

Curi, Andréa Zaitune; Menezes-Filho, Naercio

Article

The relationship between school performance and future wages in Brazil

EconomiA

Provided in Cooperation with:

The Brazilian Association of Postgraduate Programs in Economics (ANPEC), Rio de Janeiro

Suggested Citation: Curi, Andréa Zaitune; Menezes-Filho, Naercio (2014) : The relationship between school performance and future wages in Brazil, EconomiA, ISSN 1517-7580, Elsevier, Amsterdam, Vol. 15, Iss. 3, pp. 261-274, <https://doi.org/10.1016/j.econ.2014.08.001>

This Version is available at:

<https://hdl.handle.net/10419/179576>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/3.0/>

The relationship between school performance and future wages in Brazil

Andréa Zaitune Curi^{a,*}, Naercio Menezes-Filho^b

^a *EESP-FGV, Brazil*

^b *Inspere and University of Sao Paulo, Brazil*

Received 17 September 2013; received in revised form 8 June 2014; accepted 6 August 2014

Available online 27 August 2014

Abstract

This paper examines whether the quality of learning, measured by the scores obtained by the 1978 generation at the end of high school (1995), affects the wages received by this cohort five years later, measured in the 2000 Census. We compute average wages and proficiency for each state/race/gender cell and correct the selection bias caused by migrations and by the high educational level of the sample using Dahl (2002)'s model. School inputs are used as instrumental variables for proficiency. The results show that the average test scores impact wages, with an estimated elasticity of 0.3.

© 2014 Production and hosting by Elsevier B.V. on behalf of National Association of Postgraduate Centers in Economics, ANPEC.

JEL classification: I20

Keywords: School performance; Wages; Selection bias; Education production function; Roy model

Resumo

Neste trabalho nós investigamos a relação entre o desempenho na escola e os salários para os jovens brasileiros. Nós calculamos a média salarial e o desempenho para cada grupo definido por estado/gênero/cor para examinar se a qualidade do aprendizado, medida pelos resultados dos exames de proficiência da geração nascida em 1978 no final do ensino médio (1995), afeta os salários recebidos por esta geração cinco anos depois, medido no censo de 2000. Corrigimos o problema de viés de seleção advindo das migrações e do alto nível educacional da amostra usando o modelo de seleção de Dahl (2002). Além disto, nós utilizamos os insumos escolares como instrumentos para a proficiência. Os resultados mostraram que o desempenho nos exames de matemática de uma geração afeta o seu salário 5 anos depois, com uma elasticidade de 0,3.

© 2014 Production and hosting by Elsevier B.V. on behalf of National Association of Postgraduate Centers in Economics, ANPEC.

Palavras-chave: Desempenho escolar; Salários; Viés de seleção; Função de produção educacional; Modelo de Roy

* Corresponding author. Tel.: +55 01131680634.

E-mail address: andrea.curi1@gmail.com (A.Z. Curi).

Peer review under responsibility of National Association of Postgraduate Centers in Economics, ANPEC.

1. Introduction

The educational process in Brazil can be described as backward (even when compared to less developed countries) and highly skewed in favor of a privileged slice of the population. The debate over the importance of education as a factor explaining Brazil's income inequality is intense, with the leading current of opinion being that schooling is the main causal factor explaining this inequality, by generating productivity differences among individuals that last throughout their lifetimes (Menezes-Filho, 2001a). Besides this, human capital is one of the main determinants of the rate economic growth and level of welfare in a country.

Many international studies show that educational quality positively influences individuals' future wages (Murnane et al., 1995; Murphy and Peltzman, 2004), their probability of continuing on to higher education (Rivkin, 1995) and countries' economic growth (Bishop, 1989; Hanushek and Kimko, 2000).

Since the 1980s there have been many institutional changes in Brazil, including various educational policies aiming to reduce inequality, expand access to schooling and improve literacy rates. Quantitative result of such policies have been observed, such as a recent rise in the years of schooling. Although there have been many studies in Brazil of the quality of teaching, principally using proficiency tests, its socioeconomic impact is still little known.

In an effort to improve school quality, governments around the world have been significantly increasing spending on education. The interest of educators and politicians to improve schools has been backed by recognition of the importance of human capital in a range of aspects, such as economic performance, political stability and public satisfaction. Many of the motivations come from theoretical and empirical analyses of the relationship between an individual's schooling, the most common proxy for level of human capital. For many commentators, however, policy initiatives should not focus on the amount of schooling, but rather on its quality (Hanushek, 2003).

The central arguments about the effects of schooling have a simple structure. First, there is a direct relation between the number of years of study and wages. Given these returns to education, the policy discussion suggests the need to invest more in human capital, which translates directly into more public spending on education. The argument to boost spending generally is not confined only to increasing the years of study, but also to improving its quality. The assumption is that more money will improve educational quality, i.e., there is a close relationship between money and quality. This relationship, however, is questioned in Hanushek (2003), as we shall see shortly.

There is a good deal of international evidence that the quality of education, as measured by student achievement tests, is related to individuals' wages and productivity and a country's economic growth, even after controlling for years of schooling (Bishop, 1989; O'Neill, 1990; Murnane et al., 1995; Boissiere et al., 1985).

Recent studies on the returns to education in Brazil emphasize the effect of average schooling, measured by years of study, on wages. However, many international studies have recognized the superiority of the quality of these years of study, measured by proficiency tests, over sheer quantity to explain the wages of students when they enter the labor market, as well as the productivity of the countries and regions and their differences. What has been observed in Brazil is that years of schooling have advanced, but not in qualitative terms. The results of the achievement test given as part of the National System for Evaluating Basic Education (SAEB)¹ documents a fall in school performance between 1995 and 1999 (Ministério da Educação, 2001).

According to a study conducted by it in 2003, 55% of students who complete fourth grade have reading performance considered critical or very critical, and in the relatively backward Northeast and North regions this percentage reaches 70% and 66%, respectively, while for the more developed Southeast and South regions these respective figures are 44% and 47%. This weak performance is also observed in nearly 40% of students who complete the third year of high school. Once again the regional disparities stand out: in the South only 29% have a reading level deemed critical or very critical, while in the North and Northeast, these percentages are much larger, at 51% and 48%, respectively (SAEB, 2004).

In recent years policies have been implemented in Brazil to increase children's access to school and reduce the number of students held back and dropping out. However, doubt has been cast on the positive results of these quantitative measures by studies evaluating teaching quality. In this paper we examine the relation between school performance of young Brazilians and their wages, looking at whether the quality of teaching, measured by the scores obtained by a generation on proficiency exams in 1995, affects their wages obtained five later. Based on a pseudo-panel model,

¹ Basic education is defined in Brazil as through eighth grade.

we correct the problem of selection bias generated by migrations and the high level of schooling of the sample used, through the model developed by Dahl (2002). Determinants of school proficiency, such as school infrastructure and the profile of teachers and principals averaged at the cell (race/state/gender) level are used as instrumental variables to identify average cell-level school proficiency.

