Chapter 3

Brazilian educational system. Institutional outline and a mapping of inequalities.

Abstract

This chapter opens the empirical investigation of this dissertation. It contains a description of Brazil's large educational system, including a brief institutional outline and an exposition of key indicators regarding recent years, which document a substantial improvement in quantitative indicators. In order to explore qualitative issues, we turn to a presentation of the SAEB database (which is employed from this chapter on), followed by a discussion of a few methodological issues concerning the application of inequality and welfare analysis tools to distributions of test scores. Finally, we map test scores inequalities in Brazil, assessing their intensity, both at the national level, and at sub-national levels.

Contents

1 Introduction

The objective of this chapter is to point out the most important facts regarding the Brazilian educational system, serving as an introduction to the empirical investigation of this dissertation.

It contains a description of Brazil's very vast educational system (section 2), including a brief institutional outline and an exposition of key indicators regarding recent years, which document a substantial improvement on quantitative grounds. In order to explore qualitative issues, we turn to test scores data in section 3. We first present the SAEB dataset[1], which is employed in this dissertation from this chapter on, followed by a discussion of a few methodological issues concerning the application of inequality and welfare analysis tools to distributions of test scores.

Then, we map test scores inequalities in Brazil, assessing its intensity, both at the national level (section 4), and at sub-national levels (section 5). We find that distributions of less advanced grades are more unequal than those of more advanced grades, particularly in recent years. When we decompose inequality space-wise, we find that inequality within regions and also within states account for large fractions of overall inequality (around 95% percent of overall inequality in both cases), while inequality within municipalities is smaller (around 4/5 of overall inequality). In our rankings of scores inequalities, the country's most populated regions - the rich South-East and the poor North-East - are identifed as the most unequal ones. States located in such regions are also the most unequal ones. São Paulo, the richest and most populated unit of the federation, is the most unequal one.

In section 6 we conclude.

2 A brief description of the Brazilian educational system

2.1 Institutional outline

The Brazilian educational system is divided in the following subsystems:

- Pre-school: for children younger than 7 years old;
- Primary school: eight years of courses, with recommended ages ranging from 7 to 14 years old:

[1]We would like to thank the Instituto Nacional de Estudos e Pesquisas Educacionais (INEP) for giving us permission to use the SAEB databases.

Lower-primary: 7-10 years old, from the 1st to the 4th grade;

Upper-primary: 11-14 years old, from the 5th to the 8th grade;

- Secondary school: three years of courses, with recommended ages ranging from 15 to 17 years old, from the 1st to the 3th grade;
- Higher education: three to six years of courses, depending on the domain, for young people older than 17 years old;
- "Others": special education (for handicapped people); technical education; and education for adults.

Starting in 2006, school became mandatory for children aged 6-14, that is, during a period of nine years, including one year of pre-school and eight years of primary school. Before that, it was mandatory for children aged 7-14.

In Table 1, we observe that the total number of students in the system has increased 38.89% in 14 years, to reach around 54 million people in 2005.[2] So clearly, it is an enormous system, reflecting the country's large and young population, and also a series of policies implemented in recent decades aiming at bringing children, teenagers, young people, and adults (back) to the school.

Both the public and the private sectors participate in the task of educating Brazilians. Considering all levels of education together, in 2005 the private sector had almost 10 million students, a share of almost 1/5 of the total. The private sector's role is thus far from negligible, and cannot be considered simply as a complementary system as it happens to be the case in other countries, particularly in some developed countries. The share of students enrolled in the private sector has not changed considerably when we compare 1991 (17.70%) to 2005 (18.04%). However, the participation of the private sector has increased particularly in higher education, rising in one and a half decade, from an already large share of 61.30% in 1991, to almost 3/4 in 2005. In turn, the share has decreased in other levels, especially secondary school (from 27.02% to 12.15%).

The public sector, in turn, is sub-divided in different layers of government (federal, states', and municipal) each of which has educational duties, which were vaguely defined in the past, but have tended to become clearer in recent years. Municipalities' priority is to provide basic education (pre-school and primary; lower-primary in particular). State's priority is secondary education, and to some extent, upper-primary. The role excepted to be played by the federal

[2]Students enrolled in the cateogory "Others" are not included in this section's figures. Yet, it should be noticed that it is not a negligible group, since in 2005, they were 6.7 million people, corresponding to around 11.00% of the total number of students enrolled in the Brazilian educational system.

Table 1: Enrollment in Brazilian educational system. 1991 and 2005.

Educational Level	School Type	Government Level	Year 1991 Enrolled	Year 1991 *Share*	Year 2005 Enrolled	Year 2005 *Share*
Pre-School	Public	Federal	17,240	0.38	2,561	0.04
		State	424,480	9.44	266,265	3.70
		Municipal	2,742,849	60.97	4,887,641	67.84
	Private		1,313,868	29.21	2,048,546	28.43
	Total		4,498,437	100.00	7,205,013	100.00
Primary School	Public	Federal	95,536	0.33	25,728	0.08
		State	16,716,816	57.24	12,145,494	36.22
		Municipal	8,773,360	30.04	17,986,570	53.64
	Private		3,618,012	12.39	3,376,769	10.07
	Total		29,203,724	100.00	33,534,561	100.00
Secondary School	Public	Federal	103,092	2.73	68,651	0.76
		State	2,472,964	65.55	7,682,995	85.07
		Municipal	177,268	4.70	182,067	2.02
	Private		1,019,374	27.02	1,097,589	12.15
	Total		3,772,698	100.00	9,031,302	100.00
Higher Education	Public	Federal	320,135	20.46	579,587	13.02
		State	202,315	12.93	477,349	10.72
		Municipal	83,286	5.32	135,253	3.04
	Private		959,320	61.30	3,260,967	73.23
	Total		1,565,056	100.00	4,453,156	100.00
Total*			39,039,915		54,224,032	

Source: INEP/MEC (www.inep.gov.br).

* Not included: special education, technical education, education for adults.

government is “redistributive and complementary” in any level where it becomes necessary, and to “organize higher education”.[3]

It is hard to believe that a perfect division of labor between the three layers (plus the private sector) will be reached, at least in the short run, But a trend of “specialization” (i.e., each layer of government taking care of a specific level of education) is taking place:

- Municipalities increasingly take care of pre-school (share of pre-school pupils enrolled by municipal schools in 2005: 67.84%) and primary education (share of pupils in 2005: 53.64%).
- States are more and more responsible for secondary education (share of pupils in 2005: 85.07%), but still educate 36.22% of primary school pupils (mainly upper-primary pupils) and 10.72% of higher education students (especially in richer states, which have enough fiscal revenues to fund their own higher education system).
- The federal government’s share in terms of provider of education remains important exclusively in higher education, with more than half a million students. Relatively, it has lost ground, given that its share has gone down from 20.46% to 13.02%.

So a private schooling system co-exists with a public one, with contrasting funding and managing characteristics. Private schools are neither funded nor managed by the public authority, which implies their decision-making environment and the nature of their budget constraints are quite distinct from those of public schools. Private schools compete with other schools in the market, whereas public schools are managed by the government and respond to bureaucratic rules. For example, there are striking differences between the private and the public systems concerning the functioning of teachers’ labor market (a topic we turn to in chapter 5). In private schools, there is an almost unconstrained teachers’ labor market: recruitment procedures and wage settings are decided on a decentralized basis, subject to some constraints imposed by unions and collective bargaining rules. Each school is considerably free to make decisions related to teacher hiring and performance rewarding. In the public system, in turn, there is no free market for teachers. Recruitment should preferably be done by means of public contests, but many teachers are allocated to their jobs by politicians or through other, non-competitive, means. Wages depend on general guidelines stated by the federal authority (Ministry of Education, or MEC), which are filtered and adapted by state-level and/or municipal-level decision-makers. Public schools cannot decide

[3]Source: Brazil’s Ministry of Education website (www.mec.gov.br).

to pay higher salaries in order to attract better teachers, nor to increase rewards to good teachers.

