Rochester Center for
Economic Research

The Causes and Effects of Grade Repetition: Evidence from Brazil

Gomes-Neto, Jo?o Batista, and Eric Hanushek

Working Paper No. 295
September 1991

University of
Rochester

The Causes and Effects of Grade Repetition: Evidence from Brazil

Joao Batista Gomes-Neto and Eric Hanushek

Rochester Center for Economic Research
Working Paper No. 295

The Causes and Effects of Grade Repetition: Evidence from Brazil

by João Batista Gomes-Neto* and Eric A. Hanushek**

Rochester Center for Economic Research
Working Paper No. 295

August 1991

ABSTRACT

Developing countries are frequently plagued by problems of grade repetition in their primary schools, but the formulation of coherent policies founders on a lack of knowledge about either the determinants of repetition or the outcomes of repetition. This paper, relying upon a unique panel data set for northeast Brazil, documents the central role of student achievement in affecting repetition. The evidence further suggests that grade repetition improves individual performance, although in a very expensive way. Finally, preliminary investigation of mandatory promotion suggests that such policies diminish aggregate student performance but may be warranted if quality improvements in schools are not forthcoming.

*Universidade Federal do Ceará

**University of Rochester

The Causes and Effects of Grade Repetition:

Evidence from Brazil

by João Batista Gomes-Neto and Eric A. Hanushek

Even though the problem of grade repetition is high on the policy agenda of virtually every developing country, virtually nothing is known about either the causes or the educational effects of repetition. Fundamental disagreements about the nature of the problem have clearly inhibited the development of sensible policies. This research, relying on a unique data set for northeast Brazil, considers how the schooling system and individual students interact in determining the enrollment patterns in the primary schools of Brazil. This lays the groundwork for analysis of alternative policies.

High drop-out and repetition rates, collectively referred to as wastage, have been identified as one of the main failures of the Brazilian education system, in part because the rates appear above those in other countries. But disagreement about the nature of the wastage problem arises immediately. There are important discrepancies even among estimates of the level of aggregate wastage and its components. The Brazilian Ministry of Education (MEC), for example, estimates the drop-out rates for the first grade of primary education to be around 25 percent and the repetition rates to be around 30 percent in 1982, suggesting that dealing with drop-outs is the first step toward fixing the Brazilian education system. On the other hand, Fletcher and Ribeiro [1988], who use a statistical model (PROFLUXO), estimate the drop-out rate for the first grade at about 2 percent and for the repetition rate at about 55 percent, leading them to conclude that repetition is the main problem in the

Brazilian education system.[1] Other researchers[2] and even other government agencies (e.g., the *Fundação Instituto Brasileiro de Geografia e Estatística*, or IBGE) also question the MEC estimates and policy conclusions. Such differences also show up in more detailed investigations for specific regions of Brazil (see, for example, Kafuri *et al.*[1985]). The aggregate estimates of drop-out rates differ not only in level but in pattern. The estimates of the Ministry of Education suggest that drop-out rates decline with grade level, while the other estimates indicate drop-out rates increasing with grade level.

Disagreements continue about what are the major causes of wastage. Some concentrate on problems with the school system, while others turn to factors outside of the control of schools. A variety problems have been identified as the main out-of-school causes for school failures, and, importantly, each is directly related to socio-economic status of the student[3]. High direct costs -- for example, for buying uniforms, writing materials, textbooks, and the like -- and sensitivity to the opportunity costs of attending school are more likely to strike the children from impoverished backgrounds. Other authors also identify malnutrition, which is clearly related to the social and economic status, as one of the causes of the school failures (Cunha [1981], Carvalho [1983]).[4]

The in-school explanations concentrate on specific resource constraints and the general low quality of some schools. Many researchers have pointed to problems of low quality teachers as measured by low levels of education, low salary and motivation, and poor attitudes and expectations

---

[1] For comparisons of the MEC, IGE, and PROFLUXO estimates, see Fletcher and Ribeiro[1988].

[2] See Fletcher and Castro[1986], Fletcher and Ribeiro[1988], Verhine and Melo[1988], and Schiefelbein[1989].

[3] Verhine and Melo[1988], for example, suggest that factors external to schools are the primary cause of first grade drop-out behavior. An alternative view concentrates on underlying political and social incentives, but these arguments go beyond our inquiry. Specifically, some authors posit that the educational system exists and was built in such way to maintain the status quo in Brazil's unequal order (Popovic [1980], Oliveira [1981], Garcia[1982]).

