Income Inequality in Brazil: Is Education Quality an Important Determinant?

by

Catherine Michaud Leclerc

B.A., Laval University, 2012

Project Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

in the
Department of Economics
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY

Summer 2015

Approval

Name: **Catherine Michaud Leclerc**

Degree: **Master of Arts (Economics)**

Title: ***Income Inequality in Brazil: Is Education Quality an Important Determinant?***

Examining Committee: **Chair:** ***Simon Woodcock***
Associate Professor, Department of Economics

Fernando Aragón
Senior Supervisor
Assistant Professor, Department of Economics

Brian Krauth
Supervisor
Associate Professor, Department of Economics

Andrew McGee
Internal Examiner
Assistant Professor, Department of Economics

Date Defended: July 29, 2015

Abstract

This paper examines the contribution of education quality to income inequality in Brazil. More specifically, I look at the average income of municipalities in Brazil to observe how much the disparities in education quality across municipalities account for income inequality. Three commonly used methods to study income inequality are implemented: decomposition by population subgroups, a regression-based approach and the Oaxaca-Blinder decomposition. Using different measures for education quality, I find that the provision of a good measured education among public schools contributes little to the variation of average income across municipalities.

Keywords: Brazil; Income Inequality; Education Quality

À mes deux anges qui m'ont toujours grandement *supportée dans mes études.*

Acknowledgements

I first want to thank my professors from Laval University, Fernão Perestrelo and Francisco Pereira de Lima, for giving me the passion to study Portuguese and the culture of Portuguese speaking countries. My desire to study the educational system in Brazil began in their lectures, and this project would never have started without their initial guidance.

There are several faculty members from the Department of Economics at Simon Fraser that I also need to acknowledge. Without their help, my project would not be what it is. I want to thank my senior supervisor, Fernando Aragon, for his patience, guidance and for pushing me to develop my critical thinking skills in empirical research. I also want to thank Brian Krauth for his precious help and time with my project and especially for being understanding during my first semester at SFU. This gave me another chance to believe in my own capacities after a difficult moment I went through.

I want to specifically thank Andrew McGee for giving me the opportunity to work as his research assistant. Through this experience, I have learned a lot about research in Economics. It has greatly motivated me to pursue my studies in Economics, and to pursue new research ideas of my own.

Finally, I want to thank Gregory Dow, Erik Kimbrough, Luba Petersen, Stephen Easton, Simon Woodcock, Hitoshi Shigeoka, David Jacks and Lucas Herrenbrueck for their valuable advice and support while I was applying for my Ph.D.

There are also some friends and family members I would like to acknowledge. Without their help and support, I could not have been through the last two years. I first want to thank Jason Kiraly who helped me improve my English and make better decisions. I am very grateful to my friends Jean-François Morissette, Gabrielle B. Bergeron, Marie-Ève Jean, Maude Côté and Alexandra Demers-Roberge who were an incredible support during the hard times. A special thanks to my good friend Chitra Skelton who always understood that I had to study more and more, but who also made sure I had a social life during my studies.

I want to thank Frank Maynard for his active listening, his advices and for being there when I needed. I am deeply grateful to Michèle Tourigny, without whom I could not have been through it. I also want to thank Terence Cantin who followed me in this adventure in British Columbia. I hope you have learned and enjoyed this experience as much as I did.

Finally, to my dear mother Stéphanie Michaud: thank you for simply being who you are. You gave me the strength to face this world.

Table of Contents

List of Tables

List of Acronyms

FUNDEF	Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério / Fund for Maintenance and Development of the Fundamental Education and Valorization of Teaching
FUNDEP	Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização de Profissionais de Educação / Fund the Development of Basic Education and Appreciation of the Teaching Profession
GDP	Gross Domestic Product
GE	General Entropy
IBGE	Instituto Brasileiro de Geografia e Estatística / Brazilian Institute of Geography and Statistics
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais / National Institute of Studies and Educational Researches
OECD	Organisation for Economic Co-operation and Development
PNAD	Pesquisa Nacional por Amostra de Domicilios / Brazilian National Household Sample Survey
PNE	Plano Nacional de Educação/National Education Plan
SAEB	Sistema Nacional de Avaliação da Educação Básica / National Assessment of Basic Education
SFU	Simon Fraser University
UNESCO	United Nations Educational, Scientific and Cultural Organization
US	United States

Chapter 1.

Introduction

Brazil is one of the most unequal countries in the world. Although income inequality has greatly declined over the last twenty years (Ferreira, Leite and Litchfield, 2008), it is still at a high level and remains an important issue to address.[1] Many authors have written on the relationship between income inequality in Brazil and their educational system and claim that educational attainment is the most important determinant of the high levels of inequality across households.[2] However, education quality is often omitted in the analysis of the relationship between income inequality and education. This is a relevant shortcoming given the importance of quality of education on several economic outcomes.[3]

In this research paper, I fill this gap in the literature by examining the importance of education quality as a determinant of income inequality. More specifically, I observe how the disparities in the quality of primary and secondary public schools across municipalities affect the average earnings between these municipalities.

[1] According to the World Bank, the Gini Coefficient in 2012 was 52.7.

[2] See, among others, Lam (1999); Salardi (2005); Ferreira, Leite and Litchfield (2008); Barros et al. (2010); Lusting, Lopez-Calva and Ortiz-Juarez, (2013) and Azevedo, Inchauste and Sanfelice (2013).

[3] Many questions have been raised in the literature concerning the importance of education quality in Brazil. For example, Ferreira and Velez (2004) denounced the inefficiencies in the educational system where important chronic repetitions and dropouts are observed, especially among the poor. This is supported by Gomes-Neto and Hanushek (1994) who discussed the causes and consequences of the high retention rates in Brazil. They mention the importance of factors such as government policies, the low-quality of the teachers and the lack of material resources in schools as factors for a low school-quality. Lam (1999) also discusses the importance of the unequal distribution of education quality in Brazil.

The contribution of this paper is to introduce a variable that has not been studied yet in the literature on income inequality in Brazil; namely education quality. This research has potentially important policy impacts. The Brazilian National Congress recently established the National Education Plan (PNE) in which they want to increase the public expenditure on education to 10% of GDP by 2020, a notable target given that the average spending in the OECD countries in 2011 was slightly above 6% of GDP (OECD, 2014). It is crucial for the government and policy-makers to be wise in the allocation of this significant spending and consider the overall impact of the educational system on the economic development of the country.

To examine the contribution of education quality to income inequality, I combine two major Brazilian datasets: the Demographic Census of 2010 and the School Census of 1995. I use these datasets to construct measures of school quality and income at municipal level. In order to link income to school quality, I focus on the average income of individuals aged 22-34 in 2010. These are individuals who were attending primary and secondary school in 1995, the year for which I have data on school quality.

I use three different methodologies to observe the relationship between education quality and income inequality. First, I run the decomposition technique by population subgroups introduced by Shorrocks (1984) and subsequently developed by Cowell and Jenkins (1995). This method allows me to calculate the share of the overall inequality which can be explained by the differences in the average level of schooling quality across municipalities. Second, I use a regression-based approach, the Fields decomposition, to evaluate how much inequality can be accounted for by measured education quality. Finally, I implement the Oaxaca-Blinder decomposition to study in details the income differences between municipalities with different levels of measured education quality.

The results suggest that measured education quality contributes little to income inequality across Brazilian municipalities. The decomposition by population subgroups shows that between approximately 0.01 to 9.92 percent of income inequality can be explained by the differences in the quality of the public education provided across municipalities. Furthermore, the results from the Fields decomposition show that between 0.04% to 0.33% of income inequality is explained by measured education quality. In comparison, the region of residence and educational attainment respectively account for approximately 23% and 20% of income inequality. Similarly, the results from the Oaxaca-Blinder decomposition show no significant income gap between municipalities above and below the average pupil-teacher ratio index. Further, the income gap between municipalities above and below the average repetition rates index is only due to the composition effect, or the differences in the characteristics themselves.

