ASSESSING THE GENDER GAP ON ACADEMIC ACHIEVEMENT IN BRAZIL

Eduardo Tillmann[1]; Flavio V. Comim[2]

Área 2: Desenvolvimento Econômico

Abstract

This research examines the factors associated with gender inequality on 5th and 9th grade student's scholastic achievement in Brazil, exploring students', teachers' and schools' characteristics in two different types of decomposition methods, one that explores differences in mean achievement and another that assesses the entire test score distribution. The results indicate boys and girls have very distinct educational production functions. Moreover, despite both genders having similar family and socioeconomic characteristics, the main contributor towards the learning differences is the return of these characteristics in terms of test scores for each gender, which, therefore, reduces the role of teachers' and schools in diminishing these inequalities. On the policy perspective, our study findings indicate that a more efficient policy aiming at the reduction of test score differences between boys and girls should look beyond resources.

Keywords: Gender inequality. Grade decomposition. Academic achievement in Brazil.

Resumo

Este artigo examina os fatores associados com a desigualdade de gênero no desempenho escolar de alunos do 5º e 9º ano no Brasil, explorando características dos alunos, professores e escolas através de dois métodos diferentes de decomposição, um que explora a diferença de média dos resultados e outro que analisa toda a distribuição de notas. Os resultados indicam que meninos e meninas têm funções de produção educacionais distintas. Isto é, apesar de ambos gêneros possuírem características de contexto socioeconômico e familiares parecidas, o principal fator contribuinte para as diferenças de aprendizado está no retorno destas características em termos de nota para cada gênero, o que, portanto, reduz o papel do professor e da escola na diminuição destas desigualdades. Em termos de políticas públicas, os resultados indicam que uma política mais eficiente voltada para a redução da diferença de notas entre meninos e meninas deve considerar questões além de recursos.

Palavras-chave: Desigualdade de gênero. Decomposição de notas. Desempenho escolar no Brasil.

JEL Classification: I21. I24. J16.

[1] Federal University of Rio Grande do Sul (FURG). eduardotillmann@furg.br
[2] IQS School of Management, Ramon Llull University and Centre of Development Studies, University of Cambridge. fvc1001@cam.ac.uk

1 Introduction

Gender inequality is a reality for most of the developed and developing world. Women tend to be disadvantaged in a wide variety of achievements and opportunities, from health and education to jobs and wages (WORLD BANK, 2012). Gender unequal societies are less humane, less efficient and with lower potential for economic growth (SEN, 2009; SCHULTZ, 2002 AND KLASEN, 2002).

In the Brazilian context, the theme of gender inequality has been explored in a wide range of aspects, such as wage differentials (KASSOUF, 1998; MADALOZZO, 2009), occupational segregation (RASCHE JÚNIOR, 2015; MADALOZZO; ARTES, 2017), and on the decision between work and study (TILLMANN; COMIM, 2016). However, few studies have investigated gender differences on academic achievement.

Recent studies have argued for the inexistence of gender gaps in school achievement at earlier ages, but that test scores differences arise and intensify as children get older (FRYER JR; LEVITT, 2010). This, in turn, suggests that these gaps may be related to pedagogical and cultural differences in the learning process that might be associated with observable educational characteristics and not to any intrinsic distinct characteristics between boys and girls.

In order to assess this hypothesis, this paper applies the Oaxaca-Blinder decomposition technique (OAXACA, 1973; BLINDER, 1973) to measure how much of the mean gap between girls and boys depends on differences in endowments, and how much can be attributable to differences in the return of these endowments, namely, in their unexplained component. An extension of this method is the Recentered Influence Function (RIF) proposed by Fortin et al. (2011) and Firpo et al. (2009) which allows us to study the contribution of specific covariates at different parts of the test scores distributions.

Decomposition methods[1] are useful to explain the distribution of test scores by a set of factors that are correlated with educational achievement. Gender gaps are decomposed into a part that is due to gender differences in the magnitudes of the determinants of the test score in question, and a part that is due to gender differences on the effects of these determinants. By doing so it is possible to find out whether the gap is related to differences between boys and girls in resources, or whether they are due to differences in the way boys and girls use these resources for school success. This can be an important guide for education policies, because a policy that aims to diminish inequality between the two genders might have a limited effect on raising performance if it focuses only on the distribution of resources, and not on their use.

Given this context, this paper investigates the factors associated with gender differentials in Math and Portuguese academic achievement in Brazil. We use the two already mentioned types of decomposition methods and apply them to a pool of four cohorts of 5th graders, and to a pool of 9th graders. Thus, we are able to examine the gender test score gap for two distinct age patterns and educational levels. The idea is to assess students, teachers and schools characteristics in order to identify the sources of these gaps, and to verify if there is any difference between the two grades.

This paper contributes to the existing literature, first, by exploring the factors associated with differences on gender achievement in Brazil. Also, because we employ mean and quantile decomposition analysis, we are able to identify the covariates contributing to the gender gap at various points in the test scores distribution.

The main result of this paper is that boys outperform girls in math, an advantage that reverses in literacy. This happens because boys and girls have very distinct educational production functions. Indeed, their background characteristics are differently translated into academic achievement, as it can be seen by the unexplained or coefficient component of the Oaxaca-Blinder decomposition. Moreover, by assessing the whole distribution of test scores, we verify that in all percentiles boys and girls tend to perform relatively more similar in Math than in Portuguese, and on 5th grade than on 9th. In sum, the results show that regardless of the distributional statistic, the majority of the gender difference is related

[1] Two descriptions of decomposition methods can be found on Firpo (2017) and O'Donnel et al. (2008).

to how boys and girls translate their background characteristics into achievement, and not to their characteristics in itself.

The rest of the paper is structured as follows. The next section presents our dataset and describes the two decomposition techniques we use. Section 3 presents the results, separated by each method, followed by the concluding remarks.

2 Data and Methods

3.1 Data

This paper uses four cohorts of the Brazilian Ministry of Education assessment of Math and Portuguese learning for 5th graders and, also, for 9th graders in public schools (2009, 2011, 2013 and 2015). This assessment is biannual, and includes questionnaires for Principals, Teachers, Students and School characteristics. The students' questionnaire includes detailed socioeconomic information and a Math and Portuguese test. The Teachers' questionnaire, besides having their background information, has also questions regarding their contract with the school, and other work related questions. We also match the dataset from these four assessments to the corresponding Brazilian Education Census, in order to obtain data on school infrastructure.

The variables we use on our estimates are based on other studies that try to identify student school success in Brazil, such as Franco and Menezes-Filho (2008), Biondi and Felício (2007), Soares and Candian (2007) and Soares et al. (2016). Therefore, the students characteristics[2] we are interested in are race, if the student has the expected age for this grade, lives with his mom and dad, work, spends more than one hour of domestic tasks, attended kindergarten, lives in an urban area, doing the subjects' homework and if at least one of the parents have a graduate degree. We also include two indexes[3], a socioeconomic status index (SES), and an index for family involvement with the student. The school characteristics are if at the beginning of the school year the students had their books and an infrastructure quality index. The teachers characteristics are sex, graduate degree, work on other activity besides teaching, has more than 10 years of experience, works more than 40hrs a week and if he or she have a permanent teaching contract. We also include year dummies and for each of the five Brazilian regions in our estimations.

Table 2, below, presents descriptive statistics of the number of students and test scores for 5th and 9th graders.

Table 1 - Descriptive Statistics of sample Math and Portuguese test scores.

		Students	Mean	Std Dev	10th	50th	90th
5th grade Math	Boys	1994489	221.42	48.50	159.34	219.42	286.62
	Girls	2075855	215.82	46.06	158.30	212.89	278.07
5th grade Portuguese	Boys	1999704	198.03	46.51	139.09	195.17	260.91
	Girls	2081876	206.37	46.63	146.09	204.85	268.60
9th grade Math	Boys	1640355	253.71	46.64	191.75	254.00	313.70
	Girls	1933266	244.70	44.40	187.63	243.04	303.09
9th grade Portuguese	Boys	1628734	238.68	47.59	174.94	238.83	301.38
	Girls	1920328	251.32	45.32	190.61	252.43	309.87

Source: Elaborated by the authors using SAEB (2009-2015).

