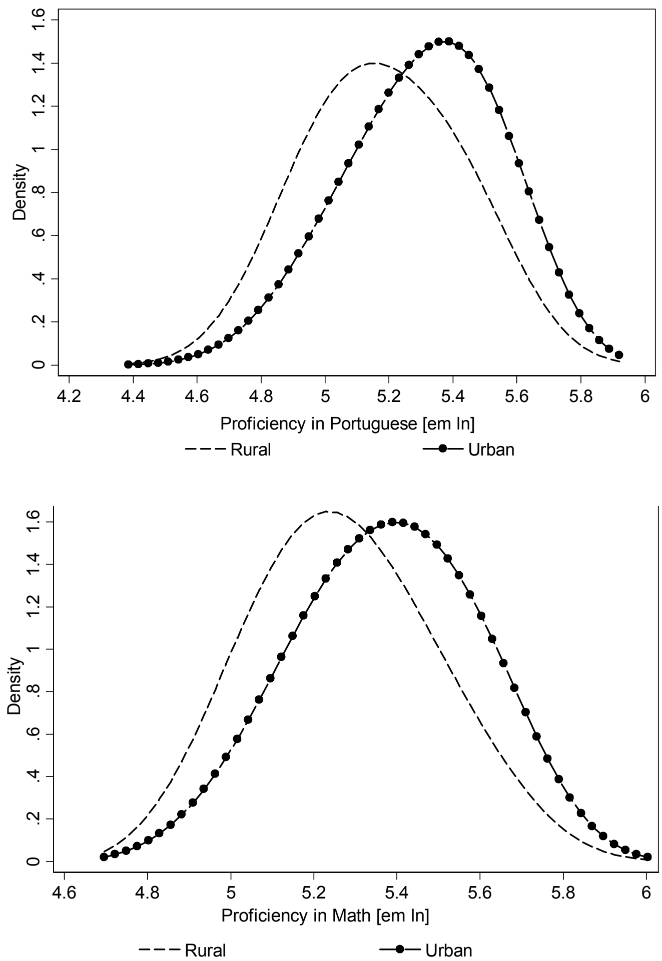
[4] See Hanushek and Woessmann (2011, 2012) and WößMann (2003) for a more detailed information about the Educational Production Function.

[5] See Apêndice.

[6] Table 02 (Descriptive Statistics – Rural and Urban – Brazil, 2015), can be requested from the authors.

Table 1. Variables Description.

	Variables	**Definition**
Dependent Variables	grade_por	Score on Portuguese (SAEB/97 scale)
	grade_mat	Score on Math (SAEB/97 scale)
	ln_por	ln of Portuguese score
	ln_mat	ln of Math score
Student's Characteristics	Age	Student age
	gender	1 – Male; 0 – Female
	Race	1 – White; 0 – non-White
	never_fail	1 – Never failed a grade; 0 – otherwise
	never_drop	1 – Never dropout; 0 – otherwise
	right_age	Correct age for the 5th grade? (1 – yes; 0 – no)
	reading	Reading habits (1 – reads frequently; 0 – no)
	comp	Has a computer at home? (1 – yes; 0 – no)
	homework1	Always do Portuguese homework? (1 – yes; 0 – no)
	homework2	Always do Math homework? (1 – yes; 0 – no)
	job	Has a job? (1 – yes; 0 – no)
Family Background	FES	Family Economic Status (FES)
	Eduf1	Father illiterate or less than Primary School (1 – yes; 0 – no)
	Eduf2	Father with Primary School (1 – yes; 0 – no)
	Eduf3	Father with High School (1 – yes; 0 – no)
	Eduf4	Father with college degree (1 – yes; 0 – no)
	Edum1	Mother illiterate or less than Primary School (1 – yes; 0 – no)
	Edum2	Mother with Primary School (1 – yes; 0 – no)
	Edum3	Mother with High School (1 – yes; 0 – no)
	Edum4	Mother with college degree (1 – yes; 0 – no)
	Both	Both parents at home? (1 - yes; 0- no)
	Incentive	Parents encourages going to school? (1 – yes; 0 – no)
Teachers' Characteristics	College	Proportion of teachers with college degree
	Experience	Proportion of teachers with more than 6 years of experience
	Wage	Prop of teachers earning more than R$ 3,152.01 (4 minimum wages of 2015)
Schools' Characteristics	Urban	School location (1 – Urban; 0 – Rural)
	Gov	School administration responsibility (1 – County; 0 – State)
	Size	Number of enrollments (all grades)
	student_teac	Ratio Students/Teachers (1^{st} to 5^{th} grade)
	Duration	Average class duration (in minutes)
	material	Lack of school supplies? (1 – yes; 0 – no)
	Director	How is the Principal selected? (1 – appointed/elected; 0 – otherwise)
	Teachers	Insufficient number of teachers? (1 – yes; 0 – no)
	ITR	Indicator of Teachers Regularity
	ISI	Index of School Infrastructure
	IBMS	Indicator of Bad Maintenance of the School