[4] On this, however, our previous studies did not find a strong relationship between malnutrition or receiving school lunch and student test achievement (Armitage *et al.* [1986], Harbison and Hanushek [forthcoming]), suggesting that any effects on wastage did not come directly from low performance induced by malnutrition.

(see, for example, Melo [1982], Brandao [1983], Armitage *et al.* [1986], Verhine and Melo[1988], McGinn *et al.*[1991]). Other analyses concentrate on specific school resources such as lack of writing materials and textbooks, insufficient material resources, and too little time in school (see, for example, Melo [1982], Armitage *et al.* [1986], McGinn *et al.* [1991], World Bank [Finance Primary Education-1986].) These arguments are frequently bolstered by data on aggregate expenditures. According to the World Bank (Finance of Primary Education [1982]) per pupil spending across states in Brazil ranges from US$24 to US$227 (see also Xavier and Marques [1984], Armitage *et al.* [1985].)[5]

This analysis employs data from the rural northeast of Brazil to test the various hypotheses about the determinants and effects of grade repetition. While the northeast region is extreme in its deprivation and, as such, is a reasonable starting point from a policy perspective. Further, we believe that many of the basic findings are transferable to other parts of Brazil and to other developing countries.

The paper begins with a short description of northeast Brazil, the laboratory for this analysis. Section 2 considers underlying factors affecting grade repetition including the availability of appropriate grade level instruction and the probability that an individual is retained in grade. Section 3 then turns to the learning that is accomplished through repetition. Section 4 employs the basic learning and promotion data from the student panel to investigate the potential impact of mandatory promotion policies.

[5] The concerns about the availability of resources are heightened by the arguments of Heyneman and Loxley [1982]. After comparing many educational systems, they conclude that the poorer is the country, the greater is the effect of the school in the student performance.

1. Brazil's Rural Northeast[6]

Brazil is politically divided in five regions with the northeast being the poorest. The northeast region encompasses 18 percent of the Brazilian land area and about 30 percent of the Brazilian population in 1990. But, it generated only 13 percent of the national product. Mean earnings in 1988 in the rural northeast were 28 percent of the national average. While 20 percent of the population in Brazil has less than one year of schooling, this figure jumps to 39 percent in the northeast. Moreover, in the Northeast 39.7 percent of the population over age 15 were illiterate, compared to 21 percent for all of Brazil.

Table 1 compares the Fletcher and Ribeiro[1989] estimates of repetition, drop-out and participation rates between Brazil and the Northeast. It also displays the sizable discrepancies between urban with rural areas in the Northeast. At each grade, there is more repetition in the northeast than in the rest of Brazil, with the repetition rates in rural areas approaching double those of Brazil as a whole. Dropout rates rise across grades--something that is not particularly surprising given the low overall levels of completion. And, again, the rural northeast presents a bleak picture compared to other areas.

---

[6] Statistical information in this section is drawn variously from: (i) Anuário Estatístico do Brasil - 1990, IBGE; (ii) Educação - indicadores sociais - volume 1, IBGE; (iii) Pesquisa Nacional por Amostra de Domicílios, 1982, IBGE; (iv) PROFLUXO model, developed by Fletcher and Ribeiro; and (v) World Bank publications, including specifically Brazil: Economic Survey Report: Northeast Region: Development Issues and Prospects (Report No. 6894-BR; July 20, 1987).

Table 1 - Repetition, Drop-out, and Participation[a] rates in Brazil and Northeast Brazil[b]

	BRAZIL	NORTHEAST			
		Total	Urban	Rural	Rural Low-income
Repetition rate (% enrollment)					
1st grade	.54	.65	.58	.73	.74
2nd grade	.33	.45	.42	.51	.53
3rd grade	.26	.37	.33	.48	.50
4th grade	.20	.32	.30	.44	.49
Drop-out rate (% enrollment)					
1st grade	.02	.04	.03	.05	.06
2nd grade	.04	.07	.04	.12	.14
3rd grade	.07	.09	.06	.16	.18
4th grade	.18	.16	.11	.29	.30
Participation rate (% generation)					
1st grade	.90	.79	.90	.68	.64
2nd grade	.86	.71	.85	.55	.50
3rd grade	.81	.63	.80	.42	.36
4th grade	.73	.53	.72	.29	.23

Note: **a** - Participation rate is defined as the percentage of a broad age cohort enrolled in each grade.
b - Statistics based upon the PROFLUXO model (see Fletcher and Ribeiro[1989].)