As we can see from Table 2, the number of girls and boys are fairly well distributed in 5th grade, and there are a slight higher number of girls at 9th grade. Also, girls tend to perform worse than boys in

[2] Table 1A in the Appendix contains all variables we use in this study and their description.

[3] Both indexes are built from the first component of Principal Component Analysis, a procedure commonly used as a dimensionality reduction technique.

Math, a difference that is greater at the highest grade. However, girls tend to outperform boys on language, a difference that is also increasing with the grade. All the percentiles tend to reproduce these patterns, the exception is the lowest percentile in Math for 5th graders, where the gender difference is almost negligible.

3.2 Empirical Specifications

But how do the characteristics of boys and girls influence their Math and Portuguese performance? In addition, is there any meaningful difference between younger and older students? These questions can be answered first by employing the traditional Oaxaca-Blinder decomposition (OAXACA, 1973; BLINDER, 19743), a counterfactual decomposition technique, which estimates a linear education production function for each gender, such as:

$$Y_{ij} = \alpha_{j0} + \sum_{k=1}^{K} X_{ik}\beta_{jk} + e_{ji} \qquad (1)$$

Where Y_i is the Math of Portuguese tests scores of each *i* individual and *j* gender, boys or girls, and, as discussed above, X_{ik} is the vector of the K covariates used in this paper. Therefore, the overall mean difference in outcomes between boys and girls can be written as:

$$E[Y_{ib}/X_{ib}] - E[Y_{ig}/X_{ig}] = (\hat{\alpha}_b - \hat{\alpha}_g) + \left(\sum_{k=1}^{K} \bar{X}_{bk}\hat{\beta}_{bk} - \sum_{k=1}^{K} \bar{X}_{gk}\hat{\beta}_{gk}\right) = \Delta \qquad (2)$$

If we add and subtract $\sum_{k=1}^{K} \bar{X}_{bk}\,\hat{\beta}_{gk}$, Δ can be rewritten as a sum of two different components:

$$\Delta = \sum_{k=1}^{K} (\bar{X}_{bk} - \bar{X}_{gk})\,\hat{\beta}_{gk} + \left[(\hat{\alpha}_b - \hat{\alpha}_g) + \sum_{k=1}^{K} \bar{X}_{bk}\,(\hat{\beta}_{bk} - \hat{\beta}_{gk})\right] \qquad (3)$$

The first term on the right hand side of the equation is called the characteristics or explained effect, as it portrays the differences on the distributions of the covariates between the two genders. The second term is called the coefficient or unexplained effect, since it reflects differences on the returns of each characteristic, between boys and girls. Both of these elements rely on $\sum_{k=1}^{K} \bar{X}_{bk}\,\hat{\beta}_{gk}$, a counterfactual estimate for boys, namely, how boys would score on their Math and Portuguese tests if they had the same estimated coefficients as girls.

One issue regarding this technique is due to the identification of the contribution of categorical variables. The problem is that the choice of the reference group affects the decomposition results[4]. We are able to contour this issue by using normalized regressions, a method proposed by Yun (2005) that consists in expressing the estimations as deviations from the mean.

Another limitation of the Oaxaca-Blinder decomposition, as discussed by Barsky et al. (2002), is that they may not provide consistent estimates of the coefficients and characteristics effect when the conditional mean is a non-linear function. This happens because the counterfactual mean will not be equal to $\sum_{k=1}^{K} \bar{X}_{bk}\,\hat{\beta}_{gk}$. One possible solution to this problem is to use a reweighting approach as in DiNardo, Fortin and Lemieux (1996) that reweight the sample of boys so that the distribution of their characteristics is similar to that of girls, using the following reweighting function:

[4] For more information see Jann (2008).

$$\psi(X) = \frac{P(g = 1/X)/P(g = 1)}{P(g = 0/X)/P(g = 0)} \tag{4}$$

Where $P(g = 1/X)$ represents the probability of a student being a girl and, $P(g = 0)$ and $P(g = 1)$, are, respectively, the samples proportions of boys and girls. The reweighting factor is then applied to the boys sample to calculate the counterfactual test score distribution.

Firpo et al. (2009), Fortin et al. (2011) and Chernozhukov et al. (2013) expand the Oaxaca-Blinder decomposition of the means and provide a comprehensive approach to study the entire distribution function. Their method consists of two stages. First, they transform the outcome variable, using influence functions, in order to obtain what they call Recentered Influence Function (RIF), which can be computed for most distributional statistics and, therefore, allows their decomposition to assess quantiles, variance, and even inequality measures. The second stage uses these estimates to generate Oaxaca-Blinder decomposition for the measures of interest.

The approach, in its simplest form, assumes that the conditional expectation of the $RIF(Y; Q_t)$ can be modelled as a linear function of the explanatory variables:

$$E[RIF(Y; Q_\tau)/X] = X\gamma + \varepsilon \tag{5}$$

Nonetheless, the RIF function can be defined as:

$$RIF(Y; Q_\tau) = Q_\tau + \frac{\tau - \boldsymbol{I}(Y \leq Q_\tau)}{f_Y(Q_\tau)} \tag{6}$$

Where $\boldsymbol{I}(Y \leq Q_\tau)$ is an indicator variable for whether the outcome variable is smaller or equal to the quantile Q_τ. Running a linear regression of $\boldsymbol{I}(Y \leq Q_\tau)$ on X is a distributional regression estimated at $y = Q_\tau$.

According to Fipo, et al. (2009) to run regressions of the RIF on the vector of covariates one should plug in the estimates of the sample quantile, $\hat{Q}_\tau$, and of the density at that point, $\hat{f}(\hat{Q}_\tau)$, into Equation (6). Letting the coefficients of the unconditional quantile regressions for each group be:

$$\hat{\gamma}_{g,\tau} = \left(\sum_{i \in J} X_i X_i^T\right)^{-1} \sum_{i \in J} \widehat{RIF}\left(N_{ji}; Q_{j,\tau}\right) X_i \tag{7}$$

We can write the equivalent of the Oaxaca-Blinder decomposition for any unconditional quantiles as:

$$\hat{\Delta} = \bar{X}_b\left(\hat{\gamma}_{b,\tau} - \hat{\gamma}_{g,\tau}\right) + \left(\bar{X}_b - \bar{X}_g\right)\hat{\gamma}_{g,\tau} \tag{8}$$

The right hand side of Equation (8) displays the coefficient and the characteristics effect. The identification of these effects by quantile, allow us to analyze, with greater detail the influence of the students, schools, teachers and families in Math and Portuguese achievement. However, RIF regressions are also subject to the two issues regarding the Oaxaca-Blinder decomposition, namely, the identification of categorical variables and the nonlinearity of the conditional expectation. Hence, we also apply normalized regressions and Dinardo, Fortin and Lemieux (1996) reweighting approach to perform this analysis.

4 Results

4.1 Mean decomposition of test scores differences

In order to assess whether a test score determination process differs according to gender, Table 3 shows how much of the gender gap in achievement is explained by differences in the student's actual characteristics and how much is due to the returns of these characteristics. This latter portion of the gender gap is called unexplained, and consists of differences in the estimated coefficients for each gender. It is important to highlight that we calculate the difference with girls as the reference group, therefore the gap is positive if it favors boys and negative otherwise.

Table 3 - Oaxaca-Blinder decomposition 5th and 9th grade students in Brazil (2009-2015).