There are several reasons that potentially explain the insignificant results. First, it is possible that the proxies of schooling quality are incorrect for Brazil. It would be prudent to explore the question using other relevant measures such as public spending per municipality, teacher salaries and test scores. Second, this research paper assesses the disparities in schooling quality compared to the average income level at the municipality level. The most common way to study the determinants of income inequality in Brazil is at the household level. There is a lack of data at this level, so I was unable to include this analysis in my paper. It is important to note that the question I try to answer is how the disparities in the average quality of the public education contribute to the variations of income earnings we observe between municipalities. This has important policy implications given that more than 50% of primary public schools were administrated and funded by municipal governments in 2009 (Bruns, Evans and Luque, 2012).

The literature on income inequality in Brazil is extensive; however, Lusting, Lopez-Calva and Ortiz-Juarez (2012) claimed that “the decline in inequality cannot be taken for granted” (p.138) and it is important to study the evolution of inequality and its determinants. Further, Bruns, Evans and Luque (2012) argued that although access to education is important, its quality is the linchpin.

The paper is structured as follows: Section 2 provides a general background of the educational system in Brazil and discusses its relationship with income inequality. Section 3 presents the two main datasets used and describes the procedure to construct an index for education quality. Section 4 develops the three methods used and the results are detailed in Section 5. Section 6 concludes.

Chapter 2.

Background

2.1. Education in Brazil

The educational system in Brazil is divided into two main levels: basic education (which includes pre-school, primary and secondary education) and higher education. Public education is free and is administrated either by the municipal, state or federal government. Since the 1990s, there have been considerable efforts and policy changes in order to improve schooling access and quality. For example, in 1996 the government established the Fund for Maintenance and Development of the Fundamental Education and Valorization of Teaching (FUNDEF). This act obliged municipalities and states to spend a minimum amount per student, which had significant impact on enrollments and teachers' wages, especially in the northern, northeastern and central western regions of the country (Bruns, Evans and Luque, 2012). In 1995, the National Institute of Studies and Educational Researches (INEP) developed the National Assessment of Basic Education (SAEB) to assess the performance of the students on standardized tests in mathematics and in Portuguese. Finally, in 2001 the government introduced the program *Bolsa Escola*[4], a conditional cash transfer program to provide financial aid to poor households who send their children to school. This program has greatly contributed to the access to education in Brazil. Barros et al. (2010) found it contributed to more than 13% of the overall decline of income inequality from 2001 to 2007.

A major issue concerning education in Brazil is that there exists a great disparity in the quality of the education between public schools and private schools at the primary and secondary levels. Private schools offer a better quality education overall; while

[4] It became the *Bolsa Familia* in 2003.

approximately 80% of the children coming from rich households attend private schools, only 3.5% of the children from poor households attend these institutions (Curi and Menezes-Filho, 2010). Public schools are free and funded by the government, but they do not provide an appropriate preparation for the public university entrance examination. Consequently, the children of disadvantaged classes who are often limited to public education will usually not satisfy the requirements of the high-ranked and funded public universities and will have to pay more for a higher private education (Akkari, 2013). This highlights the importance of studying the impact of education quality on the economic development of Brazil.

2.2. Education and Income Inequality in Brazil

The literature on income inequality and education in Brazil has identified two effects that explain how the educational system contributes to income inequality: the price effect and the quantity effect. The former refers to the return to education. According to Barros et al. (2010), the fall in the return to education would have contributed to 35 percent of the fall of inequality in Brazil. This is caused by an increase in the relative supply of skilled workers and a decrease in the number of low-skilled or unskilled workers (Lusting, Lopez-Calva and Ortiz-Juarez, 2013). Some of the literature discusses the convexity of returns to education and claims that the first years of schooling and tertiary education have the greatest impact on labour earnings. This is due to a concave correlation between educational attainment and labour earnings in the first years of education that later becomes convex with higher education (Barros et al., 2010; Blom, Holm-Nielsen and Verner, 2011; Battiston, Garcia-Domench and Gasparini, 2014).

The quantity effect is another channel through which education affects the distribution of labour earnings. Barros et al. (2010) defines it as the increase in the inequality in labour earnings due to the increase in workers' education. The effect of this is debated in the literature. Battiston, Garcia-Domench and Gasparini (2014) found that the increase in years of education in the 1990s and the 2000s increased inequality, while Azevedo et al. (2013) had the completely opposite conclusion. Furthermore, Ferreira and Velez (2004) claimed that educational attainments vary widely by race and the years of schooling, and it accounts for 29 percent of excess inequality in Brazil.

2.3. Education Quality and Earnings

While there is great certainty about the relationship between educational attainment and income earnings, there is no general consensus concerning the relationship between education quality and earnings. As Hanushek (2002) claimed, “[t]he economics effects of differences in the quality of graduates of the elementary and secondary schools are less understood than the effects of quantity” (p.13). Bedi and Edwards (2002) provide a complete discussion of the reasons that explain why some empirical research papers find that education quality positively affects earnings and others find no significant relationship. Some examples include Card and Krueger (1992) who studied the relationship between the quality of education and income earnings in the United States. Using US Census data, they showed that men educated in states with higher quality education (measured by the pupil-teacher ratio, the average term length and the relative pay of teachers) have a higher return to an additional year of school. On the contrary, Psacharopoulos and Velez (1993) surprisingly found that high repetition rates in all levels of schools in Columbia are not associated with lower income earnings. In the context of Brazil, Behrman and Birdsall (1983) studied the return of education by introducing the quality of schooling (measured by the average schooling of teachers) and found that it has a positive impact on earnings.

Chapter 3.

Data

Two major datasets are used to study the relationship between education quality and income inequality: the Demographic Census of 2010 and the School Census of 1995. One advantage of using the 1995 Census is that the major federal policy changes to improving a minimum level of schooling quality occurred after 1995.[5] We can therefore expect to observe disparities in the level of school quality across municipalities.

The Brazilian National Household Sample Survey (PNAD) is conducted annually by the Brazilian Institute of Geography and Statistics (IBGE). This survey collects general information on education, labor, income, housing and other characteristics of Brazilian households. Once every ten years, the PNAD is replaced by the Demographic Census where more households are interviewed.[6] This survey includes the municipality where each individual lives and has lived. This allows me to construct a new dataset at the municipality level where I include only individuals who have always lived in the same municipality and a variety of variables such as average income, proportion of mixed populations[7], proportion of rural areas, average educational attainment, and proportion of different economic activities.[8]

[5] FUNDEF/FUNDEP started in 1996, the assessment of the performance of the students started in 1995 with the SAEB and the Bolsa Escola/Bolsa Familia in 2001.

[6] In 2010, the Demographic Census interviewed 67.6 million of household units in 5,565 different municipalities of the country as opposed to 153,837 households in the 2009 PNAD.

[7] The IBGE divides race into five categories: White, Black, Asian, Indigenous and *Pardo*, which can be defined as the Brazilian mixed race. I will use the term ‘mixed populations’ throughout the paper as a reference to *Pardo*.

[8] See Appendix A for a complete description of the variables selected.

Since 1995, the National Institute of Studies and Educational Researches (INEP) has collected data on schools all over the country. The School Census provides general information about school location, whether the school is public or private, teachers' level of education, number of classrooms, workers, and teachers, enrollments at the pre-school, primary and secondary levels, number of repeaters, etc. Because the students in pre-school are mostly between 4 and 7 years old (which makes them between 19 and 22 years old in 2010), I keep only the schools that have at least a primary and/or secondary level. Given that university usually begins when people are 18 years old and that it takes between 3 to 6 years to complete a bachelor degree, it is less likely that pre-school students were on the job market in 2010. According to the School Census, the students in primary school are typically between 7 and 15 years old and between 15 and 19 years old in the secondary school. This means that in 2010, those individuals are mostly between 22 and 34 years old. Finally, the School Census 1995 includes 243,637 schools from 4,975 different municipalities.