2 - The Causes of Student Repetition

This section provides separate analyses of two components of grade repetition. First, because schools with appropriate grade levels are not necessarily available in rural areas, we study the underlying causes for a school not providing advanced grades. By comparing schools which do not provide instruction past the second grade with schools providing at least fourth grade, we obtain some insights into the determinants of schooling opportunities for students. We hypothesize that school and county characteristics will be the most important factors affecting the probability of a school in providing advanced grades. Second, we analyze the underlying factors affecting individual student grade repetition by comparing students retained in the second grade for two years with other students. This allows investigation of the separate effects on student repetition patterns of student

characteristics, family socioeconomic background, teacher and school characteristics, and community factors.

These analyses are feasible using a unique data source which permits tracking schools and students over time--a key element in any analysis of student flows. The EDURURAL data set, the basis for the micro analysis in this paper, was constructed to permit evaluation of programs related to a major educational loan from the World Bank to the northeast.[7] The sampling design included primary schools in areas that received loans and aid and that did not. All of the schools were found in impoverished rural areas in the states of Ceara, Pernambuco, and Piaui. The design called for repeated follow-up of sampled schools, visiting them initially in 1981 and then returning in 1983 and 1985.[8]

The EDURURAL data set was not designed to answer the two main problems treated here, i.e., which are the causes of repetition and what are the effects of repetition. Nevertheless, the sample design of the EDURURAL evaluation provides a unique opportunity to address these problems. Because the design involved visiting each sample school twice, it was possible to construct a panel, albeit limited, of students who were sampled in successive surveys. In 1983, of the 2,619 sampled students in the second grade, 506 were sampled again in 1985. From this latter group 127 students were still in second grade, forming the panel of grade repeaters employed in this and the

[7] The EDURURAL project was a US$92 million undertaking of the Brazilian government launched in 1980. It received US$32 million in loans from the World Bank and involved a comprehensive set of resources supplied to specific schools. The analysis here is not, however, concerned with the specifics of the evaluation but instead merely relies upon the data generated to evaluate that project. More complete information about the EDURURAL data set can be found in Armitage *et al.*[1986], and Harbison and Hanushek [forthcoming].

[8] A central problem in the sampling was that no effort was made to locate individual students. In each year, a random sample of students was drawn from each school. Since our analytic design employs students who are found in succeeding samples, any student who is not randomly selected in both 1983 and 1985 will be lost. A special sample following students, instead of schools, was constructed for selected schools in Ceara in 1987, but those data are not employed here because of the small samples of grade repeaters and the special problems of sample design.

following sections.[9]

2.1 - The Provision of Advanced Grades

A prerequisite for school attendance is the existence of a school with appropriate grades of instruction within a reasonable distance. School survival from year to year is not assured, as demonstrated in Harbison and Hanushek [forthcoming]. Additionally, given that a school has survived, it is important to know if it provides grades for further progress. A student cannot progress in a school that does not provide advanced grades. The absence of advanced grades has obvious implications for repetition patterns.

The sampling scheme of the EDURURAL project does not allow investigation of the general question of what determines whether or not a school exists for any individual student, but it does allow tracing the history and analyzing the existence of fourth grades for those schools sampled.[10] To do this we use a probit model to capture how school grade structure -- as measured by whether on not the school provides second grade as the most advanced grade -- is affected by various external factors.

Table 2 summarizes the results of estimates based on the school sample from the EDURURAL data base. The explanatory variables used in the models can be divided in three categories: school characteristics, county economic conditions, and the governmental support. For expositional purposes, the results of the estimation are translated into estimates of marginal probabilities evaluated at the means of the separate variables. (Variable definitions are found in the

---

[9] The sample involved surveying second and fourth graders at two year intervals. Only the 1983-85 matched sample, however, provides sufficient numbers of students repeating the second grade. The remaining 379 matched students were promoted to the fourth grade and provide the basis for estimating achievement value-added models; see Harbison and Hanushek [forthcoming].

[10] The important distinction here is that the sample is school based and not student based. Therefore, it is not possible to describe the availability of schools for the typical student but instead only for those initially in a sampled school.