	5th grade		9th grade	
	Math	Portuguese	Math	Portuguese
Boys	221.419***	198.055***	253.714***	238.693***
	(0.131)	(0.111)	(0.124)	(0.117)
Girls	215.816***	206.411***	244.695***	251.339***
	(0.128)	(0.111)	(0.118)	(0.110)
Difference	5.603***	-8.356***	9.019***	-12.646***
	(0.068)	(0.064)	(0.082)	(0.077)
Explained	-0.032	-0.190***	0.232***	0.215***
	(0.028)	(0.028)	(0.025)	(0.026)
Students	-0.023	-0.165***	0.013	-0.029
	(0.020)	(0.020)	(0.019)	(0.020)
Teachers	0.002	0.002	-0.002	-0.001
	(0.003)	(0.003)	(0.002)	(0.002)
Schools	0.004	0.005*	0.015***	0.019***
	(0.002)	(0.003)	(0.004)	(0.003)
Regions	0.007	0.008	0.003	0.013
	(0.013)	(0.010)	(0.011)	(0.010)
Year	-0.022***	-0.041***	0.204***	0.212***
	(0.005)	(0.009)	(0.014)	(0.015)
Unexplained	5.635***	-8.166***	8.786***	-12.860***
	(0.063)	(0.057)	(0.078)	(0.073)
Students	1.663***	-0.575**	2.517***	3.604***
	(0.246)	(0.226)	(0.251)	(0.243)
Teachers	-0.014	-0.356***	-0.173	0.200
	(0.115)	(0.117)	(0.168)	(0.171)
Schools	0.071*	0.028	-0.003	-0.077*
	(0.039)	(0.038)	(0.037)	(0.039)
Regions	-0.370***	-0.286***	-0.961***	-0.709***
	(0.040)	(0.038)	(0.056)	(0.053)
Year	0.083***	0.001	-0.080***	-0.056***
	(0.009)	(0.009)	(0.017)	(0.018)
Intercept	4.202***	-6.978***	7.487***	-15.822***
	(0.266)	(0.248)	(0.319)	(0.311)
Observations	4070344	4056137	3573621	3538880

Robust standard errors clustered at the school level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Elaborated by the authors using SAEB (2009-2015).

As we can see from Table 3, boys and girls have different educational production functions for Math and Portuguese. The gender gaps on achievement are mostly due to the unexplained, or coefficient component of the Oaxaca-Blinder decomposition, and have opposite signs when we compare the two subjects. Moreover, the mean differences on achievement are smaller for 5th graders, which is in accordance with the idea that the gap widens with age, as pointed out by the literature, such as Fryer and Levitt (2010) and Bedard and Cho (2010).

In general, however, the results show that a greater part of the gender gap is not due to endowments and is more associated with the return of these characteristics, especially the ones related to student background and to the intercept. Interestingly, studies for developed countries have found a greater role of the endowments on achievement (LE; NGUYEN, 2018; GEVREK; SEIBERLICH, 2014; SOHN, 2012), and also encountered a large part of the gender test score gap being explained by the intercept (LE; NGUYEN, 2018; GOLSTEYN; SCHILS, 2014). Our results show that the endowments part of the gender test score differences of 5^{th} grades is only statistically significant for Portuguese and, indicates that the individual characteristics of girls help them outperform boys at literacy. As to the unexplained component, the detailed decomposition, as shown in the bottom of Table 3, highlights that the differences rely not only on the intercept, but also on the characteristics of students and on the distinct regions of Brazil they live in. In literacy, all of the grouped variables show an improvement of girls relative to boys on achievement, therefore contributing to widen this gap, which favors girls. Nonetheless, the return to individual characteristics in Math contributes to a better performance of boys relative to girls, indicating that they are better able to convert these educational inputs into higher scores.

When we look at the unexplained component detailed by each variable, on Table 5A in the Appendix, both subjects have “doing the subjects’ homework” and “living in urban areas” as the main contributors towards a relative better performance of boys, while “work” and “attending kindergarten” are the main contributors for girls. This last result is also observed by Le and Nguyen (2018), who highlight the importance of pre-school as a positive enhancer of girls’ educational achievement in Australia. Moreover, investigating the differences on top performers Fortin et al. (2015) also finds that family environment and labor market work during school are important in accounting for gender achievement gap. Also, regardless of the subject, teacher sex, that is, if the teacher is female, also contributes towards girls’ relative performance. This last result is consistent with the importance of gender on the dynamics between students and teachers, as highlighted by Dee (2006).

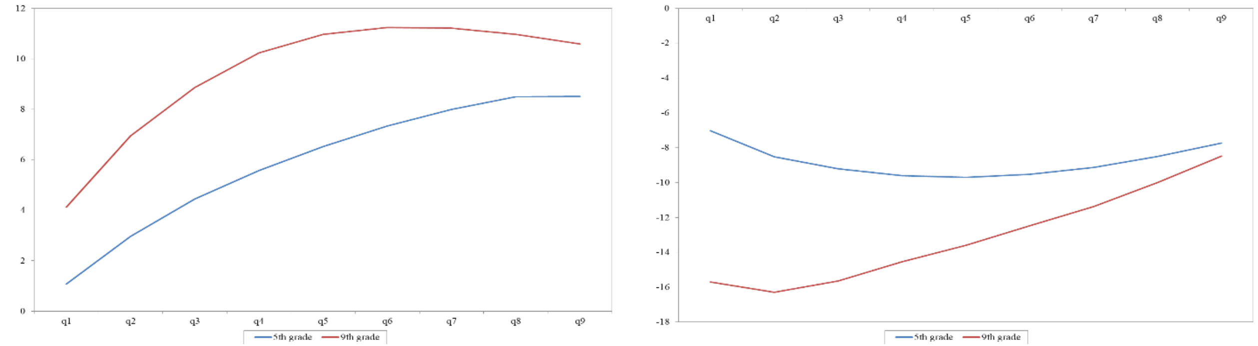
The results for the 9^{th} grade students are in line with the pattern described for 5^{th} graders. Again, the unexplained component of the decomposition is responsible for explaining the greater bulk of the gender gap. However, the explained part is statistically significant for both subjects, and both indicate an improvement towards boys relative to girls, especially due to the year effect and the schools characteristics. As to the unexplained component, the return of individual characteristics have a positive sign in literacy, meaning that, despite the fact that boys characteristics contribute to a relative better performance of them over girls in language, the effect of the intercept more than suppresses it.

When we look at the contribution of each variable for both subjects, on Table 6A in the Appendix, the ones on the coefficients component that are most important for the relative mean performance of girls are having attended kindergarten and if she lives with her mom and dad, and, for boys, are if they are living in urban areas, are working, or if at least one of the parents graduated from university. These results are in line with Biondi and Felício (2007), Franco and Menezes-Filho (2008) and Scorzafave and Ferreira (2011) who acknowledge that the characteristics of the students and of their families are the most relevant to explain academic performance in Brazil, while school, teachers and the principal characteristics display a minor role in student’s achievement. The next subsection further explores this same framework but using the RIF decomposition for percentiles of the test scores distributions.

4.2 Quantile decomposition of test score differences

After assessing the differences on mean academic achievement, we present the results capturing the gap along the distribution of the test scores at the two different grades. For that, we classify both girls and boys into percentiles of each test score distribution. Figure 4, shows the estimated differences, while Table 7A and 8A, on the Appendix, presents the detailed version of the estimated RIF decompositions.

Figure 4 - Quantile differences in Math and Portuguese achievement from RIF decomposition (2009-2015).



a) Math quantile difference b) Portuguese quantile difference

Source: Elaborated by the authors using SAEB (2009-2015).

Figure 4 highlights that, regarding the difference in performance, boys and girls tend to perform more equally in Math, as shown by the scale of both graphs. Students in 5th grade have a less unequal performance on both subjects and on all percentiles than 9th grade students. This is shown by the fact that the blue line, which indicates 5th graders, is below the one for 9th grade in Math, where boys outperform girls. Also, since the gap is calculated by the difference between boys and girls, the line for 5th graders is above the red line on Portuguese in all percentiles.