I combine the two datasets by matching all the municipalities that are present in both datasets. The reason why I use income inequality at the municipality level rather than at the individual or household level is because the available dataset does not allow me to assign a value for education quality to each individual or household.

3.1. Measured Education Quality

From the School Census, I use two sets of variables that can be used as a proxy for education quality: the pupil-teacher ratio and the repetition rates.

Pupil-Teacher Ratio

As Card and Krueger (1992) claimed, we can hypothesize that a reduction in the pupil-teacher ratio increases the quality of education received. Some empirical studies have suggested that a smaller class size increases the performance of the students to test scores.[9]

For 4,974 municipalities, I construct the pupil-teacher ratio for three groups: levels 1 to 4 and levels 5 to 8 in primary schools and all levels in the secondary school. This ratio represents the number of students enrolled divided by the number of teachers. Some observations are worth mentioning. First, there are considerable differences in the average pupil-teacher ratio between the regions. This is consistent with many research papers that found that region is an important determinant of income inequality in Brazil. Second, there are important differences according to the type of schools (i.e. public and private schools). In particular, the pupil-teacher ratio is smaller in private schools, where the level of schooling quality is greater (Akkari, 2013). Those observations are illustrated in Table B1 of the Appendix B.

Repetition Rates

Although the repetition rates can be seen as the result of school and governmental policies, they can also be used as a measure for educational outcomes. Gomes and Hanushek (1994) used data from the northeastern schools in Brazil and found that low performance on Portuguese and mathematics standardized tests significantly increased the probabilities of repetitions which have a direct impact on students' achievement. This arguably demonstrated that education quality is an important determinant of repetition rates. Furthermore, Lee and Barro (2001) discuss how UNESCO frequently uses repetition rates to compare the efficiency of the school system across countries.

The School Census lists three types of repeaters: those who failed in 1994, those who abandoned in 1994 and those who are enrolled in the same level than the previous year despite passing that level. I define the repetition rate as the number of repeaters in a given level divided by the number of students enrolled on that same level. Only levels 1 to 4 in primary schools from 1,729 municipalities are included and I did not consider

[9] See Lee and Barro (2001) for a brief review of this literature.

repetition rates in secondary school due to insufficient observations. Again, we observe in Table B1 of Appendix B a considerable difference between regions and types of schools. I also provide in Appendix B some descriptive statistics of the average income of municipalities below and above the average of each measure for education quality. Except for the pupil-teacher ratio in secondary school, I find that the average income is higher for municipalities with a lower pupil-teacher ratio and lower repetition rates.

3.2. Index for Education Quality

From the School Census, I collected three different variables for the pupil-teacher ratio and four different variables for the repetition rates, both according to the levels of schooling. I do not use all the variables in the same regression because of multicolinearity between the variables.[10] However, because they all provided relevant information to the measurement of education quality, I followed Filmer and Pritchett (2001) and used the principal component analysis procedure to construct an index for education quality. "[T]he first principal component of a set of variables is the linear index of all the variables that captures the largest amount of information that is common to all of the variables" (Filmer and Pritchett, 2001, p.116). Given a set of N variables, a_{1i}^* to a_{Ni}^* represent *N* proxies for education quality in a municipality i. I first specify each variable normalized by its mean and standard deviation: $a_{1i} = \frac{(a_{1i}^* - a_1^*)}{s_1^*}$ where a_1^* is the mean of a_{1i}^* across all municipalities and s_1^* is its standard deviation. Each normalized selected variable can be expressed as a linear combination of a set of underlying components $A_{1i}, \dots, A_{Ni}$ for each municipality i. By inverting this system, we can find the first principal component A_{1i} which represents the index for measured education quality for each municipality i: $A_{1i} = f_{11} * \frac{(a_{1i}^* - a_1^*)}{s_1^*} + \cdots + f_{1N} * \frac{(a_{Ni}^* - a_N^*)}{s_N^*}$ where $f_{11}, \dots, f_{1N}$ are the scoring factors. In Appendix C, I present the scoring factors for the index of the pupil-teacher ratio and the repetition rates.

[10] See Appendix C for the correlations between the variables

Chapter 4.

Methodology

To observe how the differences in measured schooling quality across municipalities predict levels of income inequality, I implement three different methodologies. First, the decomposition technique by population subgroups is one of the most common and simple methodologies used to study the determinants of income inequality. [11] It allows me to quantify the impact of measured education quality on income inequality. Regression-based approaches are also common in the literature as they measure how the different levels of inequality can be explained by each selected factor (Litchfield, 1999). In this paper, I use the Fields decomposition. Finally, the Oaxaca-Blinder decomposition allows one to examine the mean income gap between two groups.

4.1. Decomposition by Population Subgroups

The decomposition techniques were first described by Bourguignon (1979), Cowell (1980) and Shorrocks (1980, 1984). I use the static decomposition method developed by Cowell and Jenkins (1995).[12] Ferreira, Leita and Litchfield (2008) used this methodology to study the determinants of inequality in Brazil from 1981 to 2004 to present the share of total inequality explained by seven selected variables (including

[11] Many authors often decompose inequality by factors components or income source to identify the contribution to inequality of any given component of income (Shorrocks, 1982, p. 193). This methodology is not relevant to assess the contribution of measured education quality.

[12] The dynamic decomposition is another method developed by Mookherjee and Shorrocks (1982). It allows decomposing inequality over a period of time. Because I have access to the Demographic Census for the year 2010 only, I cannot use this methodology. It will be relevant to implement this methodology when the next Demographic Census will be available.

educational attainment).[13] It is worth mentioning that this methodology only provides an approximate picture of how the explanatory variables account for inequality.

Cowell and Jenkins (1995) show that the overall inequality (I) for a given year is composed of the inequality between the chosen groups (I_B) (for example, the inequality between rural and urban areas), and the within-group inequality (I_W) (the inequality within individuals living in rural areas for instance).[14] More specifically, they showed that for a given specific partition of the population (Π), the overall inequality (I) is a function of within-group inequality (I_W) and between-group inequality (I_B) that can be simplified as $I = I_B + I_W$. They developed the R_B and the R_W summary measures that allow estimating the share of the overall inequality that is explained by the specified partition (Π): $R_B(\Pi) = \frac{I_B(\Pi)}{I}$ and $R_W(\Pi) = I - \frac{I_W(\Pi)}{I}$.

To evaluate how the differences between municipalities with varying levels of measured schooling quality explain the overall income inequality in Brazil, I use all the variables available from the School Census 1995 as partitions Π. I divide each partition into two different groups j where group 1 includes municipalities below the average value of the partition Π and group 2 includes municipalities above its average. I also look at other characteristics such as educational attainment, region, race and location. The definition of the groups that belong in these partitions is available in the municipality level section of Appendix A.

I provide the results using the Generalized Entropy (GE) measures of inequality because they can be easily decomposed into within ($GE_W(\alpha)$) and between ($GE_B(\alpha)$) inequalities, an important requirement for the decomposition by population subgroups

[13] This methodology was also used by Bourguignon, Ferreira and Leite (2008) and Salardi (2005) in the context of Brazil, Leibbrandt, Finn and Woolard (2012) in South Africa and Dickey (2001) for the United Kingdom.

[14] They are respectively defined as $I_B = \frac{1}{\alpha^2-\alpha}\left[\sum_{j=1}^{k} f_j \left(\frac{\bar{y}_j}{\bar{y}}\right)^{\alpha} - 1\right]$, $I_W = \sum_{j=1}^{k} w_j GE(\alpha)_j$ and $w_j = v_j^{\alpha} f_j^{1-\alpha}$ where $GE(\alpha)$ is the measure of inequality, f_j is the municipality share and v_j the income share of each group j in the partition Π where $j = 1, 2, \dots, k$ (Litchfield, 1999).