Appendix.)

Table 2 - Factors Influencing the Probabilities That a School Ends with 2nd Grade, 1983-1985

Variables	1983-1985
School Characteristics	
No. of Students	-0.0014
Hardware	-0.3157
School in Teacher's House	(0.1198)
Economic Conditions	
Percentage Selling Crops	(0.0038)
Participation in "Emergencia"	(0.0024)
Organizational/governance Factors	
OME index	(0.1210)
States	
Piaui	(-0.1857)
Ceara	(-0.0353)
State - Program	
EDURURAL - Piaui	(0.0691)
EDURURAL - Ceara	0.2381
EDURURAL - Pernambuco	(-0.0671)

Notes: 1) Estimated marginal probabilities are calculated at means of variables and holding constant other factors contained in probit equations that exclude school control measures.
2) Estimates that are not significantly different from zero at the 5 percent level are reported in parentheses.

Two school factors are systematically related to the terminal grade in the school. Schools serving a larger number of students and schools with better facilities ("Hardware") are more likely to have a fourth grade (i.e., have a lower probability of ending at the second grade).[11] Schools located in the teacher's house -- a particularly marginal type of school -- are more likely to end at the second grade, but the estimated effect is not statistically significant.[12]

Local economic conditions have small and insignificant effects on the chances of having a

[11] The Hardware index measures physical facilities in schools such as number of classrooms, existence of multi-purpose room, kitchen, secretary/principal office, and seat for students.

[12] This is the marginal effect of the school's being located in the teacher's house *after* considering the state of facilities, which typically are below average.

fourth grade. Local conditions are measured by the percent of families that sell a portion of their crops and by the percentage of families participating in the *Emergencia* program--an employment program related to the severe draughts in the northeast that limited agricultural production.

The remaining factors relate to the organization and governance of the schools. Differences in support staff were not significantly related to the school's grade structure. Specifically, beyond paying for building, teacher salaries and instructional equipment, governmental support for schooling typically involves both routine managerial control, inspection, pedagogical supervision, and technical assistance. The Orgão Municipal de Educação (OME) is the specialized county-level government agency established to systematize and institutionalize these functions of education administration. The specific measure of OMEs is an index including both quantity and quality of staff, but variations in this had little effect on the underlying probabilities of grades beyond the second grade.

Of the measures of state and program status, the only significant difference was found in the EDURURAL program counties of Ceara, where schools were much more likely to end at the second grade. These estimates, which give comparisons to nonprogram areas in Pernambuco, indicate that schools are 24 percentage points more likely to end at the second grade in Ceara areas covered by the EDURURAL program. The underlying reasons for these differences are, however, not known.

2.2 - Influences on Student Repetition

We now turn our attention to student repetition. Whether individual student performance is related to repetition probabilities is a central issue in our analysis. This is extremely important for policy purposes, because it offers insight into how to assess different proposals for dealing with dropout and retention rates and their mirror image, promotion rates. Specifically, if retention is only slightly related to actual student performance -- that is, the people being left behind are about as good academically as those being promoted or dropping out -- then high repetition rates and high dropout

rates indeed represent wasted resources. Direct, regulatory efforts to lower this wastage and increase promotions might well be called for. On the other hand, if repetition are highly related to student quality, decreasing the rates of repetition reduces wastage by continuing students with lower performance; the benefits of an external intervention program of lowering wastage would be much less.

The analysis again employs probit techniques to compare those students who repeated second grade twice with those students who followed some other path -- i.e., those promoted to the fourth grade in two years, those who dropped out of school, and those who are in the third grade. Of course, the comparison group of students is not homogeneous, and policies for drop-out students are surely different from policies for repeaters. Nevertheless, this initial analysis allows us to focus directly on the issue of repetition.[13]

Another important policy variable used in the model is the dummy variable indicating if a school provides the second grade as the most advanced grade. If a school does not have a fourth grade in 1985, then it is impossible sample a student in this grade. More importantly, the student has no place to go in that school if promotion is warranted.

The estimated repetition model, estimated by probit techniques, is found in Tables 6 and 7. Again, for expositional purposes, the results of the estimation are translated into estimates of marginal probabilities evaluated at the means of the separate variables.