As to the shapes of the lines, the 9th grade gender differentials in Math, as described by the red line, is steeper, which indicates a greater gap on the first percentiles, as compared to the blue line. Nonetheless, it also shows a slight decrease on the two last percentiles, indicating that the difference between the best girls and boys in Math tend to slightly diminish. In Portuguese, the line for 9th graders shows that the worst performing boys tend to perform much lower than the worst girls, a difference that tend to diminish as we explore further percentiles. With regard to 5th grade, the line for Portuguese is more U-shaped, indicating that the greatest difference in performance is more towards the middle, that is, the best and worst tend to perform relatively more equally.

Tables 7A, on the Appendix, detail the decomposition by each percentile. There, it is shown that for the 5th grade Math, as the top panel of Table 7A shows, the unexplained component of Student characteristics is not statistically significant only for the first percentile. In Portuguese, at the lower panel of Table 7A, the students aggregated variable in the coefficients component reverses its sign in the lowest percentiles, showing that it contributes towards a relative better performance of girls at the first percentiles of the test score distribution, and contributes towards boys on the remaining percentiles. The main drivers of this change[5] are three variables, the intercept, having the ideal age for the grade, that in accordance with the aggregated variable, also change sings, and kindergarten, whose effect is higher and favors girls at lower percentiles.

Moreover, Table 8A describes the factors behind 9th grade gender differences by percentile. The results are very similar to the ones described by the Oaxaca-Blinder decomposition, and indicate a uniform importance of students' characteristics at the unexplained component of the decomposition for both subjects. Also, and specifically for Math, there is an increasing importance of the intercept as the main driver of the gender difference. This indicates that our model is able to account for most of the gender gap at lower quantiles, but other factors explain a greater bulk of this gap at higher percentiles.

Similar to our results, Ellison and Swanson (2010) highlights that for the United States the gender gap in Math skills widens at the top of the distribution. Besides that, Robinson and Lubiensky (2010) also find that in reading, gaps favoring females generally narrow but widen among low-achieving students. This is an interesting contrast with the Brazilian case, as we have seen that differences in literacy narrow among bottom and top performers, the latter especially on 9th grade. Moreover, Le and Nguyen (2018) investigating Australian test scores from 3rd to 7th grade find, in contrast with our results, sparse significant gender differences on reading, but they also find a more distinguished difference

---

[5] The detailed decomposition by each variable is not shown due to space restrictions, but are available upon request from the authors.

favoring males in math over virtually the whole distribution and in all grades, especially at the upper end of the distribution. Nonetheless, in relation to these other authors, our findings show that the gender differences in the Brazilian case are much more pronounced. One possible explanation for that, according to Guiso et al. (2008), is the role that social and cultural gender-biased environments play in these sex differences, as he indicates, by investigating a range of countries, that the gender gap in math disappears in more gender-equal societies, because girls improve their test scores, while the gap on literature, which favors girls, expands.

Overall, regardless of the grade, our results indicate that the differences in achievement are not due to the fact that one of the genders comes from a better background, or have better teachers or study in better schools, but are associated with the return of these characteristics. Also, despite the fact that we are able to disentangle the gender gap and point out some important contributors of this difference with the covariates in our dataset, most of the gap is still due to the intercept, and therefore require additional investigations.

5 Concluding Remarks

This paper aims at identifying the sources of gender differences in Math and Portuguese achievement in 5th and 9th grade. The inclusion of the more advanced grade is justified by the search for differences when considering older students. In order to attend our goals, we use a pool of four cohorts of 5th graders, and, separately, a pool of 9th graders, to employ two different types of decomposition methods on a dataset that includes students, teachers and schools characteristics.

Moreover, we apply the Oaxaca-Blinder decomposition technique that allow us to assess how much of the mean gap between girls and boys depends on differences in endowments, and how much of it is attributable to differences in the return of those endowments, namely in the unexplained component. The other decomposition method we use is an extension of the Oaxaca-Blinder that allows us to study the contribution of specific covariates at different parts of the test scores distributions.

The results we find indicate that the gap in Math favors boys, and the one in literacy favors girls. Besides that, we highlight that boys and girls have very distinct educational production functions, as the unexplained or coefficient component of the Oaxaca-Blinder decomposition, which corresponds to the greater bulk of both gaps, have a different sign for each subject. Moreover, from the aspects we are able to measure, we identify that the Students characteristics is the main constituent of this coefficient component. In sum, our results show that the majority of the gender difference is related to how boys and girls translate their background characteristics into achievement, and not to their characteristics in itself, and that this effect favors boys in Math and girls in Portuguese.

Also, despite the fact that the quantile decompositions results are very similar to the ones described by the Oaxaca-Blinder decomposition, however, the coefficient effect of the students' characteristics on 5th grade Portuguese academic achievement, changes its sign after the 40th percentile, showing that it contributes towards a relative better performance of girls at the lowest percentiles, and does the opposite on the highest quantiles. Therefore, indicating a change in the production function that is not captured by the mean decomposition. Nonetheless, with the percentile decomposition we are also able to identify that boys and girls tend to perform relatively more equally in Math than in Portuguese, and on 5th grade than on 9th.

Overall, regardless of the grade, our results indicate that the differences in achievement are not due to the fact that one of the genders comes from a better background, or have better teachers or study in better schools, but is associated with the return of these characteristics. Therefore, on the policy perspective, our study findings indicates that a more efficient policy aiming at the reduction of test score differences between boys and girls should look beyond resources. It also reduces the role played by teachers and schools in narrowing these test score differences. Moreover, despite the fact that we are able to disentangle the gender gap and point out some important contributors of this difference with the covariates in our dataset, most of the gap is still due to the intercept, and therefore require additional investigations, a task we intend to continue performing on the next articles of this thesis.

6 References

ALVES, M. T. G.; SOARES, J. F.; PEREIRA, F. Desigualdades educacionais no ensino fundamental de 2005 a 2013: hiato entre grupos sociais. **Revista Brasileira de Sociologia**, Sergipe, v. 4, n. 7, p. 49-82, 2016.

BAKER, M.; MILLIGAN, K. Boy-girl differences in parental time investments: Evidence from three countries. **Journal of Human Capital**, Chicago, v. 10, n. 4, p. 399-441, 2016.

BARSKY, R. BOUND, J.; CHARLES, K. K; LUPTON, J. P. Accounting for the black-white wealth gap: a nonparametric approach. **Journal of the American Statistical Association**, New York, v. 97, p. 663-673, 2002.

BEDARD, K.; CHO, I. Early gender test score gaps across OECD countries. **Economics of Education Review**, Cambridge, v. 29, n. 3, p. 348-363, 2010.

BERTRAND, M.; PAN, J. The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. **American Economic Journal: Applied Economics**, Washington, v. 5, n. 1, p. 32-64, 2013.

BHARADWAJ, P.; DE GIORGI, G.; HANSEN, D.; NEILSON, C. The gender gap in mathematics: Evidence from Chile. **Economic Development and Cultural Change**, Chicago, v. 65, n. 1, 2016.

BIONDI, R. L.; FELÍCIO, F. **Atributos escolares e o desempenho dos estudantes**: uma análise em painel dos dados do SAEB. Brasília: INEP, 2007. (Texto para Discussão, 25).

BLINDER, A. S. Wage Discrimination: Reduced Form and Structural Estimates. **Journal of Human Resources**, Madison, v. 8, n, 4, p. 436-55, 1973.

BUCHMANN, C.; DIPRETE, T. A.; MCDANIEL, A. Gender inequalities in education. **Annual Review Sociology**. v. 34, p. 319-337, 2008.

CAHILL, L. Why sex matters for neuroscience. **Nature Reviews Neuroscience**. v. 7, p. 477-484, 2006.

CONTINI, D.; TOMMASO, M. L. D.; MENDOLIA, S. The Gender Gap in Mathematics Achievement: Evidence from Italian Data. **Economics of Education Review**, Cambridge v.58, p. 32-42, 2017.