In Section 2 we present the econometric methodology developed in this work. In Section 3, we describe the data. In Section 4 we present the results of the estimations, and in Section 5 we conclude.

2. Econometric methodology

The econometric analysis we carry out in this work has two peculiarities: the use of a pseudo-panel and selection bias. The literature on pseudo-panels was pioneered by Browning et al. (1985), and the technique is used by researchers who do not have panel data, but instead various mutually independent cross sections in which different individuals are interviewed in each period. The objective of this technique is to overcome the limits of the cross sections, utilizing the advantages of these cross sections in relation to panel data. The limitations result from the fact that the researcher does not have lagged values for the variables, which in principle would make it impossible to control for the specific effects and identify dynamic models. The advantages involve the greater temporal coverage of these types of statistics, the fact they are not susceptible to attrition problems and the reduction in measurement errors due to working with more aggregated data (Menezes-Filho, 2001b).

In this work, the pseudo-panel consists of a grouping of individuals according to their birth cohort, state of residence, race and gender. We analyze people born in 1977 and 1978 at three moments in their lives: in 1982 (at 4 and 5 years old), in 1995 (in the last year of high school, at 17 and 18), and in 2000 (at 22 and 23, when they are in the labor market). This grouping is done by the mean of the dependent variable and of the regressors for each moment in time. Hence, instead of following the same individuals, we follow the generations over time, through these cohorts.

The selection bias problem appears twice. The first is because the people who make up the SAEB sample, that is, who took the achievement test in their senior year of high school, have completed ten years of schooling and are still in school. In other words, the sample is made up exclusively of students with a high educational level. The second bias is that of migration. The 2000 sample excludes those who migrated after reaching 17 or 18, in order to ensure that they took the achievement test in the same state where they live and work 5 years later. Thus, the sample used in this work is not random, but is instead a selected sample. We correct for the two selection biases by applying the model proposed by Roy (1951) and further developed in Dahl (2002).

In simplifying the estimate of the migration probabilities, Dahl (2002) assumes that the mean wages and mean preferences are the same for similar types of persons. He supposes a vector of variables containing all the relevant attributes about a type of person, such that individuals with the same values for these variables are identically affected by state-by-state differences in the sub-utility function. Comparative advantages motivate this approximation and the prediction is that individuals with different skills and characteristics will follow different migration paths on average.

If the vector describing a type of individual is composed only of discrete variables (such as age, schooling, race, sex, marital status, presence of children in the household, etc.), then similar individuals can then be grouped by cells. With this formulation, the sub-utility functions only depend on the individual's cell. An individual's migration probability is simply the fraction of individuals in a single cell who migrate between 2 states. The advantage of grouping individuals in cells is that the form of the latent sub-utility function is not specified appropriately if individuals defined as similar have the same average preferences and average earnings. Additionally, because of the cells, the estimation of the probability of migrating does not require assumptions about the error terms of the utility indices.

We first estimate the pseudo-panel model by weighted least squares at the cell level, with the dependent variable being the logarithm of the real hourly wage (lw), and the explanatory variables are: proficiency in the 11th grade ($lprof$), various observed cell characteristics, such as parental education and college graduation (X_c) the bias-correction variables (MB and EB), state, gender and race dummies and a random error (u):

$$lw_{c,2000} = \alpha + \beta_1 lprof_{c,95} + \beta_2 X_{c,t} + \beta_3 MB_{c,2000} + \beta_4 EB_{c,95} + \delta_s + \gamma_g + \lambda_r + u_{c,2000} \quad (1)$$

where c represents the cells ($c = 1, \dots, 108$).

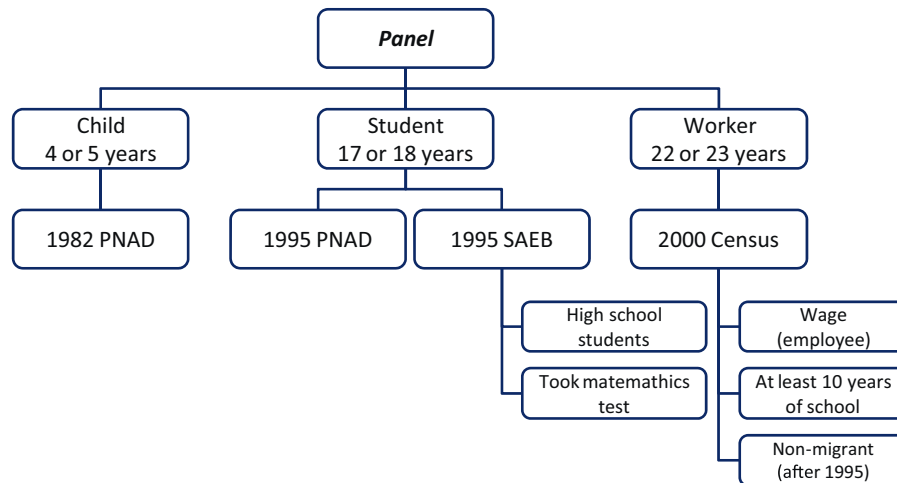


Fig. 1. Formation of the pseudo-panel.

We then instrument proficiency with the school input variables observed in 1995. The identification assumptions are that the teachers, principals and schools are correlated with average proficiency and uncorrelated with the average cohort-specific error term of the wage equation:

$$E[u_{c,2000}|Inputs_{c,95}] = 0 \quad (2a)$$

$$E[lprof_{c,95}|Inputs_{c,95}] \neq 0 \quad (2b)$$

3. Data

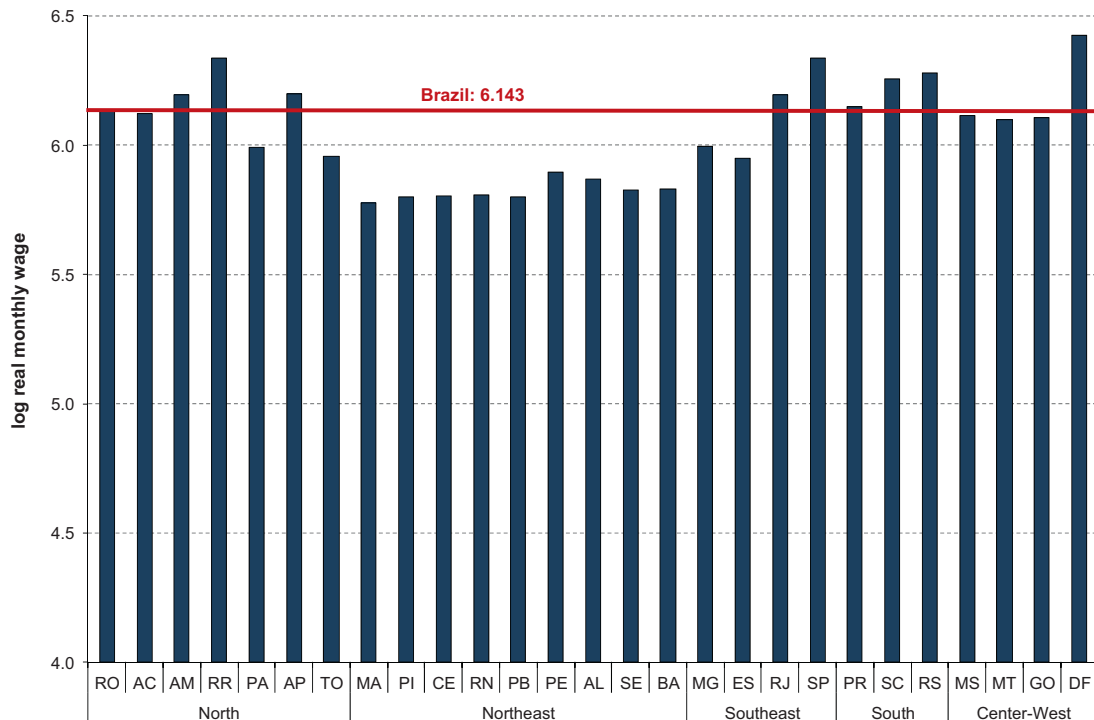
The data we use here comes from three sources. The first is the National Household Survey (PNAD) conducted by the Brazilian Institute of Geography and Statistics (IBGE) in 1982 and 1995. The second is the proficiency test given in 1995 by the Anísio Teixeira National Institute of Studies and Research (INEP), as part of the National System to Evaluate Basic Education (SAEB). The third is the 2000 Census conducted by the IBGE.²

To analyze the relationship between school performance and earnings in Brazil we form pseudo-panels (databases) that accompany the generation born in 1977 and 1978 at various moments of their lives. Just as in Dahl (2002), we assume that the average wages, the proficiencies and the average utilities are the same for similar types of persons. We use a vector of variables containing the three attributes considered relevant regarding a type of person: sex, color and state of residence. In this way, similar individuals are grouped by cells, forming a pseudo-panel model in which the generation is accompanied over time.

To form the database through the 2000 Census, we obtained information on the individuals of the 1977 and 1978 cohorts who have income, studied at least 10 years and did not migrate after 1995 onward. To this database we group the average score obtained by this generation on the SAEB test in 1995. We also group some characteristic variables of this generation that are in the 1982 and 1995 PNADs.

All the variables are grouped by cells, starting from their measures, by three characteristic vectors: sex, color and residence state, generating 108 observations (cells) in each of the databases. In this way we define the variables: *SEX* (man or woman), *RACE* (white or others) and *STATE* to form these cells representing the 108 types of individuals whose cohort we follow at three stages of their lifetimes. We chose these three characteristic vectors due to the results widely disclosed for Brazil and other countries on the wage differences between men and women, between races (O'Neill, 1990), and between states, since the return to education is not equal among the states (Dahl, 2002). Fig. 1 details the sample's composition, indicating the data used to form the pseudo-panel and its characteristics.

² Sites of the IBGE – www.ibge.gov.br – and INEP – www.inep.gov.br.



Graph 1. Real monthly wage in 2000 by state.

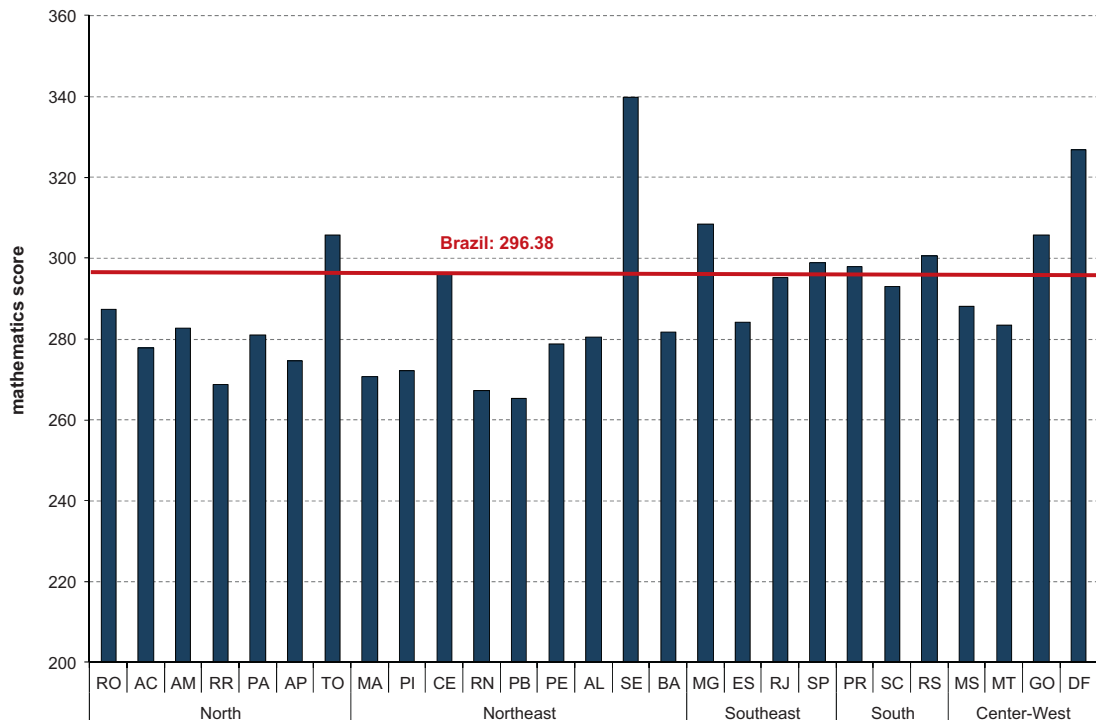
The sample used from the 2000 Census represents the individuals in the working phase of their lives. This covers young adults 22 and 23 years old, with 10 or more years of schooling and who did not migrate after 1995, the year when the youths of this generation took the SAEB proficiency test. In other words, the sample consists of those who responded that they were born in the state where they were living and working in 2000, and those who were not born in that state, but lived there without interruption since 1995. Regarding wages, we use the logarithm of the real hourly wage (*LRHW*) in the regressions, and the log of the real monthly wages in the graphs to facilitate visualization. The deflator used is the National Consumer Price Index (INPC) for the Census. We also weighted the wages by the cost of living in the regions.³

Graph 1 shows the logarithm of the average wage received in 2000 in each state by the young Brazilians 22 and 23 years of age, with 10 or more years of schooling and who did not migrate after 1995 (selected sample).⁴ According to the data, the monthly wages in Brazil in 2000 were, on average, equal to 6.14 in log form. Even deflating the wages by regional indices, the wage inequalities among the states and regions are glaring. Wages in the North region were, on average, the lowest in Brazil, around 5.8 in logarithm, while in the South region the average was 6.2. The states with the highest average wages in 2000 were São Paulo, the Federal District and Roraima. We also point out that among the states of the Northeast, the average wages are very similar, while in the Southeast there is a large disparity among the states.