Because of the random sampling of students in the schools in each year, it is possible for an individual to be retained but not to be included in the sample. To deal directly with this, the probit model includes the number of students in the schools, since the probabilities of being missed by the sampling are directly related to the number of students in the school. The school size measure (not

---

[13] This analysis complements the analysis of on-time promotion found in Harbison and Hanushek [forthcoming]. That analysis contrasts sampled students who were promoted from second to fourth grade in the two years between 1983 and 1985 with all other students sampled in the second grade in 1983.

shown) is significantly negative in the probit model, reflecting this sampling within schools.

Student and Family Characteristics. Student backgrounds should directly affect repetition probabilities. Students in families with, for example, better educated parents are supposed to be less likely to repeat a school year when compared with those whose parents have less education or are illiterate. Students with higher previous achievement are expect to be promoted than those with lower previous achievement.

The table 3 summarizes the marginal probability associated with student and family characteristics used in the model.

Table 3 - Effect of Student and Family Characteristics on the Repetition Probabilities, 1983-1985.

Characteristics	1983-1985
Female student	(-0.0042)
Student's age	(-0.0015)
Portuguese test score	-0.0010
Mathematics test score	-0.0005
Father's education	(0.0002)
Mother's education	(-0.0035)

Notes: 1) Estimated marginal probabilities are calculated at means of variables and holding constant other factors contained in probit equations that exclude school control measures.
2) Estimates that are not significantly different from zero at the 5 percent level are reported in parentheses.

The most interesting part of the model is the relationship between second grade test scores and repetition probabilities.[14] As displayed in table 3, lower test scores consistently lead to greater

[14] The Portuguese and mathematics tests employed here were developed specifically for the EDURURAL project by a team of psychometricians from the Fundação Carlos Chagas. The tests, developed in 1981 and improved in later years, were criterion referenced to minimally acceptable levels of performance in second and fourth grade mathematics and Portuguese. The test reliability, ascertained by constructing Cronbach's Alpha coefficients, shows reliability coefficients of 0.9 or better with the exception of the fourth grade Portuguese scores. Moreover, the test reliability tends to be stable over time and across states. For more information on the tests, see Harbison and Hanushek [forthcoming].

repetition probabilities; this suggests that promotion has some basis in merit. Each 10 points on Portuguese test, which has a standard deviation of approximately 25 points, decreases the repetition probabilities by about 1 percent. The effect of the mathematics test is half of this. Since the mean observed repetition rate in the sample is only 4 percent in 1983, these are significant differences due to merit. These results also confirm the finds in Harbison and Hanushek [forthcoming], where achievement on the second grade test was found to be positively related with the student on-time promotion probabilities.

Girls and boys have the same probability in repeating a school year. Students age has no effect on their repetition probabilities. These are both surprising, because these two variables were found to affect the student's on-time promotion probability (see Harbison and Hanushek [forthcoming]). Mother's and father's education are also not significantly related with the repetition probabilities.[15]

Grade provided by schools. As described previously, the availability of a school with advanced grades is not assured. Our specific concern is whether a portion of repetition is related simply to lack of other schooling opportunities. Simply stated, a student, who was in a second grade in 1983, cannot be promoted in schools where second grade is the most advanced grade provided. Our probit model include a dummy variable which equals 1 if the highest grade provided by the school is the second grade and equals 0 if the school provides grades for further progress. Not surprisingly, students in a school with second grade as its highest grade are significantly more likely to be retained in the second grade. In fact, student placed in such schools have their repetition probabilities increased by 2.3 percentage points, which is a huge compared with the mean observed repetition rate

[15] Verhine and Melo[1988] emphasize the importance of socio-economic factors, something that does not seem too consistent with these estimates. The difference from our results may partially be explained by the restricted sample used here; all students are from poor, rural families.

in the sample of 4 percent.[16]

Economic Conditions and Governmental Support. As summarized in the table 4, students are more likely to repeat a year in richer counties, i.e., those with a higher socioeconomic index. We do not have a clear explanation for this except that the opportunity cost of attending school in wealthier counties is higher and thus students are more likely to be absent. Unfortunately, we lack direct information on absenteeism. (The alternative view is that wealthier counties can afford better to make investments in schooling--an hypothesis predicting the opposite sign of that observed.)

Table 4 - Effects of County Economic Conditions and Governmental Supports on Repetition Probabilities, 1983-1985

Characteristics	1983-1985
Socioeconomic index	0.0822
OME index	-0.0418
States	
Piaui	0.0274
Ceara	0.0474

Notes: 1) Estimated marginal probabilities are calculated at means of variables and holding constant other factors contained in probit equations that exclude school control measures.
2) Estimates are significantly different from zero at the 5 percent level.