CORSEUIL, C.; FOGUEL, M. N.; GONZAGA, G.; RIBEIRO, E. P. The effects of Youth Training Program on Youth Turnover in Brazil. São Paulo: REAP – Rede de Economia Aplicada, 2012. (REAP Working Paper, 042).

DEE, T. S. The Why Chromosome: How a Teacher's Gender Affects Boys and Girls. **Education Next**, Cambridge, v. 6, n. 4, p. 68-75, 2006.

DICKERSON, A.; MCINTOSH, S.; VALENTE, C. Do the maths: An analysis of the gender gap in mathematics in Africa. **Economics of Education Review**, Cambridge, v. 46, p. 1-22, 2015.

DINARDO, J.; FORTIN, N. M.; LEMIEUX, T. Labor Market Institutions and the Distribution of Wages 1973-1992: A Semiparametric Approach. **Econometrica**, Chicago, v. 64, p. 1001-1044, 1996.

ELLISON, G.; SWANSON, A. The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions. **Journal of Economic Perspectives**, Nashville, v. 24, n. 2, p. 109-28, 2010.

FIRPO, S.; FORTIN, N. M.; LEMIEUX, T. Unconditional Quantile Regressions. **Econometrica**, Chicago, v. 77, p. 953-973, 2009.

FIRPO, S. Identifying and measuring economic discrimination. **IZA World of Labor**, Bonn, v. 347, 2017.

FORTIN, N. M.; LEMIEUX, T; FIRPO, S. Decomposition Methods in Economics. In: ASHENFELTER, O.; CARD, D. (Eds.). **Handbook of Labor Economics**, Amsterdam, Elsevier, 2011.

FORTIN, N.; OREOPOULOS; P.; PHIPPS, S. Leaving Boys Behind: Gender Disparities in High Academic Achievement. **Journal of Human Resources**, Madison, v. 50, n. 3, 2015

FRANCO, A. M. P.; MENEZES-FILHO, N. A.; Os determinantes do aprendizado com dados de um painel de escolas do SAEB. In: Encontro Nacional de Economia, 2008. **Anais...** Salvador: Anpec, 2008.

FRYER JR, R.; LEVITT, S. D. An Empirical Analysis of the Gender Gap in Mathematics. **American Economic Journal: Applied Economics**, Washington, v. 2, n. 2, p. 210-40, 2010.

GEVREK, Z. E.; SEIBERLICH, R. R. Semiparametric Decomposition of the Gender Achievement Gap: An Application for Turkey. **Labour Economics**, Amsterdam, v. 31, p. 27-44, 2014.

GOLSTEYN, B, H.; SCHILS, T. Gender gaps in primary school achievement: a decomposition into endowments and returns to IQ and non-cognitive factors. **Economics of Education Review**, Cambridge, v. 41, p. 176-187, 2014.

GUISO, L.; MONTE, F.; SAPIENZA, P.; ZINGALES, L. Culture, gender and math. **Science**, Washington, v. 320, p. 1164-1165, 2008.

HEDGES, L. V.; NOWELL, A. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. **Science**. v. 269, n. 5220, p. 41-45, 1995.

HUSAIN, M.; MILLIMET, D. L. The mythical ‘boy crisis’? **Economics of Education Review**. v. 28, n. 1, p. 38-48, 2009.

HYDE, J. S.; FENNEMA, E.; RYAN, M.; FROST, L. A.; HOPP, C. Gender comparisons of mathematics attitudes and affect: a meta-analysis. **Psychology of women quarterly**, New York, v. 14, n. 3, p. 299-342, 1990.

JANN, B. The Blinder-Oaxaca decomposition for linear regression models. **The Stata Journal**, College Station, v. 8, n. 4, p. 453-479, 2008.

KASSOUF, A. L. Wage gender discrimination and segmentation in the Brazilian labor market. **Economia Aplicada**. Ribeirão Preto, v. 2, n. 2, p. 243-269, 1998.

KLASEN, S. Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development**. The World Bank Economic Review**, 2002.

LE, H. T.; NGUYEN, H. T. The evolution of the gender test score gap through seventh grade: new insights from Australia using unconditional quantile regression and decomposition. **IZA Journal of Labor Economics**, Amsterdam, v. 7, n. 2, 2018.

LEAHEY, E.; GUO, G. Gender differences in mathematical trajectories. **Social forces**. Chapel Hill, v. 80, n. 2, p. 713-732, 2001.

MADALOZZO, R. **Market and Home Production**: Gender Differences in Brazil. São Paulo: Insper, 2009. (Insper Working Paper, 168**)**.

MADALOZZO, R.; ARTES, R. Escolhas profissionais e impactos no diferencial salarial entre homens e mulheres. **Cadernos de Pesquisa**, Brasília, v. 47, n. 163, 2017.

OAXACA, R. Male-Female Wage Differentials in Urban Labor Markets. **International Economic Review**, Pennsylvania, v. 14, n. 3, p. 693-709, 1973.

O’DONNEL, O.; VAN DOORSLAER, E.; WAGSTAFF, A; LINDELOW, M. Explaining differences between groups: Oaxaca decomposition. **Analyzing Health Equity Using Household Survey Data**, Washington-DC: World Bank, 2008.

RASCHE JÚNIOR, F. **Privação de Liberdades Femininas:** uma análise da participação das mulheres nas áreas de ciência e tecnologia no Brasil entre 2004 e 2013. 2015. 73f. Undergraduate Thesis

(Graduation in Economics) – Faculty of Economics, Rio Grande do Sul Federal University, Porto Alegre.

ROBINSON, J. P.; LUBIENSKI, S. T. The Development of Gender Achievement Gaps in Mathematics and Reading During Elementary and Middle School: Examining Direct Cognitive Assessments and Teacher Ratings. **American Educational Research Journal**, Washington, p. 1035, 2010.

SAX, L. **Why gender matters**: What parents and teachers need to know about the emerging science of sex differences. New York: Doubleday. 2005.

SCHULTZ, T., P. Why Governments Should Invest more to Educate Girls. **World Development**, Oxford, v. 30, n. 2, p. 207-225, 2002.

SCORZAFAVE, L. G.; FERREIRA, R. A. Desigualdade de Proficiência no Ensino Fundamental Público Brasileiro: Uma Análise de Decomposição. **Revista EconomiA**, São Paulo, v. 12, n. 2, p. 337-359, 2011.

SOARES, J. F.; CANDIAN, J. O efeito da escola básica brasileira: As evidências do PISA e do SAEB. **Revista Contemporânea de Educação**, Rio de Janeiro, v. 2, n. 4, 2007.

SOARES, J. F.; DELGADO, V. M. S. Medida das desigualdades de aprendizado entre estudantes de ensino fundamental. **Estudos em Avaliação Educacional**, Brasília, v. 27, n. 66, 2016.

SOARES, S. S. D.; NASCIMENTO, P. A. M. M. Evolução do desempenho cognitivo dos jovens brasileiros no Pisa. **Cadernos de Pesquisa**, Brasília, v. 42, n. 145, p. 68-87, 2012.

SOARES, J. F.; ALVES, M. T. G.; XAVIER, F. P. Effects of Brazilian schools on student learning. **Assessment in Education: Principles, Policy & Practice**, Oxford, v. 23, n. 1, p. 75-97, 2016.

SOHN, K. A new insight into the gender gap in math. **Bulletin of Economic Research**, Hull, v. 64, n. 1, p. 135-155, 2012.

SPELKE, E. S. Sex differences in intrinsic aptitude for mathematics and science? A critical review. **American Psychologist**. Washington, v. 60, n. 9, p. 950-958, 2005.

TILLMANN; E. T.; COMIM, F. Os determinantes da decisão entre estudo e trabalho dos jovens no Brasil e a geração Nem-Nem. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 46, n. 2, p. 47-78, 2016.

WORLD BANK. **World Development Report 2012**: gender equality and development. Washington: World Bank Publications, 2011.