In general, primary and secondary private schools offer higher wages for teacher, principals, and staff; have better infra-structure; and educate kids of richer families. Parents pay tuition fees, which usually are quite expensive[4], supplying schools with resources that allow some schools to operate in an environment that resembles that of good schools in developed countries (e.g., with good libraries, computer facilities and so on.). Public schools, in turn, must cope with much more constrained budget sets.

Studying in public schools is socially stigmatizing too, such that parents who can afford it, prefer to opt out of public schools and enroll their children in private schools, generally identified as "good schools".[5]

The situation in terms of quality and status of private versus public *higher education institutions* is different from the description made above. Some public universities still show good quality and have good status, in contrast with the general picture of the schooling system. In the fast-growing private sector, both relatively good and relatively bad institutions in terms of quality and status co-exist. As we do not analyse higher education in this dissertation[6], we shall not provide further details regarding that level of education.

Expenditure with public education[7] accounted for 3.9% of Brazil's GDP in 1994, and has risen to 4.1% in 2002, and to 4.4% in 2003. (For a comparison, the average for OECD countries in 2003 was 5.2%.) In 2003, public expenditure with education was decomposed as follows: 0.6 by federal government, 2.0 by states, and 1.8 by municipalities. For that same year, Soares (2002) estimates the expenditure with private education was as high as 3.0% of the GDP (average for OECD countries in 2003: 0.7). According to such figures, Brazil would be spending with education the equivalent of almost 7.5% of its GDP,

[4]Part of the money spent with tuition fees returns to parents by means of tax credits (Ioschpe, 2004). So the federal state indirectly subsidizes expenditures with private schooling, which are granted after tax, in a kind of implicit voucher system (thus, different from the direct funding of public schooling).

[5]Oliveira and Schwartzman (2002) report a survey that documents the perception Brazilians have about the quality of each type of school. When asked in what type of school they would prefer to enroll their children, 94.3% of private school teachers said they would choose private schools. And a great fraction of (municipal) public schools teachers themselves would choose private schools (69.9%), instead of public schools.

[6]Since higher education is not crucial in an 'educationist' approach, as it has been defined in Part I.

[7]The sources of figures exposed in this paragraph are Soares (2002) and INEP's website (http://inep.gov.br/estatisticas/gastoseducacao/).

more - in relative terms, of course - than the 5.9% spent on average by OECD countries.

The role of the federal authority, in many instances (e.g., regarding private higher education), is restricted to evaluative and regulatory duties. Even for the primary school level, no national curriculum has been defined - there is only a very general guideline set by the federal authority. As this vast and complex system of provision, funding and regulation suggests, the locus of policy-making is not centralized at all, the number of agents involved is enormous, and the structure of motivation and incentives each of them responds to can be quite different. Reforms cannot be thought of without a minutious diagnosis of its main deficiencies, and without taking its complexities into account.

2.2 An incomplete 'quantitative revolution'

Table 2 indicates that, in a sense, a 'quantitative revolution' has been taking place in Brazilian education in the past two decades. The improvements are pervasive, both in terms of flow variables (decreased average delay, decreased fraction of delayed children, increased gross and net enrollment rates) and in terms of stock variables (increased average years of schooling, reduced illiteracy rates).

Some dynamic implications of this 'quantitative revolution'are still to come (e.g., possibly a decrease in the returns to schooling, as mentioned in this dissertation's Introduction), but some of them are being felt right now (e.g., increased requirements for an expansion of the supply in advanced levels of the system).

From 2004 to 2005, for the first time in Brazil's recent history, there has been a decrease (-1.5%) in the number of students enrolled in secondary school.[8] To interpret this as good news - namely, as an eventual stabilization of the size of the system - we would have to be sure that such decrease is only due to the combination of past-policy effects (e.g., policies which aimed at correcting the flow of pupils through the system) and of demographic features (e.g., more recent cohorts are smaller than older cohorts), and not because drop out rate is going up. The drop out rate for lower-primary education has gone down considerably lately but remains high (from 10.2% in 1999 to 6.5% in 2004); at upper-primary it has fallen more modestly (from 12.9% in 1999 to 10.6% in 2004); finally, in the secondary education level itself it has oscillated very slightly (from 16.4% in 1999 to 16.0% in 2004). And from Table 2, we realize that less than half of 15-17 year-old pupils are actually enrolled in secondary education, where they ideally should be. The bottom line is that there is still considerable

[8]Source: INEP/MEC.

Table 2: Brazilian educational system: quantitative indicators. 1992 and 2005.

Indicator	**1992**	**2005**	***Change (in %)***
Average years of schooling of people aged at least 25	4.9	6.6	34.69
Illiteracy rate among people aged 15 or more, in %	17.2	10.9	−36.63
Fraction of children (7-14 years old) at school, in %	81.8	95.4	16.63
Illiteracy rate among children (10-14 years old), in %	12.4	3.2	−74.19
Average delay (actual age minus ideal age for grade)	2.1	0.9	−57.14
Fraction of 7-14 year-old delayed more than 2 years, in %	37.4	12.2	−67.38
Gross enrollment rate* at school, in %	67.2	85.8	27.68
Gross enrollment rate* at primary school, in %	103.6	121.0	16.80
Gross enrollment rate* at secondary school, in %	41.9	89.9	114.56
Gross enrollment rate* at higher education, in %	10.6	28.3	166.98
Net enrollment rate** at school, in %	62.4	73.5	17.79
Net enrollment rate** at primary school, in %	81.4	94.7	16.34
Net enrollment rate** at secondary school, in %	18.3	46.4	153.55
Net enrollment rate** at higher education, in %	4.6	11.7	154.35

Source: PNAD (IBGE) data compiled by IETS (www.iets.org.br).

Not included: the rural areas of the Northern region, except those of the state of Tocantins.

* Gross rate: # of people enrolled in that level / # of people in that level's ideal age range.

** Net rate: # of people of ideal age range who are enrolled / # of people in the ideal age range.

room for improvements before celebrating a stabilization of the size of the system, with universal access to education and reasonable drop out rates.

Another insufficiency of the 'quantitative revolution' is that it has not led to a homogeneous situation across the country. There are huge variation across states and across metropolitan areas, as the following examples illustrate, using data corresponding to 2005.[9]

- Illiteracy rate among children (10-14 years old): 0.4% in the metropolitan area of Curitiba; 3.8% in that of Fortaleza;
- Fraction of children (7-14 years old) at school: 98.4% in the state of Santa Catarina; 89.1% in that of Maranhão;
- Average delay (actual age minus ideal age for grade): 0.4 years in the state of São Paulo; 1.7 years in that of Maranhão;
- Net enrollment rate at primary school: 97.5% in the metropolitan area of São Paulo; 92.9% in that of Salvador;
- Net enrollment rate at secondary school: 66.6% in the state of São Paulo; 22.1% in that of Alagoas;
- Net enrollment rate at higher education: 21.4% in the Distrito Federal (metropolitan area of Brasília); 9.9% in that of Salvador.

Given the elsewhere well-documented inequalities within cities in Brazil (illustrated before in this dissertation by the Gávea *versus* Rocinha indicators), we know the improvements have not been homogeneous within smaller territorial units.

We can sum up by saying that the evolution of Brazil's educational figures regarding flow and stock variables is impressive, but even in this respect there are still some challenges to be faced in order to actually provide universal access to education with reasonable drop out rates in the country.

In addition to those challenges, there is also the *quality* issue, already mentioned in this dissertation's Introduction, and which we bring back to the fore below.

2.3 The qualitative challenge

The picture we saw in section 2.2 is one of a change in flow variables, eventually leading to a change in stock variables. The figures suggest that a quantitative change has taken place, but they do not indicate it has been accompanied by a qualitative improvement. To

[9] PNAD (IBGE) data compiled by the Brazilian "think tank" IETS (www.iets.org.br).

justify the large amount of public funds spent with education (cf. figures in section 2.1), and for a series of other obvious reasons, it is not enough simply to make sure pupils go to school; they should be learning something there. Unfortunately, there are reasons to believe this is not happening in Brazil. As we said in the introduction, as compared to other countries, the quality of education in Brazil is very bad in terms of achievement in international exams such as PISA. And that is true, even if we restrict the samples to compare only the best students in each country (Franco, 2002; Waltenberg, 2005). And if instead of international comparisons, we restrict our attention to Brazilian data, the diagnosis is not different: most Brazilian pupils do not reach minimal achievement standards (defined by experts).[10]

Overviewing the recent evolution of Brazilian education, as well as its current situation, Schwartzman (2006) claims that, while issues of access, flow, and literacy (i.e., the quantitative ones shown in Table 2), should be relatively less important nowadays, they still are the most frequent topics in the public debate. In his view, issues related to contents and quality - which, according to him, are incidentally more difficult to address through policy - have not yet been granted the priority they deserve. He relates educational quality to the drop out phenomenon, questioning the common view according to which Brazilian children and teenagers quit public schools because of liquidity constraints. He suggests they drop out because their schools are 'bad' such that they have the feeling they are wasting their time there.