Source: Authors' calculations. Data from the exam Prova Brasil/Saeb, 2015.

Figure 1. Estimate Density (log scores in Portuguese and Math).

Regarding to the failure rate, 62% of the student from rural schools have never failed in any subject, against 72% in urban schools. The dropout rate is also higher among students in rural areas. Another important difference is that the proportion of students working outside the household is eight percentage point higher in rural areas, 22% against 14%.

The indicator Family Economic Status (FES) shows that students from urban areas come from households with better financial conditions than students in rural zone. Also, parents of students in rural schools have less years of schooling than urban students. Thus, in both dimensions, family background and economic status, students in rural zone are in a disadvantage condition.

Teachers from urban schools are better paid than their rural peers. Also, in urban schools 83% of the teachers have college degree, against 61% in rural schools. Moreover, urban schools have more experienced teachers than the rural ones.

Finally, looking at schools' characteristics, we notice that on average the urban schools are bigger than the rural ones, with more enrolled students. As consequence, the rural schools have on average less students per teacher. Further, urban schools have better general infrastructure (ISI).

3.1.1 Decomposition of the school performance differentials

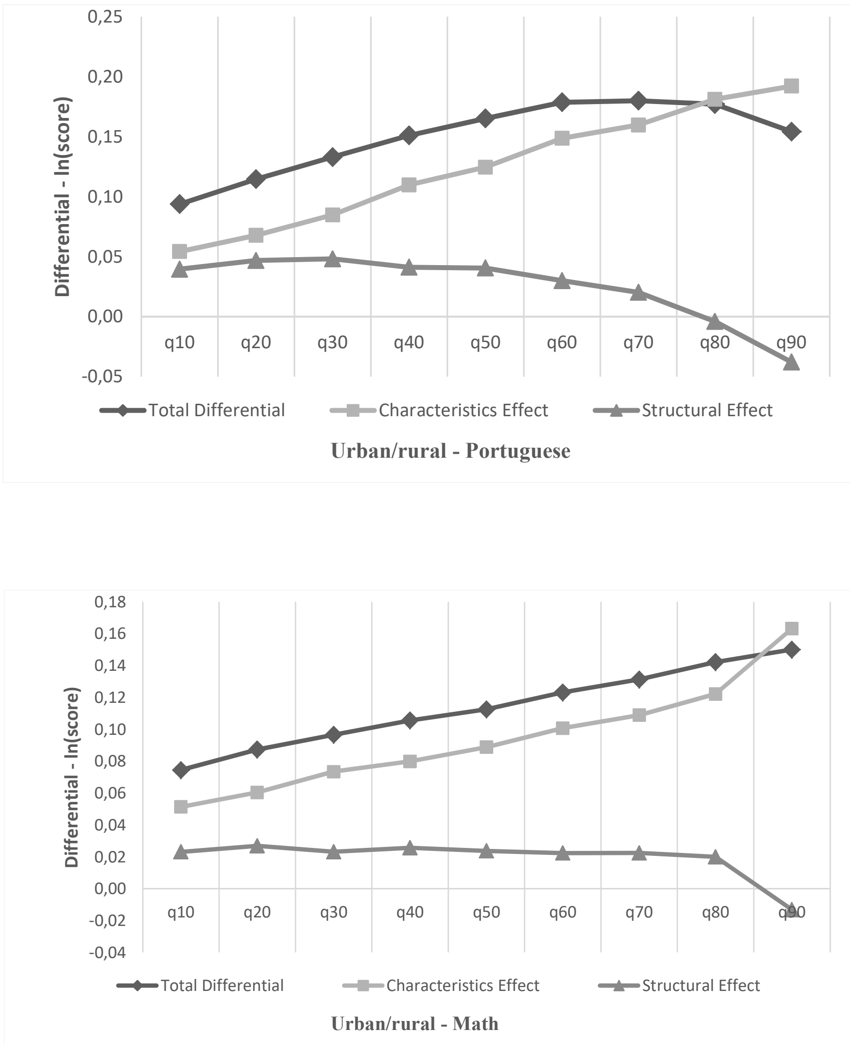
According to our methodology the first step is to estimate the unconditional quantile regression for different quantiles of the score's distribution. Detailed results for the 10^{th}, 50^{th} and 90^{th} quantiles can be seen in Tables A1 and A2 which can be requested from the authors. For comparison, we also estimate an EPF for the means of the scores.