YUN, M. A simple solution to the identification problem in detailed wage decompositions. **Economic Inquiry**, Long Beach, v. 43, n. 4, p. 766-772, 2005.

APPENDIX

Table 1A - Description of the variables included in the estimates.

Variable	Description
Qescz	index built from school has a computer lab, a science lab, a sports gym, a library and a reading room
Sesz	index built from number of bedrooms, bathrooms, computers, cars and televisions in the student household
Familiaz	index built from the parents attend school meetings, if they talk about what happened at school and, if they incentive the student to study, go to school and read
White	1 if student is white, 0 otherwise
Idealage	1 if has the ideal age for the grade, 0 otherwise
mom&dad	1 if student lives with his mom and dad, 0 otherwise
parentes_univdg	1 if at least one of the student parent has na university degree
domestic chores	1 if student reports doing household chores, 0 otherwise
Work	1 if student works, 0 otherwise
Urban	1 if student lives in an urban area, 0 otherwise
Begschool	1 if student attended kindergarten, 0 otherwise
Teachsex	1 if teacher is female, 0 otherwise
Teachhighdg	1 if teacher has university degree, 0 otherwise
Teachothact	1 if teacher works on another job besides teaching, 0 otherwise
Teachexper	1 if teacher has more than 10 years of experience, 0 otherwise
Teachhrs	1 if teacher works more than 40 hours, 0 otherwise
Hmwok	1 if student reports that he or she does his homework, 0 otherwise
Teachcontract	1 if teacher has a permanent job contrat with the school, 0 otherwise
Region	indicator variable for each of the five regions in Brazil (North, Northeast, Central-West, Southeast and South)
Bookatbeg	1 if teacher states that most of his/hers students had the classroom book at the beggining of the school year, 0 otherwise

Table 5A - Oaxaca-Blinder 5th grade detailed decomposition.

	Math	Portuguese		Math	Portuguese
	Explained			Unexplained	
qescz	0.002	0.003*	qescz	0.005***	0.001
	(0.001)	(0.002)		(0.002)	(0.001)
sesz	0.022***	0.019***	sesz	0.011***	0.002***
	(0.008)	(0.007)		(0.003)	(0.001)
familiaz	-0.029***	-0.038***	familiaz	-0.045***	-0.071***
	(0.004)	(0.005)		(0.006)	(0.005)
white	0.004**	0.004*	white	0.019	-0.048***
	(0.002)	(0.002)		(0.018)	(0.018)
idealage	-0.003	-0.004	idealage	0.016	0.111**
	(0.009)	(0.009)		(0.056)	(0.053)
livesmomdad	0.000	-0.000	livesmomdad	-0.097***	-0.075***
	(0.001)	(0.001)		(0.025)	(0.023)
escresp	0.002***	0.003***	escresp	0.230***	0.226***
	(0.001)	(0.001)		(0.054)	(0.058)
domwork	0.002	0.002	domwork	0.243***	0.144***
	(0.006)	(0.006)		(0.011)	(0.011)
work	-0.010*	-0.012*	work	-0.806***	-1.481***
	(0.005)	(0.007)		(0.092)	(0.080)
urban	0.004	0.006	urban	0.969***	0.397**
	(0.003)	(0.004)		(0.181)	(0.158)
begschool	-0.011*	-0.008	begschool	-0.473***	-0.564***
	(0.006)	(0.005)		(0.041)	(0.041)
teachsex	-0.000	0.000	teachsex	-0.241***	-0.306***
	(0.001)	(0.001)		(0.080)	(0.079)
teachhighdg	0.001	0.001	teachhighdg	0.076	-0.113
	(0.002)	(0.001)		(0.078)	(0.082)
teachothact	0.000	0.000	teachothact	0.019*	0.016
	(0.001)	(0.001)		(0.010)	(0.010)
teachexper	0.000	0.000	teachexper	-0.006	-0.013
	(0.001)	(0.001)		(0.031)	(0.028)
teachhrs	-0.000	-0.000	teachhrs	0.032***	0.021***
	(0.000)	(0.001)		(0.007)	(0.006)
mathhmwok	-0.004*	-0.135***	mathhmwok	1.597***	0.784***
	(0.002)	(0.004)		(0.145)	(0.124)
teachcontract	0.001	0.001	teachcontract	0.107***	0.039
	(0.002)	(0.001)		(0.040)	(0.035)
region	0.007	0.008	region	-0.370***	-0.286***
	(0.013)	(0.010)		(0.040)	(0.038)
year	-0.022***	-0.041***	year	0.083***	0.001
	(0.005)	(0.009)		(0.009)	(0.009)
bookatbeg	0.002	0.002	bookatbeg	0.066*	0.028
	(0.002)	(0.002)		(0.039)	(0.038)
			Intercept	4.202***	-6.978***
				(0.266)	(0.248)

Robust standard errors clustered at the school level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Region is a grouped variable for the five regions of Brazil, and its base category is the Southeast region. Year is also a grouped variable for the four years of analysis and its benchmark is 2009.
Source: Elaborated by the authors using SAEB (2009-2015).

Table 6A - Oaxaca-Blinder 9th grade detailed decomposition.

	Math	Portuguese		Math	Portuguese
	explained			Unexplained	
qescz	-0.005*	-0.002	Qescz	0.000	0.001
	(0.003)	(0.002)		(0.001)	(0.001)
sesz	0.033***	0.030***	Sesz	0.061***	0.082***
	(0.007)	(0.007)		(0.006)	(0.007)
familiaz	-0.001***	0.003***	Familiaz	-0.045***	-0.063***
	(0.000)	(0.001)		(0.003)	(0.004)
white	0.014***	0.014***	White	-0.028	-0.042**
	(0.005)	(0.005)		(0.017)	(0.017)
idealage	-0.101***	-0.103***	Idealage	-0.072***	-0.098***
	(0.013)	(0.013)		(0.013)	(0.014)
livesmomdad	-0.001*	0.000	livesmomdad	-0.202***	-0.237***
	(0.001)	(0.000)		(0.022)	(0.022)
escresp	0.036***	0.035***	Escresp	1.102***	1.216***
	(0.004)	(0.003)		(0.079)	(0.077)
domwork	0.005*	0.004	Domwork	-0.110***	0.025
	(0.003)	(0.003)		(0.017)	(0.017)
work	0.009***	0.019***	Work	0.452***	0.856***
	(0.002)	(0.003)		(0.062)	(0.058)
urban	0.011***	0.016***	Urban	1.270***	1.208***
	(0.003)	(0.005)		(0.210)	(0.192)
begschool	0.003	0.005	Begschool	-0.401***	-0.331***
	(0.003)	(0.003)		(0.058)	(0.061)
teachsex	-0.005***	-0.003***	Teachsex	-0.020	-0.035
	(0.001)	(0.001)		(0.022)	(0.029)
teachhighdg	0.000	0.000	teachhighdg	-0.349**	0.109
	(0.001)	(0.000)		(0.165)	(0.169)
teachothact	-0.000	-0.000	teachothact	-0.007	-0.008
	(0.000)	(0.000)		(0.005)	(0.006)
teachexper	0.000	-0.000	teachexper	0.006	0.025
	(0.001)	(0.001)		(0.025)	(0.023)
teachhrs	0.002***	0.003***	Teachhrs	0.024***	0.020**
	(0.001)	(0.001)		(0.009)	(0.008)
mathhmwok	0.004	-0.052***	mathhmwok	0.489***	0.987***
	(0.004)	(0.002)		(0.101)	(0.127)
teachcontract	0.001	-0.000	teachcontract	0.173***	0.089**
	(0.001)	(0.001)		(0.045)	(0.036)
regiao	0.003	0.013	Regiao	-0.961***	-0.709***
	(0.011)	(0.010)		(0.056)	(0.053)
year	0.204***	0.212***	Year	-0.080***	-0.056***
	(0.014)	(0.015)		(0.017)	(0.018)
bookatbeg	0.020***	0.021***	bookatbeg	-0.003	-0.078**
	(0.003)	(0.003)		(0.037)	(0.039)
			Intercept	7.487***	-15.822***
				(0.319)	(0.311)

Robust standard errors clustered at the school level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Region is a grouped variable for the five regions of Brazil, and its base category is the Southeast region. Year is also a grouped variable for the four years of analysis and its benchmark is 2009.
Source: Elaborated by the authors using SAEB (2009-2015).