If the probability of dropping out actually were negatively correlated to the quality of education received, the case for caring about educational quality would be reinforced, but verifying such hypothesis is not within the scope of our study. In any case, for reasons explained in this dissertation's Introduction, we are interested in studying educational quality. Besides, for reasons developed in previous chapters, we would like to work with some form of 'educational outcome'. There are many educational outcomes: examination grades, total years of schooling, different certificates or degrees, income, social positions, and so on. We shall use the performance of students in standardized tests ("test scores"), which are relevant educational outcomes since they are maybe the best available proxy for educational quality. With the caveats mentioned in section 3.2 of the dissertation's Introduction, we assume hereafter that scores express relevant skills and competences of individuals, that they are a good proxy for the accumulated human capital embodied in each individual, and that they open the way for individuals to acquire certificates, enter college, find

[10] As reveal reports analyzing the results of Brazilian pupils in PISA and SAEB exams, which have been produced by INEP/MEC technical staff in the past 10 years.

jobs, and so on. As a metric, test scores present some good characteristics: they are measurable at different points of the schooling process; quite often they are comparable across-grades, across-ages, or even across-countries.

3 Assessing educational inequality in Brazil: preliminary steps

Before turning to the empirical sections of this chapter, let us first describe the data and raise a few methodological issues.

3.1 The SAEB dataset

The data we use come from 5 waves of a survey on Brazilian pupils' achievement. The SAEB ("Sistema de Avaliação do Ensino Básico", or Basic Education Assessment System) exam is organized by INEP, a research institute which reports to the Brazilian Ministry of Education, and it has taken place every two years since 1995. While SAEB datasets are not suitable for direct international comparisons, the objectives and statistical design, as well the procedures employed in the application of the tests, have been inspired by, and do not differ very much from, well-known cross-country assessments of pupils' performance, such as PISA, TIMSS/PIRLS, and LLECE.

SAEB consists of country-wide tests in reading and mathematics, coupled with a collection of data on relevant characteristics of students, teachers, principals, and schools. It focuses on the evaluation of pupils at three key stages of the schooling system: 4^{th} grade of primary school, 8^{th} grade of primary school, and 3^{rd} grade of secondary school (labeled hereafter 11^{th}). Each of these grades corresponds to the last year of a stage in the Brazilian schooling system (end of lower-primary, end of upper-primary, end of secondary). The recommended ages of pupils attending these grades are, respectively, 10, 14 and 17 years old. It should be noticed that during the period covered by our data schooling was mandatory in Brazil for children up to 14 years (regardless of the grade they were attending).

The database consists of repeated cross-sections (not panels) of samples of schools and students. The samples for each grade and subject are representative of the whole country, as well as of subnational units (regions, states, urban x rural, private x public, etc.). In a first step, schools that took part in SAEB have been randomly chosen. In a second step, inside each of these schools, one class has been randomly chosen. Each pupil passes an exam of only one of the subjects.

Pupils' test scores correspond to subject-specific scales elaborated by INEP staff together with teachers, researchers, and national and international survey experts. Possible scores range from 0 to 500, and are supposed to evaluate skills and abilities of students. The SAEB scale is continuous and hierarchical, which means that a pupil who achieves a certain score - say, 400 in the Portuguese test - has all the literacy/reading skills held by students who achieves, say, 150, 300 or 380, plus some additional skills. For example, he might be able to understand and interpret more complex texts than his peers who scored lower. Because of the invariance of the scale, pupils' scores are comparable across years and across grades. Scores are not comparable across subjects.

Table 3 contains descriptive statistics for the scores obtained by Brazilian students in all the years, grades and subjects in the SAEB exams. Scores have been standardized such that the average for the 8th grade in 1997 is 250.00 and the standard deviation is 50.00, in both subjects.[11] Table 3 also shows that:

- The datasets we have used expand via sample weights (column 'Weight'), to the population of pupils assessed by the SAEB exams.[12] Due to drop out, grade retention, and demographic dynamics, for all the years and for both subjects, the number of pupils decrease substantially with the grade.[13]
- For a given year and subject, average scores increase as we move from 4th to 8th grade, and the same is true for the 11th grade as compared to the two less advanced ones, which is hardly surprising, given that all pupils are evaluated on the same set of skills and older, more instructed, pupils tend to be more skilled on average. Moreover, as we will comment below, some sort of selection mechanism may have already operated at more advanced stages, but not as much in the less advanced ones (the 4th grade in particular).
- For a given grade and subject, average scores usually decrease (slightly) across cohorts from 1995 until 2001, and then increase

[11] The scores have been standardized by INEP before being released for academic use, and not by us.

[12] In the appendix, there is a table containing the fraction of the pupil's population actually represented by SAEB exams' sampling scheme.

[13] From INEP/MEC website, we know that in 1995 there were 3.8 million pupils enrolled in the 4th grade, but four years later, in 1999, there were only 2.9 million enrolled in the 8th grade, figures which illustrate the spread of the dropout phenomenon.

Table 3: Descriptive statistics of *SCORE*

Grade	Subject	Year	# Obs	Weight	Mean	Std Dev	Min	Max
4	Math	1995	11,886	2,955,941	190.62	40.57	95.27	383.98
		1997	23,535	3,556,306	190.80	43.61	80.37	377.95
		1999	21,572	3,755,076	181.00	40.84	81.53	355.93
		2001	57,258	3,689,237	176.26	45.85	59.84	367.25
		2003	46,131	3,764,153	177.13	44.82	66.42	369.98
	Reading	1995	12,033	2,976,778	188.28	46.68	62.62	348.29
		1997	23,404	3,556,306	186.46	45.81	60.66	362.88
		1999	21,542	3,755,076	170.73	44.69	59.12	353.04
		2001	57,254	3,689,237	165.12	48.70	45.76	343.62
		2003	46,067	3,764,153	169.42	46.85	51.43	340.57
8	Math	1995	14,609	1,551,771	253.24	47.86	114.03	415.92
		1997	18,806	2,512,018	250.00	50.00	125.68	432.55
		1999	17,890	2,785,020	246.36	47.33	108.87	431.89
		2001	50,300	3,002,272	243.38	49.62	124.48	422.84
		2003	36,908	3,191,223	244.97	51.04	116.32	428.24
	Reading	1995	14,705	1,555,726	256.05	51.75	76.04	407.41
		1997	18,862	2,512,018	250.00	50.00	89.67	407.43
		1999	17,920	2,785,020	232.90	45.73	90.19	407.04
		2001	50,492	3,002,272	235.17	50.28	78.21	399.03
		2003	37,009	3,191,223	231.96	49.68	91.88	399.08
11	Math	1995	9,049	805,427	281.94	53.69	144.44	460.69
		1997	8,136	726,982	288.70	59.36	178.71	463.22
		1999	11,788	1,842,251	280.29	56.21	159.63	471.39
		2001	36,152	2,067,147	276.71	55.96	177.73	456.31
		2003	26,187	2,113,641	278.68	57.48	148.32	459.33
	Reading	1995	9,171	815,234	290.01	54.18	90.58	427.13
		1997	8,147	726,982	283.86	55.55	115.95	435.53
		1999	11,890	1,842,251	266.57	53.23	140.49	438.95
		2001	36,263	2,067,147	262.34	53.29	117.98	426.52
		2003	26,219	2,113,641	266.67	52.43	114.24	428.81

a bit from 2001 to 2003. However, average scores do not vary extraordinarily across cohorts.[14] So the bad performance of Brazilian pupils - attested by low average scores in PISA, mentioned in the dissertation's Introduction - seems to be persistent over time.