As expected, the estimated effect has huge variability throughout the distribution, suggesting the quantile approach is suitable. Also, the results are different for students in urban and rural schools. After the estimation of the unconditional quantile regressions we applied the Oaxaca–Blinder method to decompose the school performance differential between urban and rural schools. Figure 2 shows the performance differential in Portuguese and Math in terms of characteristic and structural effects. One can notice that the biggest part of the differential is explained by the characteristic effect, with urban students always performing better than the rural ones.

Looking at the Portuguese exam, the performance differential is increasing till the 70^{th} quantile, when the structural effect becomes negative. In Math, the differential increases monotonically with the quantiles. Table A3, which can be requested from the authors, presents the decomposition results for 9 quantiles of the scores distribution. In all cases the differential between urban and rural students is significant at 1% level. Also, for the top quantiles, 80^{th} and 90^{th}, the structural effect is negative meaning that rural students would perform better than urban student if they had equal characteristics.

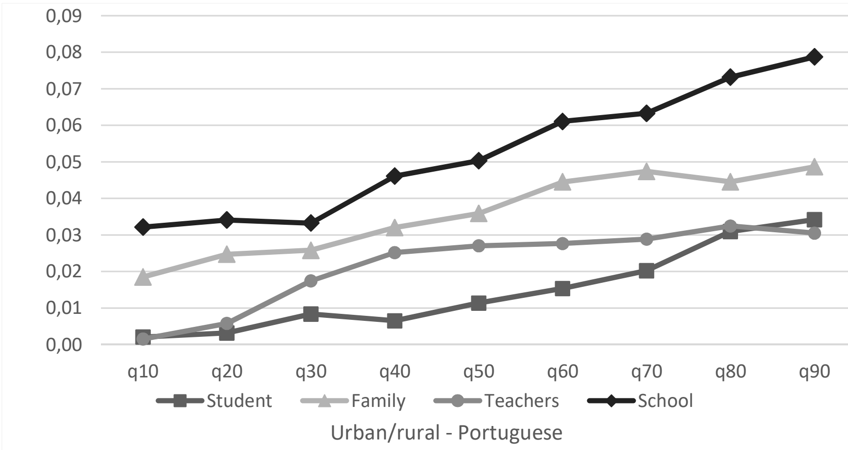
We then decompose the characteristics and structural (return) effects into different factors. These factors are the explanatory variables on our EPF and are grouped as: student profile, family background, teachers' profile, and school profile. Figure 3 presents the results for the characteristics effect. The results show that the school factor explain most of the characteristic differentials. For the 10^{th} quantile of the score's distribution, the school profile explains 43% and 59% of the characteristic differential in Math and Portuguese respectively. Family background explain around 28% and 34% in the same subjects for the same quantile.