The primary organizational measure reflects the quality and quantity of personnel in the OMEs. Students in counties with better OMEs are less likely to be retained in the second grade.

There are also distinct differences in repetition probabilities across states, as shown in table 4. The repetition probabilities in Ceara are clearly the highest among the three states. A student in

[16] McGinn *el at.* [1991] pinpoint another characteristic school organization--the use of multigrade instruction--as an important element of repetition. Direct analysis of this in our sample, however, did not support any different repetition patterns with the use of multigraded classes.

Ceara has a 4.7 percentage point higher chance of repeating the second grade twice than a student in Pernambuco (the comparison state for this analysis). Piaui also has a 2.7 percentage point higher repetition rate than Pernambuco. Again, we cannot offer any specific explanations for these differences which hold over and above any of the other factors in the model.

2.3 - Summary of Repetition Factors

Grade repetition has, according to our statistical analyses, two major components. First, government provision of suitable schools with grades for student advancement, is a prime factor. Other things being equal, the presence of grades beyond the second grade is an extremely strong determinant of student advancement. This suggests that government intervention to insure appropriate schools can have a powerful effect on repetition and wastage. Firmly established schools with adequate facilities, things that the government can influence directly, are required. Second, student achievement--as measured by tests of mathematics and Portuguese performance--is a key determinant of repetition. While some have suggested that repetition is based on factors other than student performance such as local politics, the evidence points directly to the role of student performance.

3 - The Achievement Effects of Repetition

Discussions of repetition tend to neglect one important aspect of the issue: Students who repeat a grade are in fact attending more school and would be expected to learn something during the experience. While this may be a very expensive way of organizing the learning process (the subject of attention below), it is inappropriate to assume that this is pure waste.

A simple look at the EDURURAL suggests that repetition does have noticeable learning effects. As shown in table 5, the means of the second grade repeaters in 1983 were 40.19 and 35.74 in Portuguese and mathematics, respectively. These means were more than half standard deviation

below the means of the entire second grade sample in 1983. In 1985, however, the means of achievement in Portuguese and mathematics of the repeaters were slightly above the means of the entire second grade sample.[17]

Table 5 - Means and standard deviation for 2nd graders and 2nd grader repeaters, 1983 and 1985

	All 2nd Graders		Students Remaining in Second Grade, 1983 and 1985	
	1983	1985	1983	1985
Portuguese				
Mean	58.7	59.6	40.2	61.1
Standard Deviation	23.6	25.2	25.1	22.7
N (sample size)	3944	4321	127	127
Mathematics				
Mean	51.2	49.2	35.7	52.4
Standard Deviation	24.9	25.0	25.3	25.1
N (sample size)	3944	4321	127	127

Source: EDURURAL data set.

This analysis pursues two, parallel lines of inquiry. First, we refine the estimates of the achievement gains from repetition just presented. Second, we explore whether differences among individual students in the achievement value of repeating grades can be explained in terms of student or school factors. The overall framework for analysis follows a quite standard input-output specification for the educational process, but one modified to incorporate information about grade repetition.[18]

The achievement of a given student at time $\underline{t}$ (A^t) is assumed to be related to current and

[17] Note that students identified as repeaters are all children who were sampled in the second grade both in 1983 and again in 1985.

[18] Armitage *et al.* [1986] and Harbison and Hanushek [forthcoming] provide more complete information about the achievement tests used here as well as about the specific statistical models. The general framework is described in detail in Hanushek[1978,1986].

past educational inputs from a variety of sources -- the home, the school, and the community. To highlight some of the important features, we use a general conceptual model such as:

$$A^t = f(F^{(t)}, S^{(t)}, O^{(t)}, \epsilon_t),$$

where $F^{(t)}$ = a vector of the student's family background and family educational inputs cumulative to time t;

$S^{(t)}$ = a vector of the student's teacher and school inputs cumulative to time t;

$O^{(t)}$ = a vector of other relevant inputs such as community factors, friends, and so forth cumulative to time t; and

ϵ_t = unmeasured factors that contribute to achievement at time t.