In this chapter and in the next one we sometimes compare distributions of scores over time, and also across space (essentially across Brazilian regions and states). In order to avoid a tedious series of tables and figures, in most cases we report national and regional results regarding exclusively the year 2001, the mathematics exam, and the 8^{th} grade. Let us justify this choice.

Why 8^{th} grade?:

- It corresponds with the end of primary school and of mandatory schooling in Brazil (although the latter is only a rough correspondence, because the obligation is defined in terms of age and not of grade);
- The age of 8^{th} grade Brazilian pupils is close to that of international pupils who passed the PISA 2000 exam; it is thus interesting to interpret our results having a similar international dataset as a benchmark.

Why 2001?:

- The share of the pupils' population represented by the samples has increased from 1995 to 1997, and then from 1997 to 1999, and it has been stable since then. So, each of the three most recent waves were good candidates;
- An important reform (FUNDEF) modifying the mechanism of allocation of federal funds to primary schools started being implemented in 1998; it seemed to us that it made more sense to focus on data collected after the reform had been fully implemented, so that 1999 was discarded;
- Finally, in Part III of this dissertation, we use 2001 data, because 2003 datasets were not yet available when we started working on chapters 5 and 6. For these reasons, 2001 is emphasized in this and in the next chapter too.

Why mathematics?:

[14] The results of such across-years comparisons should be taken cautiously, given that the samples of the first wave (1995) were less extensive than those of subsequent waves for all grades; and those of the second wave (1997) were less extensive than those of subsequent waves for the 11^{th} grade (cf. appendix).

- It seems to be the case that schools can make more of a difference in the teaching of mathematics than in that of reading, given that the latter is, as far as we know, determined to a larger extent by family background. If this is true, from an policy perspective, it is interesting to focus on a subject in which school could matter more;
- There is a (sparse) evidence showing that math scores are the best academic predictor of future wages, at least in the US (Bedard and Ferrall, 2003).

3.2 A brief methodological note

Measuring and comparing *income* inequality consists of the following task, according to Cowell (1995) (our italics): "how to *aggregate* (evaluate) the (change in the) distribution of an *attribute* over individual *units* of the population". Defining the income recipient unit (e.g., household or individual?), the relevant attribute or income concept (e.g., income in a given month or lifetime's income?), and the aggregation method (e.g., statistical indices or social-welfare-based indices?) are the main sets of decisions that have to be taken into account in order to measure inequality of a particular distribution of income (or other attributes) and to compare it with other distributions, be it over time or across space.

Inequality and welfare indices have been typically developed in a framework in which income is the attribute. As we are interested in using test scores as attributes, we have tried to understand if and how dominance tools (both of inequality and welfare), as well as inequality indices, could be employed to analyse test scores distributions. We believe the income toolbox is suitable to be used in the analysis of inequality and welfare in terms of test scores. Let us now make some comments on specific steps of such an analysis.

If we are interested in evaluating inequality and welfare in test scores, choosing the *recipient unit* is not a problem. There is no need to calculate a 'per capita measure of education outcome' weighted by some sort of equivalence scale, since test scores are, by definition, expressed in an individual basis and cannot be shared by individuals. There are no household economies of scale involved. So we can take pupils as the recipient units without further concerns.

As for the *attribute*, in the income case we have to: (i) define the time unit (e.g., month, year or lifetime?); (ii) make spatial corrections (accounting for regional prices); (iii) make time corrections (deflating the nominal values). The choice of the *time unit* is not as important here as it is in the case of a flow variable, such as income. Educational achievement is essentially a snapshot stock variable, so that measuring it in a week's or in a year's length of time is a senseless distinction.

Should we make some kind of *spatial adjustment* of the kind which is typically made in income distribution studies, such as corrections for regional prices differentials? On the one hand, a similar set of skills is possibly valued differently according to the average level of skills at each community and, particularly, at each local labor market. A minimal set of reading skills might be very advantageous for a person who lives in a very backward community where she could provide services, for example, to all those that are not able to read. In a more developed community where most people complete secondary school, only being capable of reading is by no means enough to find a good job and to enjoy a 'good life'. We can clearly relate this issue to more general debates, namely, on education as a positional good - especially in the philosophical literature (Brighouse and Swift, 2004) - and on the hypothesis of education as a signalling device - in the economics literature - which have challenged some of standard human capital theory's assumptions (Arrow, 1973).

On the other hand, pupils'and workers' mobility might well limit the relevance of different valuations of the skills in each community. Although Brazil is a huge country, the language spoken throughout it is the same (Brazilian Portuguese) and the intra-country cultural differences are not remarkable. This substantial homogeneity paves the way for a large potential mobility. So if a pupil achieves very high scores in tests (especially in advanced stages of schooling), but lives in a poor municipality, which does not offer her very good opportunities (e.g., well-paid jobs), in principle she could migrate to another municipality in which the set of such opportunities would be broader. The opposite case might also happen. A person having a relatively low level of skills in the local labor market where she was born could very well migrate to another place where her skills would pay off. Hence, although it is true that the human capital level expressed by scores is meaningful at the local level (in relative terms), it is also relevant at more aggregate levels (state-wide, nationally, or even internationally).

Beyond the difficulties involved in adequately accounting for the two effects mentioned above, if we wanted to acknowledge the importance of regional differences on the value of skills, in order to do some kind of spatial adjustment, such corrections would probably be very demanding in terms of data, since we would have to determine quite precisely the equivalent of the 'purchasing power of income' in the education field, that is, the *actual, multidimensional, value of education* in each community or labor market.

As far as we know, there exists no uncontestable procedure allowing us to take into account spatial heterogeneity in order to capture the actual value of scores. Because of that, in this dissertation we work with raw scores, ignoring positional good or signalling aspect of scores. In the appendix, we briefly discuss how this issue could be addressed.[15]

Still related to the definition of the attribute, *time correction* is frequently put forward when one is interested in measuring income inequality over the years. In the case of distributions of educational achievements, some kind of deflation procedure would need to be used to correct for variations over time of the 'real value of the score'. For example, a score of 300 points of score obtained ten years ago in a given place might correspond - in terms of actual value - to, say, a score of 320 points today, if for example the average score has increased since then and/or the demand for skills in the community at stake has augmented. The same difficulties related to correcting scores mentioned above (spatial adjustments) apply in the case of corrections for inflation. How should the deflation index be defined? The position we adopt is the same one we explained in the previous paragraph.

As for the *aggregation method*, studies relying on social-welfare-based indices and on the axiomatic approach allow one to understand more precisely the effects of a change in the relevant parameters, since such changes are normatively interpretable. A parameter shift in Atkinson indices represents an increase (decrease) in inequality aversion, whereas a shift in the parameters of the generalized-entropy class (for short: GE-class) indices represents a change in the relative weight attributed to inequality in specific ranges of income distributions, with indirect normative interpretations. Atkinson-class indices can be related to the CES family - Atkinson-class indices represent SEFs which range from maximin to the SEF analogous of the utilitarian SWF. A subset of the GE-class, $GE(\alpha)$, indices is ordinally (but not cardinally) equivalent to the Atkinson-class, $A(\epsilon)$, with $\alpha = 1 - \epsilon$, so that a ranking established using a GE-class index has a unique correspondence with a ranking established using an Atkinson-class index.

- $GE(-1)$ is ordinally equivalent to $A(2)$, an index which expresses a normative view closer to maximin, given that it is quite sensitive to inequality in the bottom of the distribution;
- $GE(1)$ is ordinally equivalent to $A(0)$, an index which is closer to a utilitarian-like viewpoint (homogeneously sensitive to income differences along the distribution).;

[15] Trying to integrate a positional aspect of scores can be interpreted as challenge to the appropriateness of the third crucial assumption discussed in chapter 1, namely, that of self-centeredness.

- $GE(0)$ corresponding to $A(1)$, it occupies an intermediary position, balancing distributive and aggregative concerns;
- $GE(2)$ is out of the range of Atkinson-class indices ($\epsilon \geq 0$); it is also known as the "Herfindahl index", an index which is more sensitive to inequality in the top of the distribution.