From the 2000 Census data we created two dummy variables that identify the educational characteristics of the youths of the 1977 and 1978 cohorts. One of them is a variable on years of study, which identifies the individuals with 12 or more years of schooling (*EDUCA_15*), equal to 1 for those individuals and zero otherwise, and the other is a variable indicating whether these young people were attending school in 2000 (*ATTEND.SCHOOL*), which is equal to 1 for those in school that year and zero for those not still in school.

³ Menezes, T. and Azzoni, C. “Convergência de renda real e nominal entre as regiões metropolitanas brasileiras: uma análise de dados de painel”, XXVIII Encontro da Anpec, Campinas, 2000.

⁴ Table A1 in the appendix provides a list of State names.



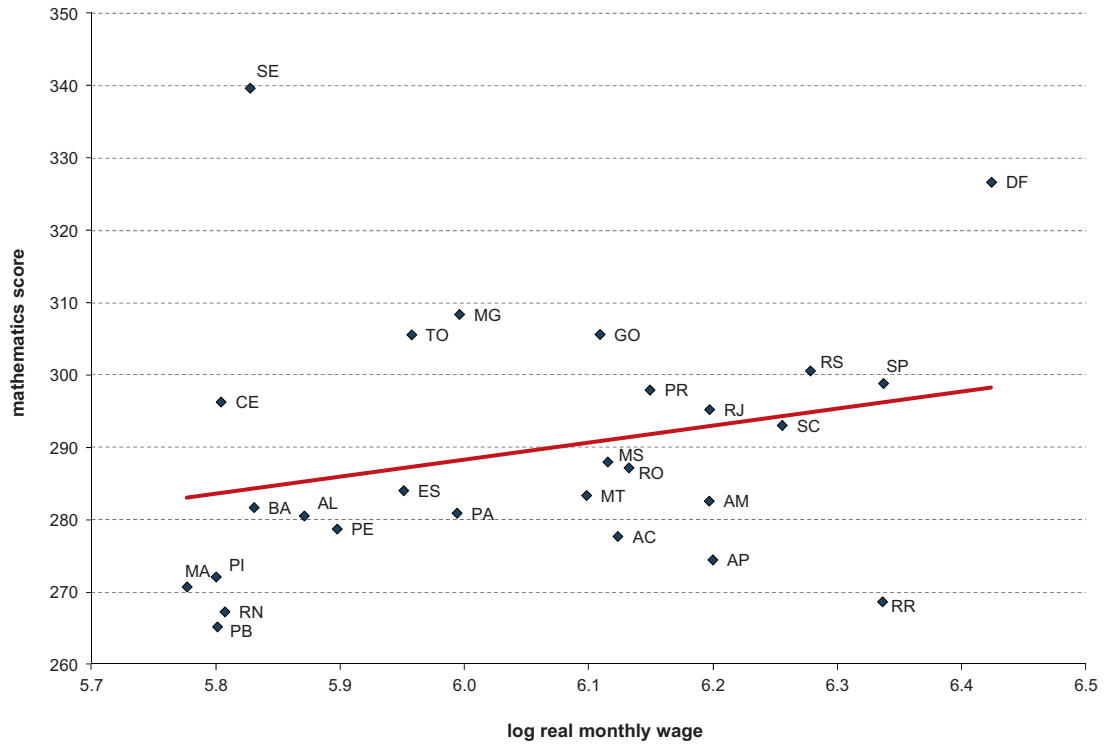
Graph 2. Proficiency in mathematics by state in 1995.

The two key variables of the model we use to correct the selection bias come from the 1995 PNAD (correction of the educational bias) and the 2000 Census (correction of the migration bias). Thus, as in [Dahl \(2002\)](#), we use a variable representing the percentage of 17- and 18-year-olds who in 1995 had exactly 10 years of schooling and were in school (*EDUCATIONAL BIAS*), that is, the percentage of individuals eligible to take the SAEB proficiency exam that year, and a variable representing the percentage of people from the 1977 and 1978 cohorts who did not move from the state where they lived in 1995 (*MIGRATION BIAS*).

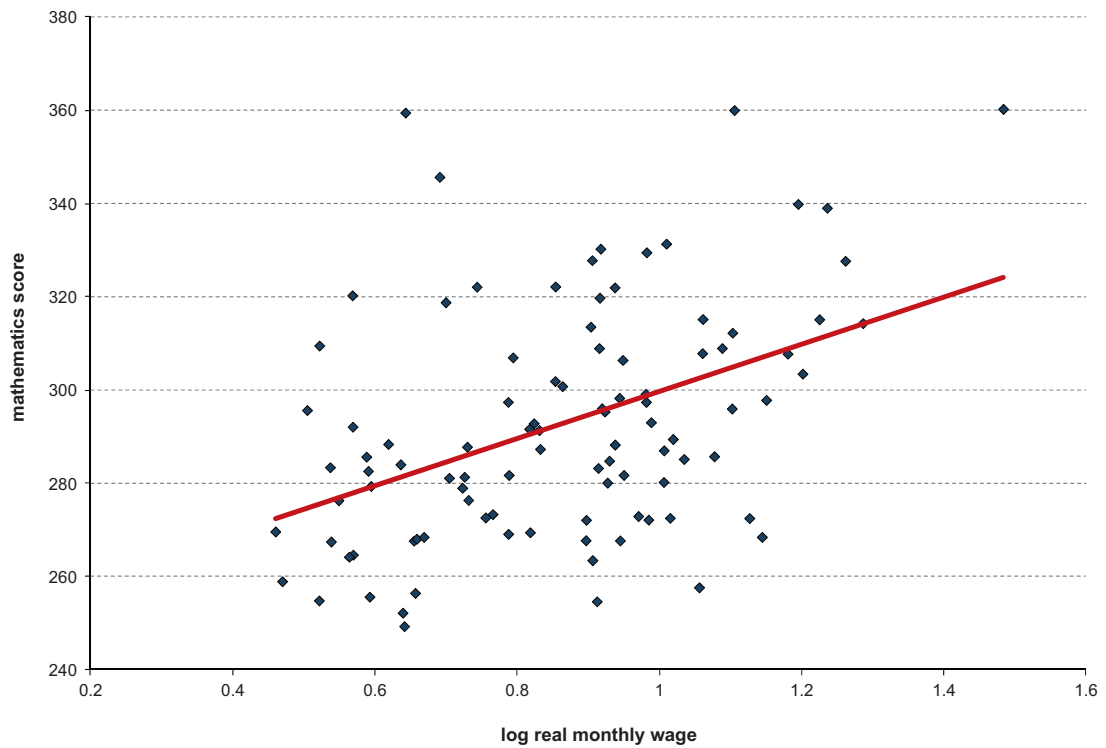
[Graph 2](#) shows the average performance in 1995 in each Brazilian state of the 17- and 18-year-old students on the mathematics proficiency test given by INEP. The highest scores occur in Sergipe, Tocantins, the Federal District and Rio Grande do Sul, all of them on average above 300 points. The lowest scores, under 270 points, occur in Roraima, Rio Grande do Norte and Pernambuco. Just as occurs with the wages received by the youths, the school performance presents regional inequalities. In all the states of the South region the average results were above 290 points, while in the states of the North and Northeast the average scores were under 280 points, with few exceptions.