(Litchfield, 1999).[15] The GE measures are defined as $GE(\alpha) = \frac{1}{\alpha^2-\alpha}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{\bar{y}}\right)^{\alpha} - 1\right]$ where y_i is the average income level in municipality $i, i \in (1, 2, \ldots, n)$, $\bar{y} = \frac{1}{n}\sum y_i$ is the arithmetic mean income and α takes the values 0, 1 or 2.[16]

4.2. Fields Decomposition

The regression-based approach allows one to assess the contribution of different explanatory factors to income inequality. Many authors have developed this as an alternative method to decomposition techniques in order to solve some difficulties and issues related to it such as the limitation to some measures of inequality (i.e. the General Entropy and the Atkinson measures), the use of discrete variables only, the arbitrary selection of subgroups, and finally the fact that it is not possible to control for endogeneity.

Fields (2003) developed a regression-based method to assess income inequality at a given point in time. This methodology is applicable to any inequality measures that are continuous and symmetric.[17]

Fields' (2003) method allows answering the question: "How much income inequality is accounted for by each factor [j]?" (p.1). In other words, this will allow me to observe how much income inequality can be accounted for by measured education quality. Basically, we begin with a typical income-generating function $\ln Y_i = X_i\beta + \varepsilon_i$ where X represents the explanatory variables. These are a dummy for the region in

[15] The Atkinson inequality measures could also be used as Cowell and Jenkins (1995) did to decompose inequality by population subgroups in the United States. However, the General Entropy measures are most commonly used in the context of Brazil and the results using the Atkinson measures were similar.

[16] Using the I'Hopital's rule, when $\alpha = 0$, GE(0) becomes the mean log deviation: $GE(0) = \frac{1}{n}\sum_{i=1}^{n}\log\frac{\bar{y}}{y_i}$ and when $\alpha = 1$, GE(1) becomes the Theil-T Index: $GE(1) = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\bar{y}}\log\frac{y_i}{\bar{y}}$. Finally, when $\alpha = 2$, then $GE(2) = \frac{1}{2n\overline{y^2}}\sum_{i=1}^{n}(y_i - \bar{y})^2$ or the half of the square of the Coefficient of Variation (CV).

[17] Fields (2003) shows that the results are independent of the measure of inequality used, which is a major advantage compared to decomposition technique.

which municipality i belongs, the proportion of mixed populations, the proportion of rural areas, the proportion of some economic activities, the average educational attainment and the index for measured education quality. By taking the variance on both sides of the equation, we can express the share of log-variance of income explained by the j^{th} explanatory factor as:

$$s_j = \frac{\beta_j * \sigma(X_j) * corr[X_j, \ln Y]}{\sigma(\ln Y)}$$

4.3. Oaxaca-Blinder Decomposition

The Oaxaca-Blinder decomposition was first developed to quantitatively assess the sources in the male-female wage gap. More generally, this method allows one to compare the differences in the mean of an outcome variable Y between two groups A and B. Those groups can be for example two regions, periods in time or levels of schooling quality.[18] In this particular case, I look at the mean log-income gap between municipalities that have inferior measured education quality (above the index average) (group A) and municipalities that have better measured education quality (below the index average) (group B).

The model begins with a semi-linear income function $\ln(Y_{ig}) = X'_i\beta_g + \epsilon_{gi}$, $g = A, B$ where $\ln(Y_{ig})$ represents log-income of a municipality i in group g, X_i represents a set of characteristics and we assume $E(\varepsilon_{gi}|X_i) = 0$. We then take the overall difference in the mean income between the two groups which becomes $\overline{\ln(Y_B)} - \overline{\ln(Y_A)} = \overline{X_B}'\widehat{\beta_B} - \overline{X_A}'\widehat{\beta_A}$. By adding and subtracting the term $\overline{X_B}'\widehat{\beta_A}$ which represents the estimates of the

[18] For example, it has been used in the context of Brazil by Bourguignon, Ferreira and Leite (2008) to study the Brazil's excess inequality compared to the United States. The authors found that the distribution and the returns of education accounted for a large part of excess inequality.

output municipalities in group B would get if their average income was determined according to the returns for municipalities in group A, we get the following equation:[19]

$$\overline{\ln(Y_B)} - \overline{\ln(Y_A)} = \overline{X_B}'(\widehat{\beta_B} - \widehat{\beta_A}) + (\overline{X_B} - \overline{X_A})'\widehat{\beta_A}$$

The first term on the right hand side represents the “unexplained” effect or the “income structure” effect and the second term represents the “explained” effect (Fortin, Lemieux, Firpo, 2011). The first term is often referred as discrimination and the second term refers to the composition effect or in other words, the differences in the means associated with the differences in the characteristics themselves (Bourguignon et al., 2008). This methodology allows me to first observe whether there is an income gap between municipalities with different levels of measured education quality. If so, I can specifically look whether it is explained by the differences in the characteristics between those two groups or whether there is some unexplained effect.

[19] We could also have $\overline{\ln(Y_B)} - \overline{\ln(Y_A)} = \overline{X_A}'(\widehat{\beta_B} - \widehat{\beta_A}) + (\overline{X_B} - \overline{X_A})'\widehat{\beta_B}$ by adding and subtracting the term $\overline{X_A}'\widehat{\beta_B}$. In this case, Group B’s coefficients are used as the reference coefficients. This is usually used in the gender wage gap literature when discrimination is believed to be towards women (group A) (Jann, 2008).

Chapter 5.

Results

5.1. Results from the Decomposition by Population Subgroups

I first provide in Appendix D some general descriptive statistics of the average income and an overview of income inequality between and within the municipalities. I observe whether some determinants of income inequality at the municipality level are important at the individual level and compare the contribution of some characteristics that have been studied in the literature. Most research papers usually find that the most important determinant of income inequality in Brazil is educational attainment, which is also what I find. Other important determinants are the region where the municipality is located and the primary race in the municipality.

The results from the decomposition by population subgroups show that the disparities in the measured education quality between municipalities explain a small percentage of the overall income inequality. The results are illustrated in Table 1. Depending on the measure for education quality and the measure for inequality used, the differences in the mean income of municipalities with lower school quality (i.e. municipalities above the average pupil-teacher ratio, repetition rates and indexes) compared to municipalities with higher schooling quality (i.e. municipalities below the averages) account for between 0.01% to 9.92% of the overall inequality. I show in Appendix E that the choice of partition does not have a significant impact on the results and that the percentages found remain small compared to other partitions, such as educational attainment, region and race (see Appendix D).

Table 1. The Percentage of Income Inequality Explained by the Differences in Measured Education Quality at the Municipality Level

	Inequality Measures		
Measures of Education Quality	**GE(0)**	**GE(1)**	**GE(2)**
Pupil-teacher ratio			
Levels 1 to 4	0.172%	0.180%	0.173%
Levels 5 to 8	0.893%	0.935%	0.895%
Secondary school	0.782%	0.837%	0.820%
Index *Pupil-teacher ratio*	0.013%	0.014%	0.014%
Repetition Rates			
Level 1	9.921%	8.921%	7.918%
Level 2	3.037%	2.938%	2.603%
Level 3	2.790%	2.707%	2.405%
Level 4	4.636%	4.471%	3.952%
Index *Repetition rates*	6.648%	6.422%	5.693%

Note. Author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995. Each measure of education quality (or partition Π) is divided into two groups. Group 1 includes municipalities below the average value of the variable and group 2 includes municipalities above the average. I report the R_B summary measure (Cowell and Jenkins, 1995) for three measures of inequality: GE(0), GE(1) and GE(2) where $R_B = GE_B(\alpha)/GE(\alpha)$.

Given the small percentages found, especially in the case of the pupil-teacher ratio index, it seems crucial to provide more evidence about the significance of the results. The Fields decomposition will shed more light on this question.