In this and in the next chapter, most often we make use of GE-class indices ($GE(0)$ in particular), a class of indices which is well-known for their good mathematical properties.

Whenever relevant, we try to establish dominance relations, by means of Lorenz and generalized Lorenz curves, which are standard tools in inequality and welfare analysis.

4 Intensity of test scores inequality in Brazil

4.1 Inequalities across grades

Welfare assessments do not make much sense for across-grades analysis, since it is obvious that the distribution corresponding with a more advanced grade - driven by much higher means - will dominate that of a less-advanced grade distribution in the generalized Lorenz sense. Older pupils are expected to perform better than younger, less instructed, ones, and they do that on average. So we proceed to compare Lorenz curves.

Graphically, it is very difficult to distinguish each grade's Lorenz curve for the 2001 mathematics exams. But by checking Lorenz curves deciles, we find both 8^{th} and 11^{th} grades Lorenz-dominate 4^{th} grade. However, nothing can be concluded in terms of dominance when we compare grades 8 and 11. So the 4^{th} grade is surely the most unequal one, but the two others could be ranked differently by different SEFs. The pattern found for 2001-math is repeated for all other years in math and reading, that is, for every year, $L_{11th} \approx L_{8th} \succ L_{4th}$.

We report in Table 4 the values obtained with a specific inequality index, $GE(0)$, as well as the ranking of grades for each year and subject, and finally the ratio between the index for the 4^{th} grade and the index for both other grades. We observe that the 4^{th} grade invariably occupies the third position (i.e., most unequal grade). Moreover, we observe that the magnitudes of the values of the index for the 4^{th} grade are substantially larger than those for the other two grades (ratios ranging from 1.20 to 2.16).

So the main finding is that the 4^{th} grade presents higher inequality than the two more advanced grades. And such pattern holds *for*

Table 4: Inequality across grades, according to $GE(0)$.

Math	Grade	$GE(0)$	Ranking	Ratio ($4^{th}/N^{th}$)
1995	4^{th}	0.0219858	3	1
	8^{th}	0.0183030	2	1.20
	11^{th}	0.0177719	1	1.24
1997	4^{th}	0.0258336	3	1
	8^{th}	0.0202355	1	1.28
	11^{th}	0.0209434	2	1.23
1999	4^{th}	0.0253769	3	1
	8^{th}	0.0186683	1	1.36
	11^{th}	0.0196061	2	1.29
2001	4^{th}	0.0342039	3	1
	8^{th}	0.0205947	2	1.66
	11^{th}	0.0197557	1	1.73
2003	4^{th}	0.0324836	3	1
	8^{th}	0.0216727	2	1.50
	11^{th}	0.0208752	1	1.56
Reading				
1995	4^{th}	0.0323691	3	1
	8^{th}	0.0230860	2	1.40
	11^{th}	0.0195005	1	1.66
1997	4^{th}	0.0304152	3	1
	8^{th}	0.0215027	2	1.41
	11^{th}	0.0207661	1	1.46
1999	4^{th}	0.0344817	3	1
	8^{th}	0.0198462	1	1.74
	11^{th}	0.0202691	2	1.70
2001	4^{th}	0.0464863	3	1
	8^{th}	0.0245373	2	1.89
	11^{th}	0.0215447	1	2.16
2003	4^{th}	0.0405682	3	1
	8^{th}	0.0242030	2	1.68
	11^{th}	0.0202472	1	2.00

all the years, and *in both subjects*. On the one hand, an optimistic reading of such result would consider it was somewhat expected, since in the first grades students are typically more heterogeneous, given that schools have not yet been able to reduce the spread of pupils's performance, an objective which would be more realistic for advanced grades' pupils. On the other hand, a more pessimistic reading would consider that students are more heterogeneous in preliminary stages in Brazil, because selection - and its probable corollary: drop out - have not yet fully intervened. Indeed, drop out is less severe a problem in less advanced grades (as the figures in previous sections reveal).

What is worrying is the serious likeliness that the selection is a *socio-economic one*, such that the probability of dropping out is correlated with pupils' socio-economic characteristics. As a quick check of that hypothesis[16], we compared across grades, the average level of pupils' mother's educational level, which is a usual indicator of socio-economic status in the economics of education literature. For pupils who passed the math exam in 2001, mother's education increases with the grade. In a scale which ranges from 1 (not studied) to 5 (college education), we have average mother's education levels of: 2.73 for the 4$^{\text{th}}$ grade; 2.90 for the 8$^{\text{th}}$ grade; and 3.04 for the 11$^{\text{th}}$ grade. That gives some credit to the hypothesis that the selection process is socio-economically biased.

4.2 Inequalities over time

Presenting SAEB descriptive statistics, we saw that for a given grade and subject, average scores do not vary very much across the five waves (1995 through 2003). Let us now understand how scores *inequalities* have evolved over the years, by trying to establish 6 time-series rankings (3 grades x 2 subjects). This is an interesting question, because the enrollment rates have increased in the 1990s, leading to more socio-economic heterogeneity of students, and presumably to increased inequality in achievement.

Once again, we first search for dominance relations by turning to pair-wise comparisons of distributions using Lorenz curves.[17] In *mathematics*, out of 30 pair-wise Lorenz comparisons (10 per grade), 19 show dominance. The distribution of 2003 is the one which is most frequently dominated, in 8 out of 12 pair-wise comparisons (3 times its Lorenz curve crosses; and once, for the 4$^{\text{th}}$ grade, it dominates 2001).

[16] A rigorous check of such hypothesis would require a more careful analysis, which is out of the scope of this study.

[17] The results of such across-years comparisons should be taken cautiously, cf. a previous footnote, and the appendix on the extent of the SAEB samples.

In *reading*, out of 30 pair-wise Lorenz comparisons only 13 indicate dominance, 8 of which in the 4^{th} grade. The 2001 and 2003 samples are more frequently dominated, while 1995, 1997, and 1999 dominate more frequently.

We report in Table 5 the values obtained with a specific inequality index, $GE(0)$, as well as, for each grade and subject, the ranking of the five waves.

Table 5: Inequality over time, according to $GE(0)$.

Year	Math	Ranking	Reading	Ranking
4^{th} Grade				
1995	0.0219858	1	0.0323691	2
1997	0.0258336	3	0.0304152	1
1999	0.0253769	2	0.0344817	3
2001	0.0342039	5	0.0464863	5
2003	0.0324836	4	0.0405682	4
8^{th} Grade				
1995	0.0183030	1	0.0230860	3
1997	0.0202355	3	0.0215027	2
1999	0.0186683	2	0.0198462	1
2001	0.0205947	4	0.0245373	5
2003	0.0216727	5	0.0242030	4
11^{th} Grade				
1995	0.0177719	1	0.0195005	1
1997	0.0209434	5	0.0207661	4
1999	0.0196061	2	0.0202691	3
2001	0.0197557	3	0.0215447	5
2003	0.0208752	4	0.0202472	2

We cannot identify a clear pattern of increasing or decreasing inequality over time for 8^{th} and 11^{th} grades, since the indices for different years oscillate very slightly from one year to another. However, for the 4^{th} grade, we identify a net increase in the extent of inequality when we move from older distributions (1995-1999) to more recent ones (2001-2003). *This result suggests the high inequality in achievement Brazil presents* - attested by comparisons with other countries that took part in the PISA exams, already and mentioned in the dissertation's introduction - *is persistent over time*.

Combining the results of this subsection with those from the previous one, we can say that the analysis of distributions of educational quality in Brazil (as expressed by test scores) shows, not only that the 4^{th} grade is the most unequal one, but also that inequality in that level is larger in most recent years. This is bad news right now and also for the coming years - a highly unequal group of pupils is evolving through the system. The good news, to the extent that there is one, is the fact that the peak of inequality in the 4^{th} grade occurred in 2001, and it has fallen from then to 2003 (a result which is confirmed by other indices). When new SAEB waves will be available, it will be interesting to check if such trend is confirmed.