[Graph 2](#) shows the state of Sergipe as having one of the highest average scores on the mathematics test, along with the state of Tocantins. Although this may at first glance appear weird, it can be explained by the fact that in these states only a small portion of the population (an unrepresentative “elite”) reach 10 years of schooling. On the other hand, in larger states, in which policies of not holding students back were applied, such as in São Paulo, a much larger part of the population reaches the senior year of high school, and they do worse on average on the math proficiency test. It should be noted that the econometric exercises below control for state dummies, which should capture different state education policies that are fixed over time.

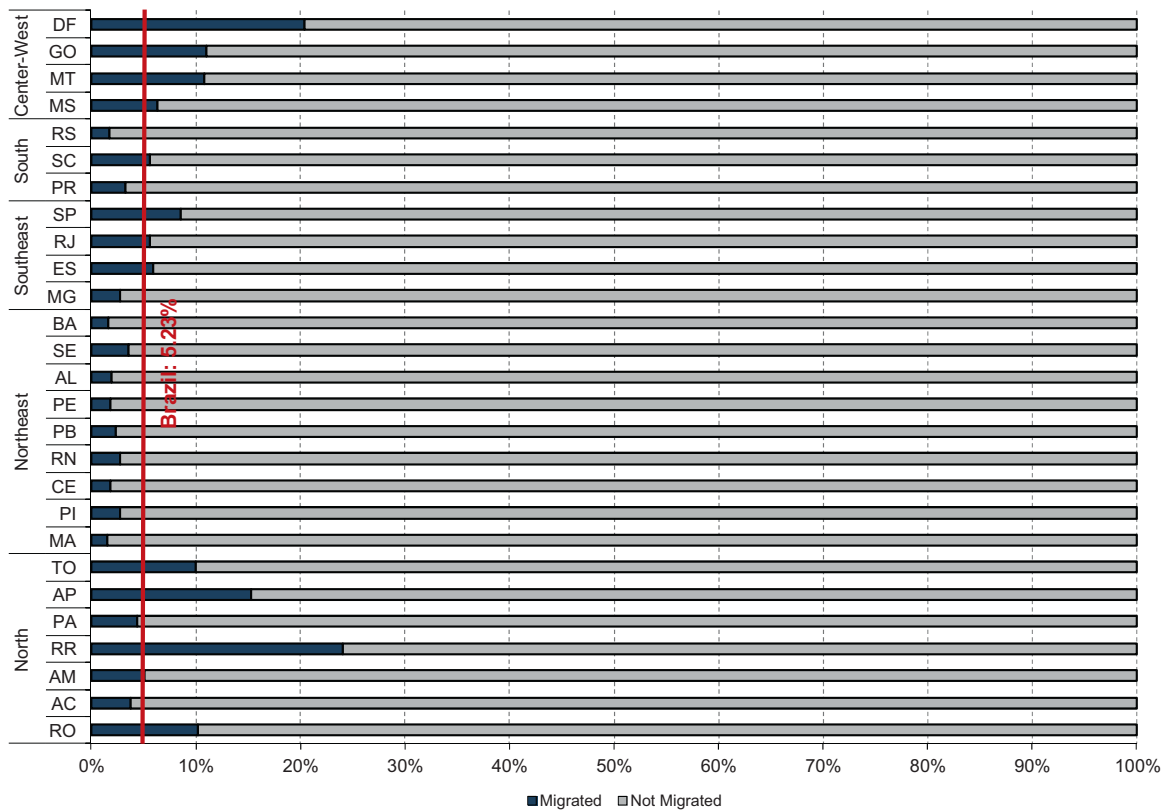
In [Graph 3](#) we cross reference the wages with proficiency by state, showing the relation between the math proficiency test score received by the 17- and 18-year-olds and the wages received by them five years later in the job market. The graph shows a positive relation for the 1977 and 1978 generation between the average scores received on the math achievement test in 1995 and the average wages received five years later. The graph indicates that on average pay levels are higher in the regions where school performance is better at the end of high school. In other words, higher scores on the proficiency tests are related to higher future wages.



Graph 3. Proficiency in mathematics versus wage by state.



Graph 4. Proficiency versus wage per cell. *Note:* the graph shows the 99 cells with 10 or more persons.



Graph 5. Percentage of migrants after 1995 per state: migration bias.

Graph 4 repeats the exercise of the preceding graph, but by cells. Each cell represents a type of individual according to sex, color and residence state, made up of individuals born in 1977 and 1978 who did not migrate after 1995, i.e., who took the achievement test in the same state where they worked in 2000. We considered only the cells composed of at least 10 people to ensure these cells are representative. We can observe that even controlling for other characteristics, the relation between wages and school performance is positive.

3.1. Migration in Brazil

Using the data from the 2000 Census, we now analyze the percentage of youths who moved after 1995 (when the test was given) to the states where they lived in 2000. This analysis is exactly the variable we use in the model to correct the migration bias of the selected sample, which represents the percentage of non-migrants after 1995.

According to Graph 5, only the states of Roraima and the Federal District had more than 20% of their population composed of 22- and 23-year-olds who had migrated there after 1995 onwards. In contrast, in Rio Grande do Sul, Bahia and Maranhão this figure is less than 2% of the population for these cohorts. The national average of youths of this age range who migrated to their state of residence in 2000 starting after 1995 is 5.23%. Comparing the regions, each of the Northeast states has at least 3.5% of young migrants who moved there after 2005. In the Center-West region, in turn, over 6% of residents of its states in 2000 were migrants who arrived after 2005.

Table 1 shows the origins and destinations of the young Brazilians born in 1977 and 1978 who migrated between states after 1995. According to this table, the states in the Southeast region were the destinations of over 48% of the people who migrated from the state they lived in after 1995 forward, and more than 55% of the migrants who were born in one state in this region moved to another in the same region. Over 55% of people born in the Northeast who left the state where they lived in 1995 went to the Southeast region. The states that attracted the fewest migrants were in the North and South regions – with only 10% of migrating Brazilians after 1995 onwards.

Table 1
Destination of young migrants after 1995 (2000 Census).