5.2. Results from Fields Decomposition

Table 2 presents the results for the OLS regression using the two indexes constructed as a measure for education quality. As we can see, measured education quality has little or no impact on income. In columns (3) and (4), I exclude the variable educational attainment as it is likely to be correlated with measured education quality and I find similar results. Only the index for the pupil-teacher ratio has a significant explanatory power. However, while one would expect that a greater pupil-teacher ratio lowers education quality and therefore tends to lower income, this is not what we observe. It appears that the pupil-teacher ratio has a positive impact on log-earnings. One possible explanation for this unexpected result could be that populated municipalities such as Rio de Janeiro and São Paulo have higher average earnings are also more likely to have a high pupil-teacher ratio. This could drive this positive result.

Table 2. OLS Regression Estimates for Log of Average Monthly Gross Income of Municipalities

Explanatory Variables	Log income (1)	Log income (2)	Log income (3)	Log income (4)
Region				
North	(omitted)	(omitted)	(omitted)	-0.095*** (0.032)
Northeast	-0.291*** (0.017)	-0.252*** (0.016)	-0.322*** (0.019)	-0.357*** (0.027)
Southeast	-0.014 (0.018)	-0.014 (0.021)	-0.004 (0.020)	-0.103*** (0.024)
South	0.074*** (0.021)	0.065** (0.031)	0.114*** (0.023)	(omitted)
Central West	0.144*** (0.020)	0.155*** (0.022)	0.153*** (0.022)	0.078*** (0.028)
Race				
Proportion of Mixed Populations	-0.454*** (0.026)	-0.463*** (0.047)	-0.620*** (0.027)	-0.526*** (0.049)
Proportion of Rural Areas	-0.252*** (0.026)	-0.398*** (0.041)	-0.429*** (0.027)	-0.498*** (0.041)
Proportion of Economic Activity				
Agriculture	0.231*** (0.047)	0.181** (0.076)	0.157*** (0.051)	0.073 (0.078)
Industry	0.885*** (0.051)	1.313*** (0.105)	1.161*** (0.054)	1.445*** (0.107)
Trade	1.040*** (0.100)	1.701*** (0.182)	1.656*** (0.105)	2.303*** (0.178)
Public Administration	-0.114 (0.142)	0.714*** (0.222)	0.940*** (0.146)	1.376*** (0.220)
Construction	0.572*** (0.143)	1.579*** (0.218)	0.524*** (0.155)	1.657*** (0.226)
Health, Education and Social Services	0.429*** (0.152)	1.078*** (0.249)	1.896*** (0.152)	1.971*** (0.241)
Educational Attainment				
Average	0.412*** (0.016)	0.279*** (0.027)	-	-
Education Quality				

Pupil-teacher ratio Index	0.015*** (0.003)	-	0.024*** (0.003)	-
Repetition rates Index	-	-0.0007 (0.0035)	-	-0.004 (0.004)
Constant	4.948*** (0.062)	5.124*** (0.099)	6.244*** (0.037)	6.150*** (0.051)
Number of observations	3,817	1,533	3,817	1,533
F-Stat	1,364.11	560.02	1,215.84	556.48
Adj. R^2	0.8334	0.8363	0.8054	0.8250

Note. Author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995. Standard errors are in parentheses. The dependent variable is defined as the average monthly gross income of individuals aged 22-34 years old living in municipality i in the year 2010. Columns (1) and (3) use the pupil-teacher ratio index as measure for education quality and columns (2) and (4) use the repetition rates index. Columns (3) and (4) do not control for the average educational attainment of the municipality. ***Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

Table 3 displays the results from the Fields decomposition and shows that measured education quality does not have a great impact on income inequality. This is consistent with the results from the decomposition by population subgroups. In fact, the results show in columns (1) and (2) that only 0.04% to 0.05% of income inequality is explained by measured education quality as opposed to 15.79% to 24.36% for the average educational attainment. In columns (3) and (4), I exclude the variable for educational attainments and find that 0.06% to 0.33% of income inequality is accounted for by measured education quality.

Table 3. The Contribution of Each Explanatory Factor to Income Inequality

Selected Factors	Inequality Shares s_j (1)	Inequality Shares s_j (2)	Inequality Shares s_j (3)	Inequality Shares s_j (4)
Region				
North	(omitted)	(omitted)	(omitted)	0.34
Northeast	19.62	16.47	21.68	23.37
Southeast	-0.36	-0.29	-0.11	-2.09
South	2.80	1.92	4.29	(omitted)
Central West	0.99	2.93	1.05	1.48
Race				
Proportion of Mixed Populations	15.48	11.58	19.97	13.14
Proportion of Rural Areas	6.01	12.16	10.22	15.21

Proportion of Economic Activity				
Agriculture	-1.49	-1.67	-1.01	-0.67
Industry	9.56	9.61	12.54	10.57
Trade	5.57	11.33	8.88	15.35
Public Administration	-0.03	0.36	0.26	0.70
Construction	0.26	1.96	0.24	2.06
Health, Education and Social Services	0.57	1.57	2.53	2.87
Educational Attainments				
Average	24.36	15.79	-	-
Education Quality				
Pupil-teacher Index	0.04	-	0.06	-
Repetition Rates Index	-	0.05	-	0.33
Residuals	16.60	16.22	19.39	17.35

Note. Author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995. I report the share of inequality $s_j = \beta_j * \sigma(X_j) * corr[X_j, lnY]/\sigma(lnY)$ that can be accounted for by each explanatory factor j based on the estimates in Table 2. Columns (1) and (3) use the pupil-teacher ratio index as measure for education quality and columns (2) and (4) use the repetition rates index. Columns (3) and (4) do not control for the average educational attainment of the municipality.

5.3. Results from Oaxaca-Blinder Decomposition

The Oaxaca-Blinder decomposition allows one comparing the mean income between two groups. I observe the differences in the mean income between municipalities that are below and above the average measured education quality. The former (group B) is defined as better education quality and the latter (group A) as inferior education quality. Tables 4 and 5 provide the means and OLS regression estimates for the decomposition when the municipalities with inferior education quality (group A) are used as the reference income structure.

Some observations are worth mentioning. In both tables, we observe in column (1) that municipalities with an inferior measured education quality have a higher proportion of mixed populations. In particular, looking at the differentials between the estimates in columns (2) and (3) from Tables 4 and 5, an increase in the proportion of mixed people decreases the income by respectively 0.6 and 15.0 percent more in municipalities with better measured education quality compared to municipalities above the average index. We can also observe that the average educational attainment is higher in municipalities with inferior measured education quality for the pupil-teacher ratio (Table 4) and higher in municipalities with better measured education quality in the case of repetition rates (Table 5).

Table 4. Means and OLS Regression Estimates of Selected Variables for Pupil-Teacher Ratio Index

Explanatory Variables	Means (1)		Better Educ. Quality Estimates (2)	Inferior Educ. Quality Estimates (3)
Education Quality	Better	Inferior		
Regions				
North	-	-	-0.193*** (0.033)	(omitted)
Northeast	-	-	-0.483*** (0.019)	-0.281*** (0.020)
Southeast	-	-	-0.187*** (0.017)	-0.023 (0.023)
South	-	-	-0.091*** (0.020)	-0.067** (0.028)
Central West	-	-	(omitted)	0.099*** (0.026)
Race				
Proportion of Mixed	0.391	0.469	-0.467*** (0.034)	-0.461*** (0.039)
Proportion of Rural Areas	0.346	0.246	-0.189*** (0.033)	-0.357*** (0.041)
Educational Attainments				
Average	3.352	3.411	0.368*** (0.022)	0.402*** (0.024)
Constant			5.141*** (0.080)	5.092*** (0.091)
Dependent Variable (Log Monthly Income)	6.267	6.270		
Number of observations	2,214	1,603	2,214	1,603

Adj. R^2			0.816	0.869

Note. Author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995. Better education quality (group B) includes the 2,214 municipalities with an index value below the mean of the pupil-teacher ratio index. Inferior education quality (group A) includes the 1,603 municipalities with an index value above its mean. Column (1) reports the mean of the selected explanatory variables that is used to compute the composition effect in Table 6. Columns (2) and (3) respectively report the estimates of the OLS regression for municipalities with better education quality and municipalities with inferior education quality. The explanatory variables for economic activity were excluded from the table. Standard errors are in parentheses. ***Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