As a summary of what we found up to this point, we can confirm that the quality of education, as expressed by test scores, in Brazil is deficient. Such deficiency express itself not only through low average performance, but also through unequal performance. There has been no improvement in recent years and there is no expectation of a decrease in the levels of inequality in the system as a whole in the close future, since the less-advanced grade is the most unequal one and will most probably, either carry out and accumulate inequality through the system, or 'eliminate it' through drop out.

Now let us identify more precisely where in the country the problems are more severe.

5 Sub-national inequality

We now report inequalities within subsets of the Brazilian territory. Most often, we measure scores inequality in the five geographical regions in which Brazil is traditionally divided for a series of pratical matters (but such division lacks a political content). These regions are further divided into 27 states, which are the political units which compose the federation.[18]

We have already alluded to intra-country inequalities in Brazil, both in this chapter (when we argued that the 'quantitative revolution' that took place in Brazilian education in recent decades has not been homogeneous across the country) and in the dissertation's Introduction (Gávea *versus* Rocinha indicators illustrating inequalities within a city). However, it should be noticed that these are just particular examples of a vast phenomenon of intra-country inequalities in

[18]Each state is composed of municipalities, which is the smallest official government unit. There are 5,560 municipalities in the country.

Brazil.[19] To illustrate our point, we present in Table 6 the population share and the GDP share of the five geographic regions in 2000, together with the average scores obtained in the SAEB (2001-math-8$^{\text{th}}$ grade exam), and the ranking of regions in terms of these average scores.[20]

Table 6: Regional indicators. Brazil, 2000-2001.

Region	Population	Share	GDP Share	Scores	Ranking
South-East	72,412,411	42.65	57.8	249.72	2
North-East	47,741,711	28.12	12.9	228.79	5
South	25,107,616	14.79	17.6	255.34	1
North	12,900,704	7.60	4.6	231.86	4
Mid West	11,636,728	6.85	7.0	244.83	3
Brazil	169,799,170	100.00	99.9	243.38	-

Population and GDP: Contas regionais, IBGE (2000).

Scores: SAEB (2001); math and 8$^{\text{th}}$ grade sample.

1. *South-East*: this is the richest region in the country, led by three dynamic cities (São Paulo, Rio de Janeiro, and Belo Horizonte), which, together with their surrounding municipalities, form the most populated metropolitan areas in the country. This region's population is higher than that of many countries, and amounts to 42.65% of Brazil's population. The share of the GDP produced in such region is very large, well above its population share. The state of São Paulo alone responds to around 1/3 of Brazil's GDP in 2000. The big cities of the South-East are very unequal in many respects (e.g., where most Brazilian *favelas* are located, side by side with upper-class neighboorhoods). Average scores in SAEB are relatively high, but not the highest.

[19] These regional inequalities are well-documented, for example, in Mendonça (2006).

[20] We have also calculated similar rankings of regional average scores for every combination of year, subject and grade (5 years x 2 subjects x 3 grades = 30 rankings). For each year, subject and grade, we have attributed 1 to the highest average score, 2 to the 2$^{\text{nd}}$ highest and so on. We then took the average out of the 30 distributions. The average values obtained are the following: South: 1.53; South-East: 1.93; Mid-West: 2.57; North: 4.47; North-East: 4.50. The most frequent ranking reversals are observed between North and North-East, and between South and South-East. According to the average ranking obtained out of these 30 rankings, the 5 regions are classified exactly as the particular ranking of Table 6, showing that 2001-math-8$^{\text{th}}$ grade ranking is, in some sense, a representative one.

2. *North-East*: for historical reasons, this is the poorest region and, in many respects, a very unequal one. The GDP share is less than half the population share. It hosts a very large population - almost 50 million people- and three metropolitan areas: Salvador, Fortaleza, and Recife. This region's average performance in the SAEB is the lowest in the country;

3. *South*: it is a relatively rich and egalitarian region, for Brazilian standards. Two metropolitan areas (Porto Alegre and Curitiba) are located there. The average performance in SAEB is the highest in the country;

4. *North*: a poor region (GDP share smaller than population share), with low density of population. The Amazon Forest is located there. The only metropolitan area is Belém. Average results in SAEB are low, very close to those of the North-East;

5. *Mid-West*: a very heterogeneous region, with both very poor territories, and very rich zones, such as the metropolitan area of Brasília (Distrito Federal).[21] In terms of test scores, this region occupies an intermediary 3$^{\text{rd}}$ position.

5.1 Decomposition of inequality

We have decomposed inequality into between- and within- relevant territorial units, employing GE-class indices. While there is some inequality of educational achievement across Brazilian geographical regions, Panel I in Table 7 indicates that the greatest fraction of measured inequality is observed within each region, accounting for around 96% of the overall inequality.

Is the fraction corresponding to between-regions inequality too small? To answer that, we would need a basis for a comparison (e.g., studies for other countries), which we do not have. We tend to think it is not so small, given that the groups we are dealing with here (the regions) contain a huge number of pupils, they were likely to present highly heterogeneous performance within them. Yet, around 4% of the overall inequality comes from inter-regional inequality.

Panel II reports inequality decompositions across Brazilian states, and the result is very similar to the one obtained for regions. In Panel III, in turn, we can see that the fraction of overall inequality observed between municipalities is much larger (around 1/5 of overall inequality).

[21]It should be mentioned that in recent years, high growth rates have been registered in this region, and the development of the agribusiness has played a major role.

Table 7: Inequality decomposition. Brazil, SAEB 2001, Math, 8th grade.

Panel I: Inequality across the 5 regions

	GE(-1)	GE(0)	GE(1)	GE(2)
Overall	0.02106	0.02059	0.02051	0.02078
Within regions	0.02017	0.01971	0.01963	0.01992
Between regions	0.00089	0.00088	0.00087	0.00087
Within/Total	95.77%	95.73%	95.71%	95.86%
Between/Total	4.23%	4.27%	4.24%	4.19%

Panel II: Inequality across the 27 states

	GE(-1)	GE(0)	GE(1)	GE(2)
Overall	0.02106	0.02059	0.02051	0.02078
Within states	0.01999	0.01954	0.01946	0.01974
Between states	0.00106	0.00106	0.00105	0.00104
Within/Total	94.92%	94.90%	94.88%	95.00%
Between/Total	5.03%	5.15%	5.12%	5.00%

Panel III: Inequality across municipalities

	GE(-1)	GE(0)	GE(1)	GE(2)
Overall	0.02106	0.02059	0.02051	0.02078
Within municipalities	0.01693	0.01644	0.0163	0.0165
Between municipalities	0.00412	0.00414	0.00419	0.00427
Within/Total	80.39%	79.84%	79.47%	79.40%
Between/Total	19.56%	20.11%	20.43%	20.55%

Inequality within sub-national units are substantial, even when we go down to the smallest territorial unit used here. For this reason, it is interesting to compare some of these sub-national units to find out in which of them inequality is more intense.

5.2 Regional inequalities

We showed above that the rankings of regional *average scores* were similar for different combinations of year, subject and grade. We also showed that a great fraction of inequality is found within sub-national units. Now as a further step in our mapping of educational inequalities, we try to understand in more detail the phenomenon of intra-country *scores inequalities*, first by ranking Brazilian regions.

Plotting Lorenz curves for Brazil's regions is not very elucidative, given that the five distributions are not sufficiently different in terms of inequality for us to observe their differences graphically. GL curves are also problematic in this respect: the only information we obtain from a visual inspection of GLs is a reproduction of the ranking of average scores for the five regions.

By comparing deciles of the distribution of Lorenz and GL curves, however, we can establish rankings of educational inequality and educational welfare (Table 8). Panel I shows educational welfare rankings. The GL curve of the Southern region ranks first, the Northern ranks fourth, and the North-Eastern ranks fifth; South-East and Mid-West GL curves cross. Thus when we take the mean score into account, we can order almost all the distributions, with only one crossing. The drawback of such ranking is that it is essentially driven by the average score, and the importance of inequality is downgraded. As a matter of fact, we observe this ranking is similar to the average score ranking.

The educational inequality ranking (Panel II) shows the crossing of Lorenz curves of three regions - Mid-West, North and South. These three regions Lorenz-dominate the North-East and the South-East, whose Lorenz curves also cross with each other.