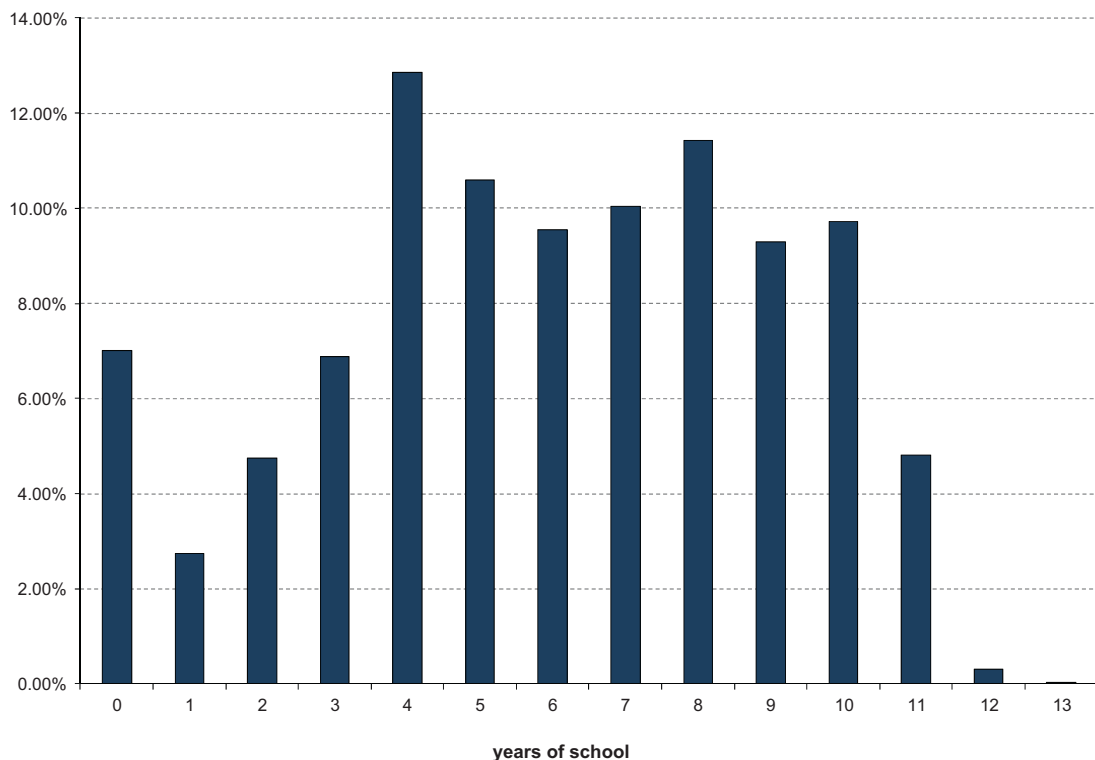
%	Destination region					Total
	North	Northeast	Center-West	Southeast	South	
Birth region						
North	48.49	9.48	26.16	13.54	2.33	100
Northeast	8.48	13.1	14.86	61.95	1.61	100
Center-West	16.09	7.7	39.73	29.67	6.8	100
Southeast	4.73	12.86	16.13	55.32	10.96	100
South	5.36	2.69	16.96	35.72	39.27	100
Total	10.42	11.2	17.7	52.27	8.42	100

3.2. Analysis of the education of the young Brazilians in 1995

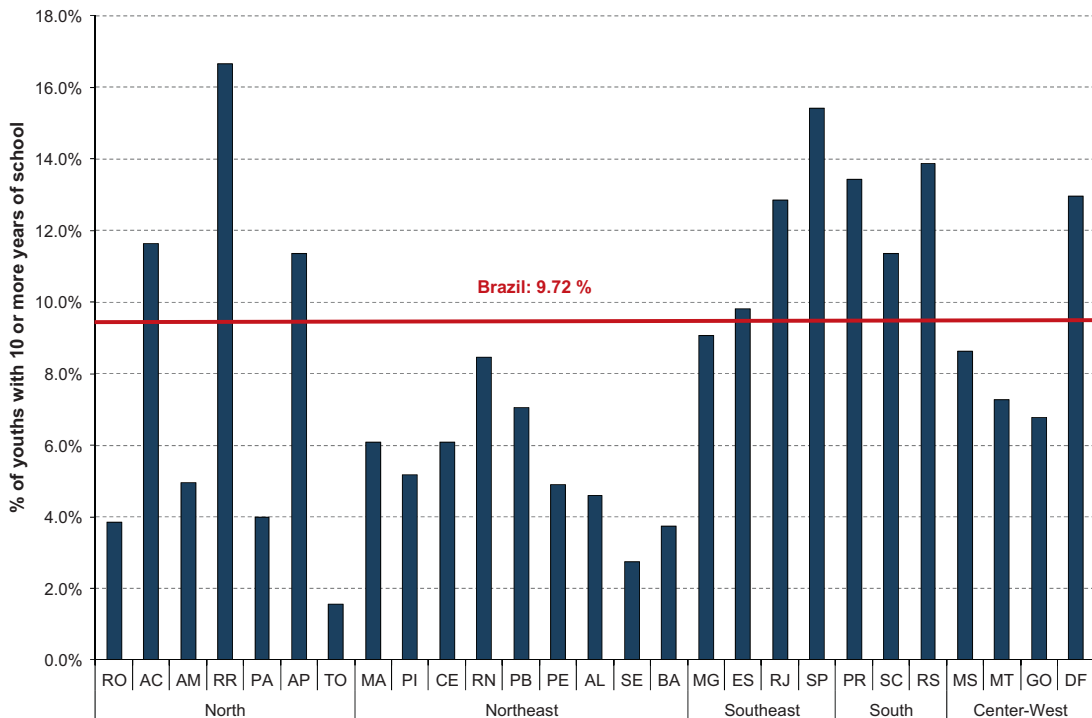
In this section we analyze the educational level of 17- and 18-year-old Brazilians in 1995. An individual who started attending school at 6 or 7 years of age and never repeated a grade must have had 10 years of schooling when reaching 17 or 18.

Nevertheless, [Graph 6](#) shows that in 1995, of those born in 1977 and 1978, 7% never even finished one year of school, and just over 12.8% of them only completed up through fourth grade. All told, just under 35% of the Brazilian population aged 17 and 18 had less than 5 years of schooling, when they should have been in their last year of high school (in the 11th year of study). Also according to this graph, fewer than 11.5% of these youths had completed junior high school (8 years of study), and fewer than 10% had completed 10 years of schooling.

[Graph 7](#) shows the percentage for each state of the 17- and 18-year-olds who had completed 10 years of schooling in 1995. In the Northeast states, except for Rio Grande do Norte, fewer than 7% of the youths from the 1977 and 1978 cohorts had 10 years of study. In contrast, in the South region, the percentage in all states is greater than the national



Graph 6. Educational level of young Brazilians in 1995.



Graph 7. Percentage of youths per state with 10 years of schooling.

average of 9.72%. The same is true of the Southeast region, except for Minas Gerais, where only 9% of the youths of these cohorts had 10 years of study.