Table 7A - RIF detailed decomposition for 5th graders.

	5th grade Math								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Boys	159.340***	178.152***	192.928***	206.307***	219.422***	232.963***	247.492***	264.187***	286.617***
	(0.130)	(0.133)	(0.143)	(0.147)	(0.153)	(0.155)	(0.157)	(0.161)	(0.168)
Girls	158.268***	175.187***	188.479***	200.736***	212.893***	225.615***	239.502***	255.691***	278.108***
	(0.112)	(0.117)	(0.128)	(0.137)	(0.147)	(0.154)	(0.156)	(0.163)	(0.179)
Difference	1.072***	2.966***	4.448***	5.571***	6.529***	7.348***	7.990***	8.495***	8.510***
	(0.112)	(0.094)	(0.094)	(0.088)	(0.093)	(0.091)	(0.095)	(0.099)	(0.119)
Explained	-0.035	-0.042	-0.044	-0.043	-0.040	-0.035	-0.030	-0.023	-0.015
	(0.025)	(0.028)	(0.030)	(0.032)	(0.033)	(0.032)	(0.031)	(0.029)	(0.026)
Student	-0.038**	-0.038*	-0.037*	-0.033	-0.028	-0.022	-0.015	-0.007	0.002
	(0.018)	(0.020)	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)	(0.020)	(0.018)
Teachers	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Schools	0.003*	0.004*	0.004	0.004	0.004	0.005	0.004	0.004	0.003
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Region	0.013	0.011	0.009	0.008	0.007	0.006	0.004	0.003	0.001
	(0.011)	(0.013)	(0.015)	(0.015)	(0.016)	(0.016)	(0.015)	(0.013)	(0.011)
Year	-0.015***	-0.020***	-0.023***	-0.024***	-0.025***	-0.025***	-0.026***	-0.025***	-0.023***
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Unexplained	1.107***	3.007***	4.493***	5.615***	6.569***	7.383***	8.020***	8.519***	8.524***
	(0.109)	(0.090)	(0.089)	(0.083)	(0.087)	(0.085)	(0.090)	(0.095)	(0.117)
Student	0.033	1.551***	2.147***	1.987***	2.571***	2.530***	2.643***	2.488***	2.073***
	(0.499)	(0.375)	(0.344)	(0.328)	(0.323)	(0.323)	(0.336)	(0.347)	(0.431)
Teachers	0.115	0.059	0.262	-0.048	0.015	0.184	-0.124	-0.323**	-0.467**
	(0.234)	(0.204)	(0.182)	(0.161)	(0.179)	(0.172)	(0.157)	(0.160)	(0.185)
Schools	0.123	0.210***	0.133**	0.083	0.099**	-0.023	0.030	0.056	-0.031
	(0.111)	(0.068)	(0.054)	(0.052)	(0.050)	(0.057)	(0.060)	(0.070)	(0.058)
Region	-0.464***	-0.462***	-0.372***	-0.327***	-0.281***	-0.342***	-0.391***	-0.359***	-0.394***
	(0.065)	(0.055)	(0.057)	(0.054)	(0.055)	(0.055)	(0.058)	(0.061)	(0.072)
Year	0.102***	0.097***	0.086***	0.076***	0.079***	0.100***	0.090***	0.069***	0.060***
	(0.020)	(0.014)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.015)	(0.014)
Intercept	1.198**	1.554***	2.238***	3.843***	4.086***	4.934***	5.772***	6.589***	7.283***
	(0.572)	(0.426)	(0.384)	(0.363)	(0.350)	(0.348)	(0.348)	(0.352)	(0.425)

Robust standard errors clustered at the school level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Elaborated by the authors using SAEB (2009-2015).

Table 7A. (continuation)

	5th grade Portuguese								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Boys	139.097***	156.024***	169.963***	182.653***	195.188***	208.141***	222.226***	238.647***	260.940***
	(0.101)	(0.117)	(0.122)	(0.131)	(0.134)	(0.136)	(0.137)	(0.139)	(0.144)
Girls	146.112***	164.538***	179.162***	192.244***	204.889***	217.651***	231.352***	247.150***	268.663***
	(0.118)	(0.123)	(0.127)	(0.129)	(0.131)	(0.133)	(0.130)	(0.130)	(0.128)
Difference	-7.015***	-8.514***	-9.198***	-9.592***	-9.700***	-9.510***	-9.126***	-8.503***	-7.724***
	(0.100)	(0.096)	(0.089)	(0.087)	(0.088)	(0.088)	(0.088)	(0.106)	(0.116)
Explained	-0.218***	-0.233***	-0.232***	-0.224***	-0.212***	-0.199***	-0.179***	-0.156***	-0.127***
	(0.024)	(0.029)	(0.031)	(0.032)	(0.032)	(0.032)	(0.031)	(0.029)	(0.025)
Student	-0.217***	-0.226***	-0.218***	-0.202***	-0.183***	-0.164***	-0.139***	-0.114***	-0.085***
	(0.019)	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)	(0.022)	(0.020)	(0.018)
Teachers	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Schools	0.003*	0.004**	0.005*	0.005*	0.006*	0.006*	0.006*	0.006*	0.005*
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Region	0.014	0.014	0.013	0.011	0.009	0.007	0.005	0.003	0.001
	(0.009)	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.010)	(0.008)
Year	-0.021***	-0.028***	-0.034***	-0.041***	-0.046***	-0.051***	-0.053***	-0.054***	-0.050***
	(0.006)	(0.008)	(0.009)	(0.010)	(0.011)	(0.012)	(0.012)	(0.011)	(0.011)
Unexplained	-6.797***	-8.281***	-8.966***	-9.367***	-9.488***	-9.310***	-8.947***	-8.347***	-7.597***
	(0.097)	(0.092)	(0.083)	(0.081)	(0.082)	(0.082)	(0.083)	(0.102)	(0.113)
Student	-9.359***	-5.823***	-2.356***	-0.360	1.158***	2.426***	3.295***	4.247***	3.848***
	(0.483)	(0.414)	(0.378)	(0.337)	(0.341)	(0.310)	(0.285)	(0.332)	(0.330)
Teachers	-0.353	-0.337*	-0.484***	-0.455***	-0.327*	-0.224	-0.147	-0.252	-0.538
	(0.226)	(0.195)	(0.170)	(0.162)	(0.167)	(0.155)	(0.152)	(0.226)	(0.336)
Schools	-0.096	-0.092	-0.102*	0.011	0.068	0.059	0.063	0.119*	0.206***
	(0.066)	(0.061)	(0.060)	(0.054)	(0.050)	(0.050)	(0.061)	(0.067)	(0.058)
Region	-0.334***	-0.430***	-0.463***	-0.440***	-0.326***	-0.312***	-0.240***	-0.112*	-0.076
	(0.061)	(0.057)	(0.054)	(0.053)	(0.053)	(0.055)	(0.055)	(0.063)	(0.070)
Year	-0.097***	-0.116***	-0.084***	-0.058***	-0.013	0.033***	0.068***	0.105***	0.114***
	(0.014)	(0.013)	(0.013)	(0.012)	(0.011)	(0.012)	(0.013)	(0.015)	(0.015)
Intercept	3.442***	-1.483***	-5.478***	-8.065***	-10.048***	-11.292***	-11.987***	-12.453***	-11.150***
	(0.561)	(0.464)	(0.410)	(0.384)	(0.374)	(0.331)	(0.302)	(0.336)	(0.389)

Robust standard errors clustered at the school level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Elaborated by the authors using SAEB (2009-2015).