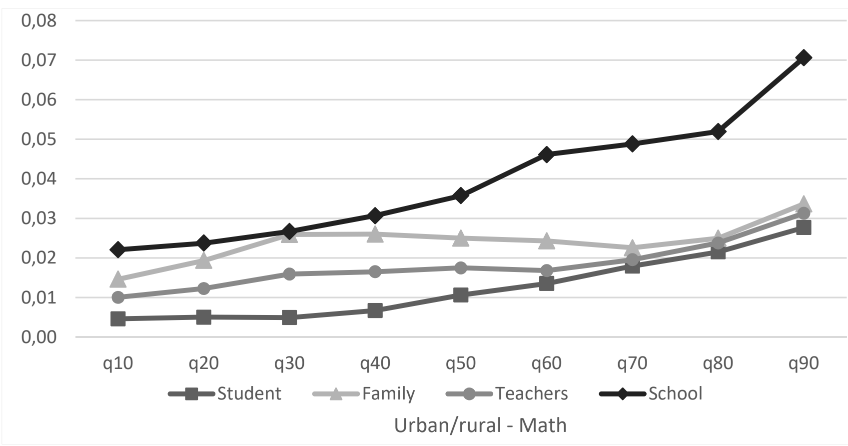
For the median student, approximately 9% and 12% of the characteristic effect is explained by individual profile, 29% e 28% by family background, 22% and 20%



Source: Authors' calculations.
Note: Significant at 1%.

Figure 2. School performance differential – Decomposition (*ln* of scores in Portuguese and Math).





Source: Authors' calculations. Data from Censo da Educação Básica–INEP/MEC and Prova Brasil/Saeb 2015.

Figure 3. Characteristics Effect – Decomposition – Urban-Rural.

by teacher's profile and 40% by the school structure respectively in Portuguese and Math. The student profile is important only for the exams top performers. In the 90th quantile, student profile accounts for 18% and 17% of the effect in Portuguese and Math (to see results for other quantiles, request the authors for table A4).

Figure 4 shows the decomposition of the structural or return effect. This is the effect relative to the estimate coefficients. There is no well-defined pattern across the distribution of the scores in both exams. For some quantiles the effect is not statistically significant. It's important to notice that for the lower quantiles in the Portuguese exam the differential relative to individual characteristics is negative. Thus, students in rural schools have a bigger return on their characteristics but have a lower level of the same characteristics. The same happens in the math exam in the lowest quantile.

In the case of the coefficients associated with the characteristics of the families, teachers and schools, they were positive and contributed to the observed differential, but this contribution becomes less important throughout the higher performance strata. That is, in the lower half of the distribution, the marginal contribution of characteristics of families and schools has a greater weight in explaining the differences in performance.

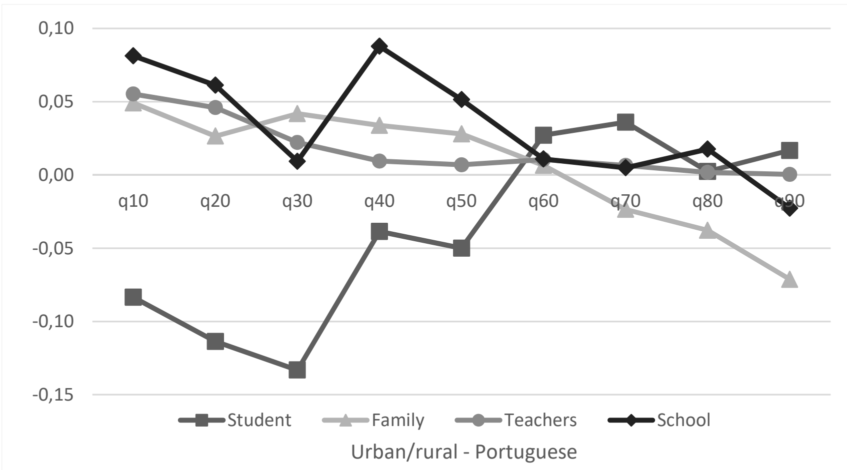
In general, at the top of the distribution, the differences between students from rural and urban schools are less explained by the returns of the characteristics considered. This may reflect a change in the composition of students, which operates to reduce differences between groups, especially among those students with greater cognitive abilities in the rural and urban areas.