Table 5. Means and OLS Regression Estimates of Selected Variables for Repetition Rates Index

Explanatory Variables	Means (1)		Better Educ. Coef. (2)	Inferior Educ. Coef (3)
Education Quality	Better	Inferior		
Regions				
North	-	-	(omitted)	-0.139*** (0.040)
Northeast	-	-	-0.305*** (0.022)	-0.337*** (0.036)
Southeast	-	-	-0.030 (0.028)	-0.147*** (0.040)
South	-	-	0.026 (0.039)	-0.012 (0.054)
Central West	-	-	0.142*** (0.028)	(omitted)
Race				
Proportion of Mixed	0.515	0.602	-0.500*** (0.061)	-0.350*** (0.077)
Proportion of Rural Areas	0.319	0.391	-0.353*** (0.055)	-0.438*** (0.061)
Average Educational Attainments	3.291	3.123	0.306*** (0.037)	0.231*** (0.040)
Constant			5.201*** (0.132)	5.344*** (0.152)
Dependent Variable (Log Monthly Income)	6.136	5.903		
Number of	906	627	906	627

observations			
Adj. R^2		0.850	0.785

Note. Author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995. Better education quality (group B) includes the 2,214 municipalities with an index value below the mean of the repetition rates index. Inferior education quality (group A) includes the 1,603 municipalities with an index value above its mean. Column (1) reports the mean of the selected explanatory variables that is used to compute the composition effect in Table 6. Columns (2) and (3) respectively report the estimates of the OLS regression for municipalities with better education quality and municipalities with inferior education quality. The explanatory variables for economic activity were excluded from the table. Standard errors are in parenthesis. ***Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

Table 6 provides the results of the decomposition for the composition effect and the income structure effect. Consistent with the results from the decomposition technique by population subgroups and Fields decomposition, the income gap is either insignificant (column (1)) or related to the differences in the characteristics themselves (column (2)). Indeed, in column (2), the composition effect accounts for 0.2235 log points out of the 0.2337 average income gap between municipalities with better measured education quality and municipalities with inferior measured education quality. Only an insignificant 0.0101 part of the gap (income structure effect) is left unexplained (Fortin, Lemieux, Firpo, 2010).

Table 6. Measured Education Quality Income Gap: Oaxaca-Blinder Decomposition Results

Reference Group:	Using Inferior Education Coef. from col. 2, Table 4 Pupil-Teacher Ratio (1)	Using Inferior Education Coef. from col. 2, Table 5 Repetition Rates (2)
Mean log income gap	-0.0031 (0.0161)	0.2337*** (0.0239)
Composition effect (explained)	-0.0064 (0.0157)	0.2235*** (0.0221)
Income structure effect (unexplained)	0.0033 (0.0082)	0.0101 (0.0119)

Note. Author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995. Column (1) provides the results using municipalities above the average pupil-teacher ratio index (group A) as the reference group and column (2) provides the results using municipalities above the average repetition rates index (group A) as the reference group. Standard errors are in parenthesis. ***Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

Chapter 6.

Conclusion and Discussion

In this research paper, I was interested in examining if education quality is a relevant and important determinant of income inequality in Brazil. From the literature, we know that educational attainment is one of the most important determinants of income inequality across households. We also observe varying levels of schooling quality. I combined the School Census of 1995 and the Demographic Census 2010 to study the relationship between education quality (measured by the pupil-teacher ratio and the repetition rates) and the disparities between average incomes across municipalities. Table 7 summarizes the findings from the three methodologies used.

Table 7. Summary of the Results

	Decomposition by Population Subgroups[a]	**Fields Decomposition**[b]	**Oaxaca-Blinder Decomposition**[c]		
Measured Education Quality			Mean Income Gap	Composition Effect	Income Structure Effect
Pupil-Teacher Ratio	0.001% to 0.129%	0.04% to 0.06%	-0.0031	-0.0064	0.0033
Repetition Rates	0.325% to 1.191%	0.05% to 0.33%	0.2337***	0.2235***	0.0101
Educational Attainments	53.4% to 55.0%	15.79% to 24.36%	-	-	-
Region	51.0% to 58.1%	21.03% to 26.91%	-	-	-

[a] Results from Table 1 and Table D2 in Appendix D;

[b] Results from Table 3. The percentage for the explanatory factor *region* was calculated by adding the percentage for each region. Fields (2003) showed that the overall effect of a dummy variable can be found by taking the sum of the s_j;

[c] Results from Table 6.

Note. ***Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level

The results show that my measure of education quality is not an important determinant of income inequality between Brazilian municipalities. Educational attainment remains the most important factor of the variations between municipal average incomes. Where the municipality is located also explains a considerable percentage of income inequality.

There are some limitations related to this paper that are worth mentioning. In this research paper, I introduced education quality as a determinant of income inequality using the three following decomposition methods: the decomposition technique by population subgroups, the Fields decomposition and the Oaxaca-Blinder decomposition. Those methods specifically look at the disparities in the mean income. Fortin, Lemieux and Firpo (2011) suggest other methodologies that could also be useful. These methods provide decomposition for other distributional statistics such as the variance, quantiles and Gini coefficient. For example, a variance decomposition could provide interesting results by observing how the variance in measured schooling quality affects the income variations across municipalities. This could be more representative of the reality as some populated municipalities such as Rio de Janeiro and São Paulo have a higher than average income with considerable variations among the individuals living within those municipalities.

Another limitation to this research is that I do not control for internal migration. This slightly affects the interpretation of the results; measured education quality has little or no impact between the variations of the average income among individuals living in the same municipality compared to the one they were born in. Card and Krueger (1992) include a state-of-birth effect and a state-of-residence effect in their analysis of school quality and the return of education in the United States by assuming individuals attend school in their state of birth. Although this assumption would be harder to implement to study income inequality across municipalities, there is potential for future research to provide a more realistic picture of the relationship.

Finally and most importantly, it could be argued my proxies for education quality are incorrect. Data limitations restrict the use of other measures commonly used such as the expenditure per student, test scores, real teacher salaries, etc. The SAEB 1995 includes some of these variables for 2,839 public and private schools. Although there is no information about the municipality where those schools are located, combining the two 1995's censuses could provide additional and relevant information about schooling quality in Brazil. Future research on education quality in Brazil should construct a more complete index for education quality that includes a greater variety of measures. On a related matter, this education issue is not only a problem for Brazil; Chile is another country with high levels of income inequality, with similar disparity issues in education quality between public and private schools. It would be relevant to examine multiple datasets in different countries and see whether a better education quality index could be constructed to study the relationship between education quality and income inequality in general.

References

Akkari, A. (2013). Blurring The Boundaries Of Public And Private Education In Brazil. Journal of International Education and Leadership Volume, 3(1).

Azevedo, J. P., Inchauste, G., & Sanfelice, V. (2013). Decomposing the recent inequality decline in Latin America (No. 6715). The World Bank.

Barros, R., De Carvalho, M., Franco, S., & Mendonça, R. (2010). Markets, the State, and the Dynamics of Inequality in Brazil. Declining Inequality in Latin America: a Decade of Progress, 134-174.

Battiston, D., García, C., & Gasparini, L. (2014). Could an increase in education raise income inequality? Evidence for Latin America. Latin American Journal Of Economics, 51(1), 1-39.

Bedi, A. S., & Edwards, J. H. (2002). The impact of school quality on earnings and educational returns—Evidence from a low-income country. Journal of Development Economics, 68(1), 157-185.

Behrman, J. R., & Birdsall, N. (1983). The quality of schooling: quantity alone is misleading. The American Economic Review, 928-946.

Blom, A., Holm-Nielsen, L., & Verner, D. (2001). Education, earnings, and Inequality in Brazil, 1982-1998: implications for education policy. Peabody journal of education, 76(3-4), 180-221.