Table 8A - RIF detailed decomposition for 9th graders.

	9th grade Math								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Boys	191.754***	212.176***	227.777***	241.407***	254.003***	266.266***	279.111***	293.644***	313.698***
	(0.145)	(0.144)	(0.143)	(0.141)	(0.138)	(0.136)	(0.134)	(0.137)	(0.159)
Girls	187.640***	205.233***	218.908***	231.166***	243.029***	255.018***	267.890***	282.673***	303.102***
	(0.101)	(0.113)	(0.119)	(0.125)	(0.131)	(0.136)	(0.139)	(0.144)	(0.159)
Difference	4.114***	6.943***	8.869***	10.240***	10.974***	11.247***	11.221***	10.971***	10.596***
	(0.138)	(0.121)	(0.112)	(0.102)	(0.098)	(0.099)	(0.098)	(0.100)	(0.114)
Explained	0.224***	0.247***	0.246***	0.247***	0.242***	0.234***	0.227***	0.219***	0.232***
	(0.022)	(0.025)	(0.027)	(0.028)	(0.028)	(0.028)	(0.027)	(0.026)	(0.026)
Student	-0.040**	-0.029	-0.016	-0.003	0.011	0.023	0.035*	0.049**	0.072***
	(0.018)	(0.020)	(0.021)	(0.022)	(0.021)	(0.021)	(0.020)	(0.020)	(0.020)
Teachers	-0.001	-0.001	-0.001	-0.002	-0.002	-0.003	-0.003	-0.004	-0.004
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Schools	0.011***	0.014***	0.015***	0.016***	0.017***	0.017***	0.017***	0.016***	0.018***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Region	0.017*	0.013	0.009	0.006	0.003	-0.001	-0.003	-0.006	-0.010
	(0.009)	(0.011)	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.011)
Year	0.237***	0.250***	0.239***	0.229***	0.215***	0.198***	0.181***	0.164***	0.156***
	(0.018)	(0.017)	(0.016)	(0.015)	(0.014)	(0.013)	(0.012)	(0.011)	(0.011)
Unexplained	3.890***	6.696***	8.623***	9.994***	10.732***	11.013***	10.995***	10.752***	10.364***
	(0.137)	(0.119)	(0.109)	(0.098)	(0.093)	(0.094)	(0.094)	(0.096)	(0.111)
Student	2.662***	2.987***	3.113***	2.954***	2.622***	2.133***	2.174***	2.108***	2.473***
	(0.461)	(0.403)	(0.381)	(0.338)	(0.322)	(0.310)	(0.301)	(0.307)	(0.328)
Teachers	0.271	-0.060	-0.178	-0.144	-0.315*	-0.252	-0.053	-0.441**	-0.251
	(0.598)	(0.272)	(0.231)	(0.215)	(0.189)	(0.182)	(0.193)	(0.190)	(0.222)
Schools	0.047	-0.002	0.027	0.015	-0.017	-0.048	-0.067	-0.039	0.032
	(0.070)	(0.060)	(0.055)	(0.053)	(0.051)	(0.047)	(0.047)	(0.048)	(0.056)
Region	-0.964***	-1.098***	-1.058***	-1.004***	-1.014***	-1.000***	-1.029***	-0.968***	-0.901***
	(0.085)	(0.080)	(0.080)	(0.077)	(0.073)	(0.071)	(0.072)	(0.074)	(0.083)
Year	0.024	0.027	-0.061**	-0.121***	-0.207***	-0.259***	-0.204***	-0.086***	0.072***
	(0.037)	(0.030)	(0.026)	(0.023)	(0.022)	(0.022)	(0.021)	(0.021)	(0.025)
Intercept	1.850**	4.842***	6.779***	8.293***	9.663***	10.439***	10.174***	10.177***	8.939***
	(0.923)	(0.518)	(0.467)	(0.423)	(0.393)	(0.369)	(0.364)	(0.357)	(0.400)

Robust standard errors clustered at the school level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Elaborated by the authors using SAEB (2009-2015).

Table 8A. (continuation)

	9th grade Portuguese								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Boys	174.949***	195.522***	211.605***	225.768***	238.854***	251.875***	265.523***	281.096***	301.387***
	(0.129)	(0.138)	(0.146)	(0.142)	(0.139)	(0.138)	(0.137)	(0.136)	(0.134)
Girls	190.644***	211.806***	227.245***	240.304***	252.444***	264.350***	276.888***	291.091***	309.867***
	(0.136)	(0.131)	(0.130)	(0.127)	(0.125)	(0.123)	(0.120)	(0.116)	(0.113)
Difference	-15.695***	-16.284***	-15.640***	-14.537***	-13.590***	-12.475***	-11.366***	-9.995***	-8.480***
	(0.132)	(0.121)	(0.120)	(0.108)	(0.101)	(0.099)	(0.094)	(0.098)	(0.104)
Explained	0.151***	0.183***	0.211***	0.228***	0.242***	0.249***	0.252***	0.248***	0.230***
	(0.023)	(0.027)	(0.029)	(0.030)	(0.030)	(0.030)	(0.029)	(0.027)	(0.024)
Student	-0.110***	-0.093***	-0.071***	-0.049**	-0.027	-0.007	0.012	0.030	0.045***
	(0.020)	(0.022)	(0.023)	(0.023)	(0.022)	(0.022)	(0.021)	(0.019)	(0.017)
Teachers	0.000	-0.000	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Schools	0.018***	0.021***	0.022***	0.022***	0.021***	0.021***	0.020***	0.019***	0.017***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Region	0.031***	0.029***	0.024**	0.018	0.013	0.009	0.003	-0.000	-0.003
	(0.009)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.009)
Year	0.212***	0.226***	0.237***	0.237***	0.235***	0.229***	0.218***	0.200***	0.173***
	(0.016)	(0.018)	(0.018)	(0.018)	(0.017)	(0.016)	(0.015)	(0.014)	(0.012)
Unexplained	-15.846***	-16.467***	-15.851***	-14.764***	-13.831***	-12.725***	-11.617***	-10.242***	-8.710***
	(0.130)	(0.118)	(0.116)	(0.104)	(0.097)	(0.095)	(0.089)	(0.095)	(0.101)
Student	0.669	3.304***	4.872***	4.979***	5.213***	5.176***	4.580***	4.044***	3.697***
	(0.496)	(0.427)	(0.407)	(0.346)	(0.312)	(0.300)	(0.279)	(0.302)	(0.308)
Teachers	0.174	0.674**	0.272	0.164	0.286	0.216	0.386*	0.197	0.136
	(0.352)	(0.330)	(0.322)	(0.277)	(0.210)	(0.194)	(0.197)	(0.197)	(0.212)
Schools	-0.236***	-0.163**	-0.055	-0.073	-0.040	-0.009	0.016	0.066	-0.006
	(0.069)	(0.065)	(0.062)	(0.057)	(0.053)	(0.051)	(0.051)	(0.052)	(0.056)
Region	-0.557***	-0.924***	-1.109***	-1.056***	-0.946***	-0.815***	-0.714***	-0.472***	-0.347***
	(0.090)	(0.082)	(0.081)	(0.076)	(0.072)	(0.074)	(0.068)	(0.070)	(0.078)
Year	-0.373***	-0.287***	-0.177***	-0.085***	0.010	0.080***	0.113***	0.144***	0.120***
	(0.039)	(0.034)	(0.031)	(0.026)	(0.023)	(0.022)	(0.022)	(0.022)	(0.024)
Intercept	-15.522***	-19.070***	-19.654***	-18.694***	-18.354***	-17.373***	-15.998***	-14.221***	-12.311***
	(0.645)	(0.598)	(0.576)	(0.480)	(0.387)	(0.367)	(0.350)	(0.366)	(0.378)

Robust standard errors clustered at the school level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Elaborated by the authors using SAEB (2009-2015).