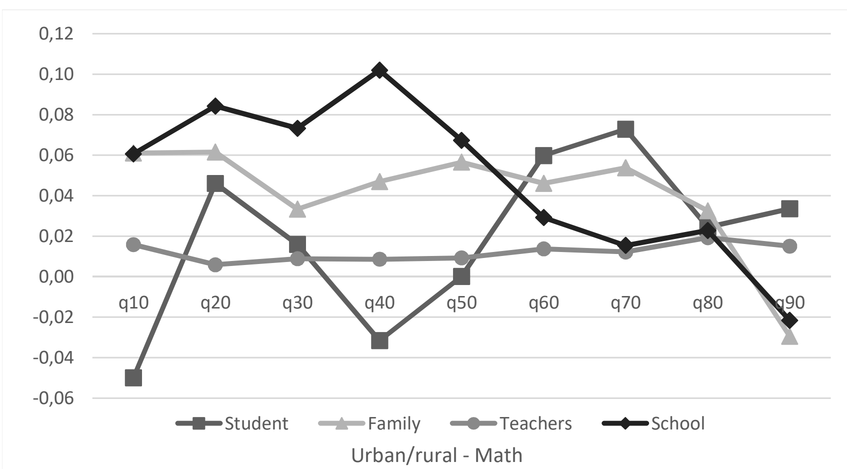
4. Final comments

Unconditional quantile regressions estimate of the educational production function at the student level point out that the contributions of the characteristics of students, family, teachers and school are not constant throughout the distribution of grades in the two exams (Portuguese and Math). The performance gap between the students in the two zones is also statistically significant in both exams. The results of the decomposition also show that, in both tests, a large part of the performance differential comes from the characteristics effect.

In addition, the decomposition exercises by quantiles revealed the increasing role of the characteristic effect, that is, the higher the performance in an exam the more important are the characteristics in explaining the educational gap between groups. The structural effect (unobservable factors), despite of having relatively low weight, also contributes to the increase of the performance differential and cannot be ignored, except for the higher quantiles.

Regarding the implementation of public policies aiming to reduce disparities between rural and urban school students, policymakers should consider that





Source: Authors' calculations. Data from Censo da Educação Básica–INEP/MEC and Prova Brasil/Saeb 2015.

Figure 4. Structural Effect – Decomposition – Urban-Rural.

asymmetric effects of student, family, teacher, and school characteristics on the quantiles of performance require a differentiated approach among students, where educational improvement initiatives must consider differences in the socioeconomic composition of students. In addition, because they play a significant role in the performance differential, teachers need to be qualified and well-paid, and schools must have a good infrastructure, especially when the family background is poor.

Finally, financial investments alone do not guarantee quality improvement and educational equity in Brazil. Initiatives to improve non-measurable aspects of schools (such as parental involvement, encouragement for students to attend the library and others) deserve attention and are equally important. Also, the possible success of a good educational policy to deal with inequality and improve the quality of public schools depends on finding the right balance between financial investment and the development of a school environment that benefits the learning process of children in social vulnerability condition.

References

Amini, C., & Nivorozhkin, E. (2015). The urban–rural divide in educational outcomes: Evidence from Russia. *International Journal of Educational Development*, *44*, 118–133. http://dx.doi.org/10.1016/j.ijedudev.2015.07.006

Firpo, S., Fortin, N., & Lemieux, T. (2007). *Decomposing wage distributions using recentered influence functions regressions.* Vancouver.

Firpo, S., Fortin, N., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953–973. http://dx.doi.org/10.3982/ECTA6822

Hanushek, E. A., & Woessmann, L. (2011). The economics of international differences in educational achievement. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbooks in economics of education* (Vol. 3, pp. 89–200). Amsterdam.