While a full dominance ranking cannot be established because of crossings of Lorenz curves, we can have an idea of which regions are more unequal than the others. The main result here is that the highly-populated South-East and North-East are the most unequal regions for a wide class of SEFs.

The values of the deciles of both Lorenz and GL curves for Brazil's regions are very similar from one region to another. Possibly, a number of the dominance relations established by means of point estimates comparisons would not be confirmed if statistical inference procedures were applied. In such case, the rankings could be even less complete.

Table 8: Regional inequalities: dominance rankings (SAEB 2001, Math, 8th grade)

Panel II: Educational 'welfare'

Ranking of Brazilian regions based on deciles of GL curves

South
||
South-East — Mid-West
||
North
||
North-East

.

Panel I: Educational inequality

Ranking of Brazilian regions based on deciles of Lorenz curves

Mid-West — North — South
||
North-East — South-East

This leads us to complement the analysis using specific inequality indices, which allow us to establish complete rankings, at the cost of more normative controversies. However, this procedure allows us to observe the sensitivity of the rankings to shifts in relevant parameters (e.g., inequality aversion, and related ones).

In Table 9 we report inequality indices for the distributions of scores both at the national (first line) and at the regional level (remaining lines).[22]

The magnitude of the indices shown in Table 9 is low as compared to those we usually observe with income distribution data. For example, the highest score-Gini found here (0.11737 for the South-Eastern region) is much lower than the income-Gini of any country in recent years.[23]

Brazil's score-Gini index shows an intermediate level of inequality for this grade-year-subject (0.11501), as compared to those of other

[22] Atkinson indices are not reported here, but are available.

[23] Barros, Henriques, and Mendonça (2000), for example, present Gini indices for around 90 countries for the late 1980s, and the lowest ones (Slovakia, Czech Republic, and Ukraine) oscillated around 0.22-0.25.

SAEB samples (not reported here).[24]

When we turn to the additive decomposable GE-class indices, we observe that Brazil's indices are extremely influenced by the highly populated and highly unequal South-Eastern region's indices. Indeed, Brazil's index is always lower than that of the South-East, and higher than that of any other region.

When we compare regions, the ranking is robust to most variations of parameters. The only exception is to be found when we turn to an index which strongly penalizes inequalities among low-performing students, namely, $GE(-1)$. In the latter case, the Northern region (slightly) outperforms the Southern region.

Table 9: Inequality indices (GE-class and Gini) for Brazil and its 5 regions. SAEB 2001, Math, 8th grade.

Index:	**GE(-1)**	**GE(0)**	**GE(1)**	**GE(2)**	**Gini**	**Ranking according to**	
Brazil	0.02106	0.02059	0.02051	0.02078	0.11501	**GE(-1)**	**Other**
Region:							**indices:**
S	0.01696	0.01646	0.01621	0.01620	0.10209	2	1
N	0.01691	0.01661	0.01656	0.01675	0.10340	1	2
MW	0.01769	0.01732	0.01725	0.01746	0.10489	3	3
NE	0.01983	0.01965	0.01981	0.02033	0.11261	4	4
SE	0.02211	0.02152	0.02133	0.02154	0.11737	5	5

The general pattern that emerges from this subsection's rankings allows us to draw some conclusions on Brazil's regional inequalities:

1. The rich South-Eastern region, which ranks second in terms of average scores, is the most unequal one according to all indices used here;
2. The poor and low-performing North-Eastern region is also one of the most unequal ones, occupying a stable fourth position;
3. In addition to presenting the highest average scores (cf. section 5), the Southern region is also the least unequal one according to most rankings;

[24] The highest score-Gini is the one associated with the 2001-reading-4$^{\text{th}}$-grade sample: 0.1673. The lowest one is for the 1995-reading-11$^{\text{th}}$-grade sample: 0.1054.

4. The Mid-Western region occupies an intermediary position, both in terms of average scores and of scores inequality;
5. While its record of average scores is bad, the Northern region shows a relatively egalitarian record in terms of test scores;
6. The magnitude of the difference is not negligible, at least between the extreme cases: the most unequal region (SE) is 30.76% more unequal than the least unequal one (S), according to $GE(0)$.

5.3 Inequalities in Brazilian states

As the final piece of evidence in our mapping of Brazilian educational inequalities, we compiled rankings of the 27 Brazilian states based on GE-class indices. If a state moves up in the ranking (i.e., it is classified as less unequal) as α moves from -1 to 2 in the GE-class indices, it means that this state has relatively more dispersion at the bottom than at the top. And the opposite holds for those states which move down. The ranking of scores inequality among Brazilian states is reasonably robust to changes in the value of the parameter α. When we move all the way from $GE(-1)$ to $GE(2)$, states occupying extreme positions do not move, and most states move less than 3 positions. The exceptions are: Bahia and Alagoas (down 3 positions), Santa Catarina (up 3 positions), Minas Gerais (up 4 positions), Maranhão (down 6 positions), and Tocantins (up 6 positions). Such results correspond to what one would expect: overall more developed states going up; overall less developed states going down. (Tocantins going up is the only somehow unexpected result.)

Table 10 shows the results for $GE(0)$ index, which as we know, occupies an intermediary position in the range of GE-class indices. Among the less unequal states, we find the barely populated Northern states of Amapá, Acre and Rondônia, but also the most important and most populated state of the Southern region, Rio Grande do Sul (6.00% of the country's population). The three states which compose the Southern region are relatively well-placed (i.e., among the eleven less unequal states), as well as most of the Northern states. The four Mid-Western states are located in intermediary positions, ranging from the 5^{th} to the 16^{th} position.

The most unequal unit of the federation is São Paulo, the state which responds to the largest share of the country's GDP (1/3) and population (1/5). The ten most unequal states all belong to only two regions, the South-East and the North-East, previously identified as the most unequal of the country. The range of measured inequality levels in the case of states is much wider than that we found for regions,

Table 10: Ranking of Brazilian states according to $GE(0)$. SAEB 2001, Math, 8th grade.

State	Region	$GE(0)$	Ranking
Amapá	N	0.01312	1
Rio Grande do Sul	S	0.01464	2
Acre	N	0.01491	3
Rondônia	N	0.01505	4
Goiás	MW	0.01513	5
Santa Catarina	S	0.01534	6
Pará	N	0.01568	7
Roraima	N	0.01661	8
Amazonas	N	0.01721	9
Mato Grosso do Sul	MW	0.01736	10
Paraná	S	0.01817	11
Espírito Santo	SE	0.01836	12
Bahia	NE	0.01844	13
Alagoas	NE	0.01862	14
Distrito Federal	MW	0.01863	15
Mato Grosso	MW	0.01886	16
Tocantins	N	0.01888	17
Sergipe	NE	0.01901	18
Maranhão	NE	0.01904	19
Paraíba	NE	0.01920	20
Ceará	NE	0.01969	21
Rio de Janeiro	SE	0.01969	22
Minas Gerais	SE	0.01978	23
Pernambuco	NE	0.02002	24
Piauí	NE	0.02118	25
Rio Grande do Norte	NE	0.02174	26
São Paulo	SE	0.02289	27

since the most unequal state is 74.47% more unequal than the least unequal one (30.76% in the case of regions).

These findings have allowed us to map more precisely where in the country score-inequalities are more intense. They reinforce those coming from the regional analysis: the most unequal regions are indeed composed of the most unequal state.

The results regarding national and sub-national educational inequalities in Brazil provide a basis for future research which would compare the figures reported here with other intra-national, national, and international figures. Future research would also help putting into perspective the importance and magnitude of the inequality levels we have computed (e.g., how high is inequality in Brazil, in its regions, and in its states?). More particularly, given that we have identified the more unequal sub-national units in our mapping exercise, our research can be a useful basis for further research on score-inequality measurement in particular regions, sets of states, states, or even municipalities.

6 Conclusions

Brazil's educational system is enormous and complex, involving more than 50 million students. While a pervasive improvement in quantitative indicators (access, flow, literacy rates) have taken place recently, there is still room for extending and deepening the improvements in those respects. In particular, the change has not led to a homogeneous situation space-wise.