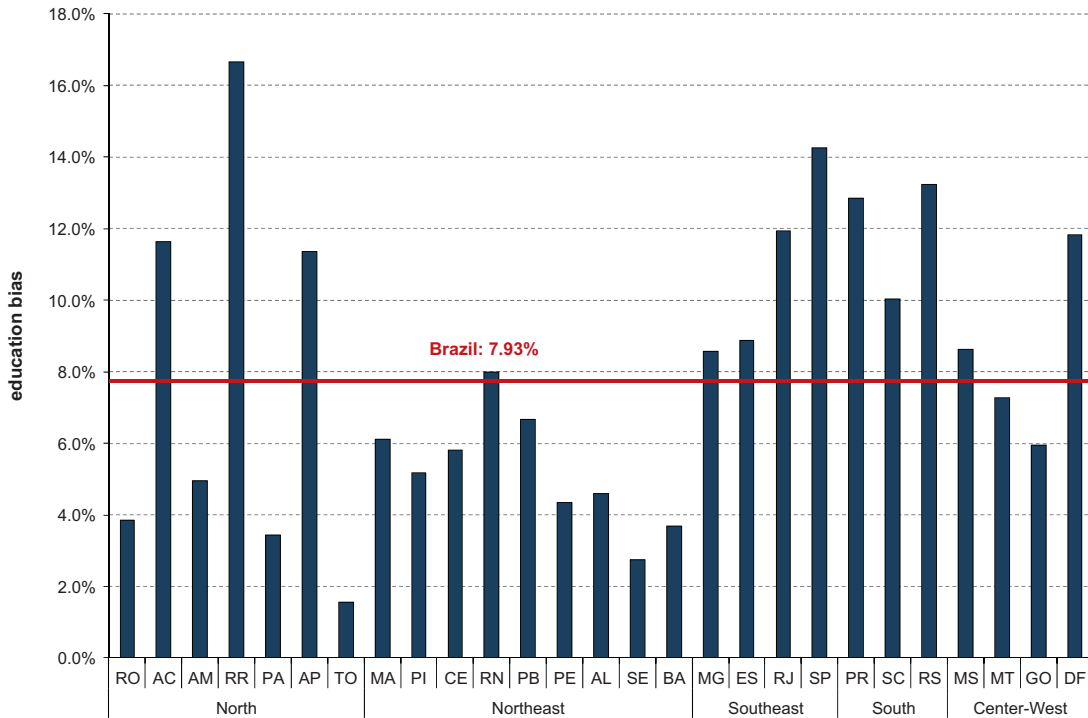
The sample used in this work is restricted to the 17- and 18-year-olds who had completed exactly 10 years of study and were attending school in 1995, because the only individuals eligible to take the SAEB proficiency test were those born in 1977 or 1978 who were high school seniors in 1995. Since this sample is not random, the model has a sample selection bias, which we correct by using a variable that indicates the percentage of the youths with these characteristics in each cell.

Graph 8 shows the percentage of youths born in 1977 and 1978 with exactly 10 years of study and who were in school in 1995, for each state. This is the variable used to correct the model's educational selection bias. In Brazil nearly 8% of the population has these characteristics. In the states of Roraima and São Paulo, less than 14% of the population was eligible to take the 1995 SAEB test. Tocantins was the only state in 1995 where fewer than 2% of the people born these years had exactly 10 years of schooling and were still attending school, and in Sergipe this figure was under 3%. This corroborates the analysis we discussed previously that these states obtained high averages on the achievement tests because only a small "elite" part of the population reaches that educational level, while in the state of São Paulo, where the policy is not to flunk public school students, a much larger part of the population reaches the last year of high school.

Tables A2 and A3 in the Appendix describe the data that will be used in the estimation procedure, by showing cells means and standard deviations. We can notice that 47% of the cells are male cells, 46% are white cells, 45% of the people in the cells still attend school and have about 11.3 years of schooling, 19% of the people in the cells have entered college, 94% remained in the same state in the last 5 years, and about 8% of the cells had 10 years of education when they were 17 or 18 years old. Fig. 2 plots the source of each of the variables used in the analysis.

4. Results

Table 2 presents the results of the pseudo-panel model of equation (1), in which we estimate the effect of the performance on the math achievement test on wages five years later. Column (I) shows that average proficiency scores



Graph 8. Educational bias.

and average wages 5 years later are positively and strongly correlated. Column (2) includes the dummies for state, race and gender, which identify each of the cells, which causes a drastic fall in the estimated proficiency coefficient that nevertheless remain significant. Column (3) includes the controls for the bias, which are not statistically significant and hardly alters the proficiency effect. This means that the state, gender and race dummies are already capturing much of the selection bias. In column (4) we include the other controls, which are not significant either, and the estimated proficiency coefficient remains close to 0.2.

Finally, in column (5) we use the instrumental variable procedure, using the school inputs as instruments for school proficiency. We perform this exercise to examine whether there are omitted variables that are correlated with the proficiency and with wages, such as the students' average (cell) ability. The instrumental variables we use are

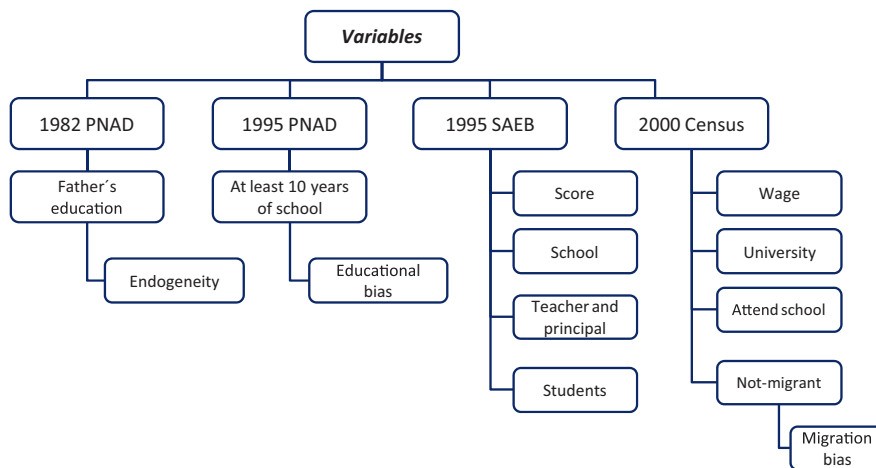


Fig. 2. Sources of variables.

Table 2
Proficiency in mathematics (1995) and wages (2000 Census).