Hanushek, E. A., & Woessmann, L. (2012). Schooling, educational achievement, and the Latin American growth puzzle. *Journal of Development Economics*, *99*(2), 497–512. http://dx.doi.org/10.1016/j.jdeveco.2012.06.004

Lavor, D. C., & Arraes, R. d. A. d. (2014). Qualidade da educação básica e uma avaliação de política educacional para o Ceará. In *X Encontro Economia do Ceará em Debate, 2014*, Fortaleza.

Lounkaew, K. (2013). Explaining urban–rural differences in educational achievement in Thailand: Evidence from PISA literacy data. *Economics of Education Review*, *37*, 213–225. http://dx.doi.org/10.1016/j.econedurev.2013.09.003

Menezes-Filho, N. (2007). *Os determinantes do desempenho escolar no Brasil* [Sumário Executivo]. Instituto Futuro Brasil, IBMEC São Paulo e Faculdade de Economia e Administração da Universidade de São Paulo.

Nieto, S., & Ramos, R. (2014). *Decomposition of differences in PISA results in middle income countries* (Working Paper 2014/08). Barcelona: Institut de Recerca em Economia Aplicada Regional i Pública. http://www.ub.edu/irea/working_papers/2014/201408.pdf

Oaxaca, R. L. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, *14*(3), 693–709. http://dx.doi.org/10.2307/2525981

Soares, J. J., Neto, Jesus, G. R. d., Karino, C. A., & Andrade, D. F. d. (2013). Uma escala para medir a infraestrutura escolar. *Estudos em Avaliação Educacional*, *24*(54), 78–99. http://dx.doi.org/10.18222/eae245420131903

Soares, S., Razo, R., & Farinas, M. (2006). Perfil estatístico da educação rural: Origem socioeconômica desfavorecida, insumos escolares deficientes e resultados inaceitáveis. In A. M. Bof (Ed.), *A educação no Brasil rural.* Brasília: INEP.

Soares, T. M. (2005). Modelo de três níveis hierárquicos para a proficiência dos alunos de 4ª série avaliados no teste de língua portuguesa do SIMAVE/PROEB-2002. *Revista Brasileira de Educação*(29), 73–87. http://dx.doi.org/10.1590/S1413-24782005000200007

WößMann, L. (2003). Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics*, *65*(2), 117–170. http://dx.doi.org/10.1111/1468-0084.00045

Apêndice.

Recentered Influence Function (RIF) regression and decomposition method

Let Y be the score of a student at the exam Prova Brasil, q_τ is the τ-th quantile value and X is the set of explanatory variables, that includes individual characteristics, family background, characteristics of the school and a measure of the teacher's quality. Applying to the quantile function we get

$$\mathbb{E}\left[RIF(Y, q_\tau) \,|\, X\right] = X\beta_\tau \,. \tag{A-1}$$

Coefficients β_τ are approximations of the marginal effects of each explanatory variable on unconditional quantile q_τ.

Having the estimates of the EPF for each area ($k =$ rural e urban) using the method described above, we can then decompose the students' performance differential using the traditional Oaxaca–Blinder technic. Assuming the model is linear and the expected value of RIF, for a given quantile τ we have

$$\begin{aligned}\hat{\Delta}^{q_\tau} &= \hat{\beta}_{\tau,\text{urban}}\left[\bar{X}_{\text{urban}} - \bar{X}_{\text{rural}}\right] + \bar{X}_{\text{rural}}\left[\hat{\beta}_{\tau,\text{urban}} - \hat{\beta}_{\tau,\text{rural}}\right] \\ &= \hat{\Delta}_X^{q_\tau} + \hat{\Delta}_S^{q_\tau} \,.\end{aligned} \tag{A-2}$$

The first term, $\hat{\Delta}_X^{q_\tau}$, is the characteristic effect, that captures the effect of differences in the observed characteristics. The second, $\hat{\Delta}_S^{q_\tau}$, is the structural effect, capturing differences on the returns (estimated coefficients) of each characteristic of each group.