The approach is to measure the different possible inputs into education and to estimate their influence on student achievement. This conceptual model explicitly incorporates a stochastic, or random, error term -- ϵ_t -- to reflect the fact that we can never observe all of the factors affecting achievement. The estimation problem is simplified considerably if there is information on achievement at two different times, for example, at time t and at an earlier time t'. It is possible then to include the prior achievement as one of the explanatory variables in the regression and to concentrate on the specific inputs over just the period t to t'. This formulation, which is often called a "value-added" specification, gets around the lack of measurement of past inputs into the process and of other individual specific (but constant) factors such as ability.

3.1 - Learning through Repetition

The simple differences in means for repeating students compared to all students (table 5) can potentially misstate the learning effects associated with grade repetition. When repeating students have special characteristics or school circumstances that differentiate them from other students, the

difference in means will misstate the separate effect of repetition.[19]

We employ a cross-sectional analysis of achievement differences to estimate the effect of repetition on student learning. Specifically, standard models that include student, family, and school factors are supplemented with information about repetition. A dummy variable which assumes 1 if the student is repeating a school year and 0 otherwise is included to capture the independent learning effects of repetition. We estimate this model for second and fourth grades in 1983 and 1985, using the two achievement tests (Portuguese and mathematics) as dependent variables.

There are some obvious problems with this approach, and thus it should be viewed as a crude approximation of the effects of repetition. Three problems arise. First, repetition is not exogenous but itself is affected by performance. This implies that causation runs in both directions and that the estimates of the pure learning effect of repetition are biased. Second, the repetition measure does not indicate how many years had been repeated. Instead it only indicates whether or not the student was in the same grade the previous year. Therefore, it averages together varying amounts of repetition. Third, because of the structure of EDURURAL data set, it is not possible to estimate the effects of repetition within a value-added context; such estimation can only be done in cross-sectional models. This heightens the chance that the estimates of the effects of repetition will be contaminated by other factors that are mismeasured.

[19] Because we do compare the same students over time, however, the difference in means comes close to a value-added measure of repetition. Thus, some of the largest sources of bias are eliminated.

Table 6 - Effect of repeating a school year on achievements (t-statistics are in parentheses)

Grade	Portuguese		Mathematics	
	1983	1985	1983	1985
Second	0.575 (0.63)	2.632 (2.81)	1.614 (1.64)	4.149 (4.41)
Fourth	4.200 (3.36)	5.562 (4.85)	4.083 (2.57)	4.385 (2.92)

The table 6 summarizes the effects of grade repetition on school achievement. It is not surprising that repetition is significant in most of the cross-section models employed here. Only for the second grade specification in 1983 for Portuguese and mathematics achievement is this not significant at the 5 percent level. According to the estimates, by repeating the second grade students can raise their achievements by 2.6 points in Portuguese and 4.1 points in mathematics. In the fourth grade estimates the effect in the means achievements ranges from 4.2 to 5.6 points in Portuguese, and from 4.1 to 4.4 points in mathematics.

Note that these are the net relationship between achievement and repetition. If students who repeat begin at a lower level of achievement than those who do not repeat a grade, the period of repetition is more than sufficient to make up for the average starting decrement. After repeating, the students have higher tests than those not repeating, holding constant family background and other factors.

This effect, however, has some costs. The student must spent at least one more year in the same grade at school. Beyond the increase in opportunity cost, the direct cost are not negligible, even in this area where the student cost is low. Assuming that all repetition only lasts one year, the average of direct costs in raising 1 point in Portuguese (mathematics) through repetition in the second grade is US$11.40 (US$7.23) and is US$6.67 (US$7.14) for repeating the fourth grade. These are

huge figures since the average student cost in the rural Northeast is only US$30.00.[20]

3.2 - Differential Learning While Repeating Grades

From the previous analysis we can conclude that students learn by repeating. We cannot conclude anything about which factors may be most important for learning during the period of repetition. Here we consider directly whether there are systematic learning differences among the grade repeaters by estimating value-added achievement models for repeaters. We use this specification in the special matched sampled 1983/1985, where we could find 127 second grade repeaters.[21]

The results from these regressions (table 7) give us little guidance about what can improve repeaters' achievement. Most of the variables used in the model are not statistically significant at 5 percent level. The main result is that students' previous achievement is consistently related with their achievement after repetition. This, of course, it was not surprise. In short, we do not have a good explanation for what makes a difference in repeaters achievements.