Beside working on those - important - quantitative aspects, time has come for researchers to care also about qualitative issues too. In terms of quality, Brazil performs poorly in international comparisons, and also according to theoretical achievement standards determined inside the country.

In order to explore qualitative issues, we have turned to the SAEB test scores database and to inequality and welfare analysis toolbox. Employing such data and tools, we have assessed the intensity of inequalities in terms of educational quality in the country, both at the national level, and at sub-national levels.

Our results suggest that distributions of less advanced grades are more unequal than those of more advanced grades, especially in recent years, a situation which is particularly challenging for the coming years.

When we decompose national inequality space-wise, we find that inequality within regions or within states account for a very large fraction of overall inequality (95-96% of overall inequality), while inequality across municipalities is smaller (around 4/5 of overall inequality).

In our rankings of scores inequalities, the rich South-East and the poor North-East - the most populated regions of the country - occupy the last positions, that is, they are the most unequal ones. States located in such regions are also the most unequal ones. São Paulo - the richest and most populated of the 27 units which compose the federation - is the one showing more inequality.

In the next chapter, we focus on a refinement of the inequality measurement exercise. More precisely, we decompose overall inequality, not only in terms of territorial units, but also in terms of personal characteristics, in order to identify the fraction of overall inequality which is acceptable (allegedly caused by differential personal effort) and the fraction which is unacceptable (allegedly caused by differential personal circumstances). By doing so, we link Part II, which is devoted to measurement, to the normative concerns of Part I.

A Appendix: Extent of SAEB samples

The samples of the first wave of SAEB (1995) were less extensive than those of subsequent waves for all grades (except 11^{th} with respect to 1997). The samples of the second wave (1997) were less extensive than those of subsequent waves for the 11^{th} grade, as Table 11 indicates.

Table 11: Proportion of the pupils' population represented by SAEB samples. 1995-2001.

	Subject and grade					
Year	**M.-4th**	**R.-4th**	**M.-8th**	**R.-8th**	**M.-11th**	**R.-11th**
1995	76.82	77.37	71.54	71.72	68.12	68.95
1997	86.45	86.45	99.41	99.41	50.31	50.31
1999	87.08	87.08	95.21	95.21	97.74	97.74
2001	84.97	84.97	93.20	93.20	96.64	96.64
2003	89.62	89.62	97.46	97.46	95.49	95.49

B Appendix: Positional or signalling adjustments

All the results we provided in this chapter were based on calculations employing the original (or raw, or absolute) scores obtained by individuals. However, for reasons explained in section 3, we thought it would be interesting to develop a procedure to adjust scores, in order to express, together with the (potential) human capital they reflect, their *local "positional" or "signalling" aspect.* Those two elements taken together would reveal the "true local value" of the score achieved by a pupil. Such correction would correspond with the time and spatial adjustments mentioned in section 3.2. As far as we know, there exists no uncontestable procedure allowing us to do so, and this is why we work in this dissertation only with raw scores.

We have thought about two alternatives. The first one would consist of correcting the scores simply by dividing them by the average score obtained in the relevant local community. However, such correction would not differ much from eliminating the "between component" in standard decomposition. That was not exactly what we were searching for, because in such case only the local relative value would matter, while what we wanted was to account, *both for the relative and for the absolute* value of scores.

A second alternative would consist of correcting the scores through a formula of the following kind:

$$AS_i = S_i \times (q_i)^\alpha \tag{1}$$

where:

- AS_i is the adjusted score of individual i;
- S_i is the original score, which represents the "embodied human capital" component of AS_i;
- $(q_i)^\alpha$ is a term aimed at correcting the original score to its local value (by means of a bonus for those individuals well-placed in their local distribution of scores). This is the "signalling (or positional) correction" component of AS_i, with:

  α ($\alpha \geq 0$): a parameter expressing the importance given to the local positional aspect; with $\alpha = 0$ there is no correction at all (whatever the q), but as α increases, the bonus increases; for a given α, the bonus increases exponentially with the quantile;

  q ($q = 1, \ldots, Q$): the quantile of performance to which individual i pertains at her local level; for those in the first quantile (with a relatively low performance), α does not contribute much to raising AS_i (or not at all for the first quantile), but for higher quantiles, it does contribute.

Table 12 shows the bonus (in percentage points) for different values of the parameters, assuming we work with deciles (the 1$^{\text{st}}$ decile corresponding with individuals with the worst relative performances).

As an example, let $\alpha = 0.050$, and consider an individual who achieved 200 points of score. If she lived in a low-performing community, such that her score placed her in the 10$^{\text{th}}$ decile, her positional score would be 224.40 (bonus of 12.20%, or 24.40 points). However, if she lived in a high-performing community, such that that same score placed her in the 1$^{\text{st}}$ decile, her positional score would not be affected by any bonus, remaining at 200. If she occupied an intermediary relative position - say, 3$^{\text{rd}}$ decile - her positional score would be 211.29 (bonus of 5.65%, or 11.30 points).

We have not yet been able to study more deeply this second alternative's details, such as the advantages and disadvantages of the particular specification given in Equation 1 above. Even if that (exponential) specification proved to be a convenient and adequate one, a number of issues would remain unsettled. For example, we have

Table 12: Bonus (in % of scores) for different values of the parameters.

α	1st decile	2nd decile	3rd decile ⋯	10th decile
0.000	0.00	0.00	0.00	0.00
0.001	0.00	0.07	0.11	0.23
0.025	0.00	1.75	2.78	5.93
0.050	0.00	3.53	5.65	12.20
0.100	0.00	7.18	11.61	25.89
0.250	0.00	18.92	31.61	77.83
0.500	0.00	41.42	73.21	216.23
1.000	0.00	100.00	200.00	900.00

no clue on how to set an appropriate value for α. Maybe we would have to estimate somehow the returns to the relative position in the local distribution of scores, possibly allowing α to vary across communities. Another path would consist of studying the literature on norm-referenced tests, in which scores of an individual are evaluated taking into account her relative performance in the test.

For these reasons, we could not conduct a distributional analysis using adjusted scores, and we leave this for future research.

References

Arrow, K. (1973): "Higher education as a filter," *Journal of Public Economics*, 2, 103–216.

Barros, R. P. d., R. Henriques, and R. Mendonça (2000): "A estabilidade inaceitável: desigualdade e pobreza no Brasil," in Henriques (2000).

Bedard, K., and C. Ferrall (2003): "Wage and test score dispersion: some international evidence," *Economics of Education Review*, 22, 31–43.

Brighouse, H., and A. Swift (2004): "Equality, Priority and Positional Goods," Equality Exchange Network (Draft).

Coleman, J., E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld, and R. York (1966): *Equality of Edu-*

cational Opportunity. US Government Printing Office, washington, dc, usa edn.

Cowell, F. A. (1995): *Measuring inequality*. Prentice Hall, Harvester, Wheatsheaf, 2nd edn.

Franco, C. (2002): "Educação das elites no Brasil: a Bélgica não existe," *Rio de Janeiro: Trabalho e Sociedade (IETS)*, 2(4), 13–16.

Henriques, R. (ed.) (2000): *Desigualdade e pobreza no Brasil*. Instituto de Pesquisa Econômica Aplicada, Rio de Janeiro, Brazil.

Ioschpe, G. (2004): *A Ignorância Custa um Mundo: o Valor da Educação no Desenvolvimento do Brasil*. Francis, São Paulo, Brazil.

Mendonça, R. (2006): "Em direção às metas de desenvolvimento do milênio. Uma análise regional," *Sinais Sociais - SESC*, 1(1), 62–113, Rio de Janeiro.

Oliveira, J., and S. Schwartzman (2002): *A Escola Vista por Dentro*. Alfa Editora, Belo Horizonte, Brazil.

Schwartzman, S. (2006): "Educação e pobreza no Brasil," *Cadernos Adenauer*, 7(2), 9–37.

Soares, S. (2002): "Avaliação educacional como instrumento pedagógico," *Rio de Janeiro: Trabalho e Sociedade (IETS)*, 2(4), 23–26.

Waltenberg, F. D. (2005): "Iniqüidade educacional no Brasil. Uma avaliação com dados do PISA 2000," *Revista Economia (ANPEC)*, 6(1), 67–118.