Bourguignon, F. (1979). Decomposable income inequality measures.Econometrica: Journal of the Econometric Society, 901-920.

Bourguignon, F., Ferreira, F. H., & Leite, P. G. (2008). Beyond Oaxaca–Blinder: Accounting for differences in household income distributions. The Journal of Economic Inequality, 6(2), 117-148.

Bruns, B., Evans, D., & Luque, J. (2011). Achieving world-class education in Brazil: The next agenda. World Bank Publications.

Card, D., & Krueger, A. B. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. The Journal of Political Economy, 100(1), 1-40.

Cowell, F. A. (1980). On the structure of additive inequality measures. The Review of Economic Studies, 521-531.

Cowell, F. A., & Jenkins, S. P. (1995). How much inequality can we explain? A methodology and an application to the United States.The Economic Journal, 421-430.

Curi, A. Z., & Menezes-Filho, N. A. (2010). Determinantes dos gastos com educação no Brasil. Pesquisa e Planejamento Econômico, 40(1).

Dickey, H. (2001). Regional earnings inequality in Great Britain: a decomposition analysis. Regional Studies, 35(7), 605-612.

Ferreira, F. H., & de Barros, R. P. (1999). The slippery slope: explaining the increase in extreme poverty in urban Brazil, 1976-1996 (Vol. 2210). World Bank Publications.

Ferreira, F. H. G., & Velez, C. E. (Eds.). (2004). Inequality and Economic Development in Brazil. Washington, DC, USA: World Bank Publications. Retrieved from http://www.ebrary.com

Ferreira, F. H., Leite, P. G., & Litchfield, J. A. (2008). The rise and fall of Brazilian inequality: 1981–2004. Macroeconomic Dynamics, 12(S2), 199-230.

Fields, G. S. (2003). Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States. Research in labor economics, 22(3), 1-38.

Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure Data—Or tears: An application to educational enrollments in states of india*. Demography, 38(1), 115-132.

Fortin, N., Lemieux, T., & Firpo, S. (2011). Decomposition methods in economics. Handbook of labor economics, 4, 1-102.

Gomes-Neto, J. B., & Hanushek, E. A. (1994). Causes and consequences of grade repetition: Evidence from Brazil. Economic Development and Cultural Change, 117-148.

Hanushek, E. A. (2002). Publicly provided education. Handbook of public economics, 4, 2045-2141.

Jann, B. (2008). A Stata implementation of the Blinder-Oaxaca decomposition.Stata Journal, 8(4), 453-479.

Jenkins, S. P. (2010). INEQDEC0: Stata module to calculate inequality indices with decomposition by subgroup. Statistical Software Components.

Lam, D. (1999). Generating Extreme Inequality: Schooling, Earnings, and Intergenerational Transmission of Human Capital in South Africa and Brazil. Research Report.

Lee, J. W., & Barro, R. J. (2001). Schooling quality in a cross–section of countries. Economica, 68(272), 465-488.

Leibbrandt, M., Finn, A., & Woolard, I. (2012). Describing and decomposing post-apartheid income inequality in South Africa.Development Southern Africa, 29(1), 19-34.

Litchfield, J. A. (1999). Inequality: methods and tools. Text for.

Lustig, N., Lopez-Calva, L. F., & Ortiz-Juarez, E. (2013). Declining inequality in Latin America in the 2000s: the cases of Argentina, Brazil, and Mexico. World Development, 44, 129-141.

Mookherjee, D., & Shorrocks, A. (1982). A decomposition analysis of the trend in UK income inequality. The Economic Journal, 886-902.

OECD (2014), Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en

Psacharopoulos, G., & Velez, E. (1993). Educational quality and labor market outcomes: evidence from Bogota, Colombia. Sociology of Education, 130-145.

Salardi, P. (2005). How much of Brazilian Inequality can we explain? An attempt of income differentials decomposition using the PNAD 2002. Universita di Pavia, Italy.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. Econometrica: Journal of the Econometric Society, 613-625.

Shorrocks, A. F. (1982). Inequality decomposition by factor components. Econometrica: Journal of the Econometric Society, 193-211.

Shorrocks, A. F. (1984). Inequality decomposition by population subgroups. Econometrica: Journal of the Econometric Society, 1369-1385.

The World Bank, World Development Indicators (2012). GINI Index (World Bank Estimtes) [Data file]. Retrieved from http://data.worldbank.org/indicator/SI.POV.GINI.

Appendix A.

Definition of the Variables Selected from the Demographic Census 2010

Variables	Description	
	Individual level	Municipality level
Income	Monthly gross income earned from labor and other sources	Average monthly gross income among individuals in each municipality
Educational Attainments	Individual are grouped into five different categories. This follows Ferreira, Leite and Litchfield (2008): • 1 if Illiterates or less than 1 year of schooling • 2 if Primary School (1 to 4 years of schooling) • 3 if Primary School (5 to 8 years of schooling) • 4 if Secondary School (9 to 11 years of schooling) • 5 if Tertiary Education (more than 12 years of schooling)	Average educational attainments among individuals. Grouped into eight categories: • [1;1.5[(0 municipalities) • [1.5;2[(11 municipalities) • [2;2.5[(77 municipalities) • [2.5;3[(1,169 municipalities) • [3;3.5[(2,487 municipalities) • [3.5;4[(1,669 municipalities) • [4,4.5[(149 municipalities) • [4.5;5] (3 municipalities)
Race	There are five different categories: • White • Black • Mixed • Asian • Indigenous	I first define the proportion of each race in a municipality and then define the most important race in each municipality : • White (2,459 municipalities) • Black (7 municipalities) • Mixed (3,039 municipalities) • Asian (0 municipalities) • Indigenous (35 municipalities)
Region	There are five regions in Brazil: • 1 if North • 2 if Northeast • 3 if Southeast • 4 if South • 5 if Central West	Same as at the individual level
Location	• Rural area • Urban area	Because each municipality is composed of urban and rural areas, I create three groups according to the proportion of individuals living in rural areas in each municipality: • [0, 0.4[(3,194 municipalities) • [0.4, 0.6[(1,348 municipalities) • [0.6, 1] (1,023 municipalities)

Appendix B.

Descriptive Statistics on Measured Education Quality

Table B1. Descriptive Statistics for Measured Education Quality: Comparison of the Averages between Regions and Public/Private Schools

Variables Measured Education Quality	Number of Municipalities	Regions					School Type	
		1	2	3	4	5	Public	Private
Pupil-teacher ratio Primary Schools Levels 1 to 4	4,967	22.9	23.0	24.7	16.3	20.7	21.9	20.3
Pupil-teacher ratio Primary Schools Levels 5 to 8	4,905	21.9	22.5	21.7	16.6	19.7	20.7	14.9
Pupil-teacher ratio Secondary School	3,831	21.4	14.4	14.6	13.7	13.5	14.5	12.9
Repetition Rates Primary Level 1	1,656	0.44	0.44	0.32	0.27	0.34	0.39	0.18
Repetition Rates Primary Level 2	1,640	0.39	0.33	0.33	0.24	0.24	0.31	0.15
Repetition Rates Primary Level 3	1,602	0.27	0.27	0.19	0.19	0.22	0.24	0.13
Repetition Rates Primary Level 4	1,562	0.26	0.23	0.15	0.19	0.18	0.21	0.12

Note. I first calculate the pupil-teacher ratios and repetition rates for each school of the School Census 1995. Then I collapse the variables at the municipality level. In this table, I provide the average pupil-teacher ratio and repetition rates among municipalities. Only public schools are included for the comparison between regions.