	OLS				IV
	(I)	(II)	(III)	(IV)	(V)
Dependent variable: log real hourly wage ^a					
Mathematics score	1.737 <i>0.386</i>	0.237 <i>0.098</i>	0.243 <i>0.099</i>	0.186 <i>0.118</i>	0.319 <i>0.154</i>
Male		0.133 <i>0.011</i>	0.137 <i>0.011</i>	0.158 <i>0.021</i>	0.146 <i>0.022</i>
White		0.228 <i>0.009</i>	0.217 <i>0.014</i>	0.173 <i>0.030</i>	0.174 <i>0.030</i>
State dummies		YES	YES	YES	YES
Migration bias			−0.082 <i>0.170</i>	−0.319 <i>0.228</i>	−0.316 <i>0.236</i>
Education bias			0.167 <i>0.153</i>	0.057 <i>0.168</i>	0.084 <i>0.168</i>
Education: 12 or more years completed				0.303 <i>0.237</i>	0.248 <i>0.237</i>
Attend school				0.025 <i>0.098</i>	0.050 <i>0.097</i>
Father's education				0.007 <i>0.011</i>	0.004 <i>0.011</i>
Constant	−8.873 <i>2.168</i>	−0.439 <i>0.556</i>	−0.423 <i>0.559</i>	0.165 <i>0.693</i>	−0.554 <i>0.871</i>
R ²	0.309	0.986	0.986	0.987	0.986
First stage – excluded instruments $F(16,50)$					3.19
Obs	100	100	100	100	100

^a Note: Standard error in italics.

the teachers' and principals' characteristics, such as salaries and education level, and school characteristics, such as number of computers, presence of science labs, library and share of private schools in the cell. The estimated coefficient increases to about 0.31, which indicates that a 10% increase in proficiency increases wages by about 3%. Finally, the F -test of joint significance of the instrument set does not reject the null that all instrumental variables (listed in Table 2) are jointly equal to zero.

5. Conclusion

In this paper we examined the relation between school performance and wages of young Brazilians. We computed average wages and proficiency by state/race/gender cells to examine whether the quality of learning, measured by the scores obtained by the cohort born in 1978 at the end of high school (1995), affects the wages received by this cohort five years later, as measured in the 2000 Census. We corrected the selection bias problems caused by migrations and by the high educational level of the sample using Dahl (2002)'s selection model. Finally we used school inputs as instruments for proficiency. The results showed that the average test scores of a generation impacts its wages 5 years later, with an elasticity of 0.3.

Appendix – Descriptive statistics.

Table A1
List of states.

DF	Distrito Federal	RJ	Rio de janeiro
GO	Goiás	ES	Espirito Santo
MS	Mato Grosso Sul	MG	Minas Gerais
MT	Mato Grosso	BA	Bahia
RS	Rio Grande do Sul	SE	Sergipe

Table A1 (Continued)

DF	Distrito Federal	RJ	Rio de janeiro
SC	Santa Catarina	AL	Alagoas
PR	Paraná	PE	Pernambuco
SP	São Paulo	AL	Alagoas
PE	Pernambuco	CE	Ceará
PB	Paraíba	PI	Piauí
RN	Rio Grande do Norte	MA	Maranhão
TO	Tocantins	AP	Amapá
PA	Pará	RR	Roraima
AM	Amazonas	AC	Acre
RO	Rondonia		

Table A2

Descriptive statistics (PNAD and Census).

Variables	Obs.	Mean	Std. Dev.
Male	100	0.474	0.502
White	100	0.463	0.501
Log real hourly wage	100	0.876	0.214
Attend school	100	0.450	0.174
Years of schooling	100	11.357	0.294
Education: 12 or more years	100	0.189	0.087
Log mathematics score	100	5.659	0.085
Mathematics score	100	291.795	25.372
Migration bias	100	0.940	0.052
Education bias	100	0.078	0.051
Father's education	100	3.226	1.285

Table A3

Descriptive statistics (instrumental variables).

Mathematics	Obs.	Mean	Std. Dev.
Teacher's salary – <R\$308	100	0.315	0.423
Teacher's salary – R\$308–1023	100	0.532	0.494
Teacher's salary – 1024–2046	100	0.116	0.354
Teacher's salary – >2046	100	0.036	0.199
Teacher's experience – less than 1	100	0.006	0.101
Teacher's experience – 1–0 years	100	0.477	0.493
Teacher's experience – more than 10	100	0.516	0.494
Teacher's education – no college	100	0.090	0.24
Teacher's education – college	100	0.637	0.471
Teacher's education – graduate	100	0.273	0.445
Principal's salary – <r\$308	100	0.070	0.2
Principal's salary – r\$308–1023	100	0.452	0.482
Principal's salary – 1024–2046	100	0.322	0.489
Principal's salary – >2046	100	0.155	0.398
Principal's experience – less than 1	100	0.007	0.083
Principal's experience – 1-0 years	100	0.190	0.331
Principal's experience – more than 10	100	0.802	0.338
Principal's education – no college	100	0.073	0.217
Principal's education – college	100	0.628	0.485
Principal's education – graduate	100	0.298	0.47
Private school	100	0.248	0.451
Library	100	0.838	0.315
Science lab	100	0.485	0.487
Computer	100	0.190	0.449

References

- Bishop, J., 1989. *Is the test score decline responsible for the productivity growth decline?* *Am. Econ. Rev.* 79 (1), 178–197.
- Boissiere, M.X., Knight, J.B., Sabot, R.H., 1985. *Earning, schooling, ability, and cognitive skill.* *Am. Econ. Rev.* 75 (5), 1016–1030.
- Browning, M., Deaton, A., Irish, M., 1985. *A profitable approach to labor supply and commodity demands over the life cycle.* *Econometrica* 59, 503–544.
- Dahl, G.B., 2002. *Mobility and the return to education: testing a royl model with multiple markets.* *Econometrica* 70 (November (6)), 2367–2420.
- Hanushek, E.A., 2003. *The failure of inputs-based schooling policies.* *Econ. J.* 113 (February), F64–F98.
- Hanushek, E.A., Kimko, D.D., 2000. *Schooling, labor force quality, and the growth of nations.* *Am. Econ. Rev.* 90 (December (5)), 11894–12208.
- Menezes-Filho, N.A., 2001a. *Educação e desigualdade.* In: Lisboa, M.B., Menezes-Filho, N. (Eds.), *Microeconomia e sociedade no Brasil. Contra Capa Livraria, Rio de Janeiro.*
- Menezes-Filho, N.A., 2001b. *Microeconometria.* In: Lisboa, M.B., Menezes-Filho, N. (Eds.), *Microeconomia e sociedade no Brasil. Contra Capa Livraria, Rio de Janeiro.*
- Ministério da Educação, 2001. *Desempenho do Sistema Educacional Brasileiro: 1994–1999*, Brasília.
- Murnane, R.J., Willett, J.B., Levy, F., 1995. *The growing importance of cognitive skill in wage determination.* *Rev. Econ. Stat.* 77 (May (2)), 251–266.
- Murphy, K.M., Peltzman, S., 2004. *School performance and the youth labor market.* *J. Labor Econ.* 22 (2), 299–325, The University of Chicago.
- O’neill, J., 1990. *The role of human capital in earnings differences between black and white men.* *J. Econ. Perspect.* 4 (October (5)), 869–895.
- Rivkin, S.G., 1995. *Black/white differences in schooling and employment.* *J. Hum. Resour.* 30 (Fall (4)), 826–852.
- Roy, A.D., 1951. *Some thought on the distribution of earnings.* *Oxf. Econ. Paper* 3, 135–146.
- Sistema Nacional de Avaliação da Educação Básica, 2004, June. *Resultados do SAEB 2003 Brasil e Rio de Janeiro*, Brasília, DF.