Table B2. Average Income of Municipalities Below and Above the Mean of Measured Education Quality

Variables for Measured Education Quality	Average Income of Municipalities	
	Below the mean	Above the mean
Pupil-teacher ratio Primary Schools Levels 1 to 4	569.15 (268.87)	545.77 (285.43)
Pupil-teacher ratio Primary Schools Levels 5 to 8	584.61 (277.09)	531.65 (270.36)
Pupil-teacher ratio Secondary School	569.33 (278.3)	619.74 (266.9)
Repetition Rates Primary Level 1	530.34 (266.20)	392.79 (194.71)
Repetition Rates Primary Level 2	499.04 (257.60)	419.03 (218.44)
Repetition Rates Primary Level 3	500.86 (259.87)	423.22 (217.52)
Repetition Rates Primary Level 4	508.88 (259.16)	415.18 (215.61)

Note. Author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995. Standard deviations are in parentheses.

Appendix C

Correlations and Scoring Factors – Education Quality Index

Table C1. Correlations between the Selected Variables from the Brazilian School Census 1995 to Measure Education Quality

Variables	Pupil-teacher ratio 1-4	Pupil-teacher ratio 5-8	Pupil-teacher ratio 2	Repetition Rates 1	Repetition Rates 2	Repetition Rates 3	Repetition Rates 4
Pupil-teacher ratio 1-4	1.00						
Pupil-teacher ratio 5-8	0.17	1.00					
Pupil-teacher ratio 2	0.14	0.14	1.00				
Repetition Rates 1	0.06	-0.03	-0.09	1.00			
Repetition Rates 2	0.03	0.00	0.00	0.42	1.00		
Repetition Rates 3	0.10	0.06	-0.01	0.34	0.57	1.00	
Repetition Rates 4	0.09	0.01	-0.05	0.29	0.47	0.57	1.00

Table C2. The Scoring Factors and Summary Statistics for the Pupil-Teacher Ratio

Variables	Scoring factors	Mean	Standard Deviations
Pupil-teacher ratio 1-4	0.5824	22.23	11.64
Pupil-teacher ratio 5-8	0.5864	20.89	12.35
Pupil-teacher ratio 2	0.5630	14.53	9.45

Note. Mean value of Index: 0.000;

Standard Deviation of Index: 1.137;

The percentage of the covariance explained by the first principal component is 43.07%;

The first eigenvalue is 1.292;

Following Filmer and Pritchett and author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995.

Table C3. The Scoring Factors and Summary Statistics for the Repetition Rates

Variables	Scoring factors	Mean	Standard Deviations
Repetition Rates 1	0.4141	0.383	0.205
Repetition Rates 2	0.5299	0.305	0.191
Repetition Rates 3	0.5432	0.237	0.179
Repetition Rates 4	0.5027	0.211	0.190

Note. Mean value of Index: 0.000;

Standard Deviation of Index: 1.529;

The percentage of the covariance explained by the first principal component is 58.43%;

The first eigenvalue is 2.337;

Following Filmer and Pritchett and author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995.

Appendix D

Descriptive Statistics and Results from the Decomposition by Population Subgroups with Other Characteristics

Table D1 provides descriptive statistics of the average income and the income inequality level among different groups. More details about the groups can be found in Appendix A. Table D2 shows the percentage of income inequality explained by the differences observed between groups at the individual and at the municipality levels. In other words, it represents the share of the overall inequality which can be explained by the characteristics listed below. Table D3 compares the results with previous literature.

Table D1. Descriptive Statistics of Municipal Average Income and Income Inequality

Characteristics	Description	Number of Municipality	Average Income	Income Inequality (Gini Coefficient)
All		5,565	546.73	0.27
Region	North	449	395.73	0.21
	North East	1,794	301.46	0.17
	South West	1,668	642.91	0.19
	South	1,188	805.76	0.17
	Central West	466	631.88	0.16
Location	Rural	1,023	411.00	0.33
	Mixed	1,348	435.70	0.28
	Urban	3,194	637.07	0.22
Race	Majority White	2,459	750.69	0.18
	Majority Mixed	3,039	384.35	0.23
	Majority Black	7	444.83	0.15
	Majority Indigenous	35	331.57	0.31
Average Educational Attainment	Group 1	0	-	-
	Group 2	11	163.74	0.21
	Group 3	77	256.98	0.20
	Group 4	1,169	301.38	0.18
	Group 5	2,487	478.89	0.22
	Group 6	1,669	789.94	0.14
	Group 7	149	1,048.16	0.12
	Group 8	3	1,026.31	0.06

Note. Author's calculations based on the Brazilian Demographic Census 2010. More details on the variables are available in Appendix A. The calculation of Gini Coefficient follows Jenkins (2010).

Table D2. The Percentage of Income Inequality Explained by the Differences in the Characteristics of the Individuals and Municipalities using the General Entropy Measures of Inequality

	Inequality Measures					
Characteristics	**GE(0)**		**GE(1)**		**GE(2)**	
	Individual	Municipality	Individual	Municipality	Individual	Municipality
Educational Attainments	23.39%	53.78%	21.24%	55.01%	4.51%	53.44%
Region	9.81%	58.07%	7.71%	56.14%	1.36%	51.06%
Urban/Rural	7.37%	15.34%	5.44%	15.60%	0.90%	14.50%
Race	7.58%	43.51%	6.39%	45.23%	1.20%	43.49%

Note. Author's calculations based on the Brazilian Demographic Census 2010. Each partition Π is divided into different groups described in Appendix A. I report the R_B summary measure at the individual and municipality levels (Cowell and Jenkins, 1995) for three measures of inequality: GE(0), GE(1) and GE(2) where $R_B = GE_B(\alpha)/GE(\alpha)$.

Table D3. Comparison with Previous Literature from the Results of the Decomposition Technique by Population Subgroups

Characteristics	**Municipality Level**[a]	**Individual Level**[a]	**Ferreira, Leite and Litchfield (2008)**[b]	**Salardi (2005)**[b]
Year	2010	2010	2004	2002
Educational Attainments	53.4 – 55.0%	21.2 – 23.4%	35 – 38 %	24 -32 %
Region	51.0 – 58.1%	7.7 – 9.8 %	8 – 10 %	7 – 10 %
Urban/Rural	14.5 – 15.6%	5.4 – 7.4%	5 – 7 %	6 – 8 %
Race	43.5 – 45.2%	6.4% - 7.6%	11 – 12 %	10 – 13 %

[a] Results from Table D2. Only GE(0) and GE(1) are reported at the individual level;
[b] Results are at the household level.
Note. This table shows that the main determinants of income inequality at the individual, household and municipality levels are *educational attainment*, *region* and *race*.

Appendix E

Results from the Decomposition by Population Subgroups using Different Partitions

Table E1. Results from the Decomposition Technique by Population Subgroups using Different Partitions

	Partitions with GE(0)			
Measures of Education Quality	**Below and above the average (1)**	**Below and above the median (2)**	**Quartiles (3)**	**Deciles (4)**
Pupil-teacher ratio				
Levels 1 to 4	0.172%	0.547%	2.503%	0.590%
Levels 5 to 8	0.893%	0.000%	1.677%	0.324%
Secondary school	0.782%	0.000%	1.146%	0.280%
Index Pupil-teacher ratio	0.013%	0.001%	0.610%	0.948%
Repetition Rates				
Level 1	9.921%	9.680%	11.493%	12.301%
Level 2	3.037%	0.000%	3.552%	4.118%
Level 3	2.790%	0.000%	3.654%	3.984%
Level 4	2.790%	1.610%	5.041%	6.254%
Index Repetition rates	6.648%	6.566%	8.320%	9.369%

Note. Author's calculations based on the Brazilian Demographic Census 2010 and School Census 1995. I report the R_B summary measure (Cowell and Jenkins, 1995) for the GE(0) measure of inequality: where $R_B = GE_B(0)/GE(0)$. Each measure of education quality (or partition Π) is divided into different groups. In column (1), I define group 1 as municipalities below the average value of the variable and group 2 as municipalities above the average. In column (2), municipalities are rather grouped whether they are below or above the median. Column (3) reports the results when municipalities are divided into quartiles and into deciles for column (4).