Beyond previous achievement, only student's age appear to be consistently affecting repeaters achievements. The effect is negative, i.e., the older students do worse than the younger ones. Mothers' education, which was consistently significant in the general achievement model (see Harbison and Hanushek [forthcoming],) does not appear to have any influence at all for repeaters' performance. The same holds with fathers' education. These results are summarized in table 7.

---

[20] See Harbison and Hanushek [forthcoming], and Xavier and Marques [1984].

[21] As described in Harbison and Hanushek [forthcoming], re-visiting sampled schools turned up a number of students sampled in two successive surveys. The models here employ the restricted sample of students in the second grade both years to understand the effects of family, community, and school factors on achievement growth. The sample, unfortunately, is quite small, making the detection of differential effects difficult.

Table 7 - Effect of Student and Family Characteristics on Repeater's Achievement, 1983/1985 (t-statistics in parentheses)

	Portuguese		Mathematics	
Variables	(1)	(2)	(1)	(2)
Personal characteristics				
Female student	4.660 (1.23)	4.026 (1.11)	-9.491 (-2.28)	-10.009 (-2.54)
Age	-1.690 (-2.28)	-1.589 (-2.18)	-1.619 (-1.98)	-1.518 (-1.92)
Parents' Education				
Mother's Education	1.024 (1.05)	0.603 (0.64)	0.044 (0.04)	-0.167 (-0.16)
Father's Education	-0.521 (-0.50)	-0.596 (-0.58)	1.218 (1.06)	1.164 (1.05)
Joint characteristics: pupil and school				
Portuguese test score, 1983	0.450 (4.18)	0.442 (4.15)	0.218 (1.84)	0.211 (1.83)
Mathematics test score, 1983	-0.002 (-0.02)	-0.001 (-0.01)	0.235 (2.29)	0.243 (2.44)

Note: Specifications (1) and (2) use different teacher characteristics.

In general, students in Ceara learned more over the period than students in Piaui and Pernambuco, the other sampled states (see Harbison and Hanushek [forthcoming].) This does not prove true for repeaters. Repeaters in all three states perform evenly. Despite huge repetition rates, none of the states appears to have any special program to them or, if they have, such programs do not appear clearly beneficial.

3.3 - Summary of Learning Effects of Repetition

The central finding from the examination of achievement is that repetition does enhance a student's learning. While students who repeat are on average below average in performance before repetition, they move to above average after repetition. Therefore, repeating a grade is not pure waste as some would suggest. On the other hand, it is a very expensive form of schooling. Among repeating students, there is, however, no information on what specific factors determine differential achievement.

4 - Mandatory Promotion

One way that is sometimes suggested for tackling the wastage resulting from high repetition is mandatory promotion. Indeed, if promotion and its mirror image, repetition, in the system is not highly related with the student school performance, then a mandatory promotion policy could diminish the wastage with perhaps low cost to the educational system. This, however, is not the case that we found in our data; promotion[22] and repetition were strongly related with the student achievement. If such a direct linkage is the case, we would expect mandatory promotion to lower the effective level of achievement associated with each grade, thus damaging overall school quality.

Nevertheless, the issue cannot be decided on *a priori* grounds. The incredibly high repetition and drop out rates existing in the Brazilian school system, especially in the primary school, increase the cost of getting a graduate, because money is spent on people who never or very slowly progress through the system. Therefore, it is worthwhile exploring more deeply this problem and trying to infer what would happen if we promoted students who fail under the current system. At the very least, this allows more accurate description of the exact nature of the trade-offs.

A central question is how student achievement is affected by repetition and, inferentially, by mandatory promotion. Our previous analyses gave some indication of the average effects of grade repetition. Here we pursue another logical approach of investigating in more detail the entire distribution of promotees and repeaters.

A total of 3,944 students were sampled in the second grade in 1983. 506 of them were sampled again in 1985; 127 were still in second grade while the other 379 were found in the fourth grade. Table 5 provided the means and standard deviations of the Portuguese and mathematics achievement scores in the second grade for students repeating the second grade. In contrast, the students promoted on time to the fourth grade had average 1983 second grade scores of 68.6 and 56.8

[22] See Harbison and Hanushek [forthcoming].

for Portuguese and mathematics, respectively. Thus, they were the .2-.4 standard deviations above mean instead of .6-.8 standard deviations below the mean that the repeaters were. By 1985, however, the means for the repeater group are slightly above the means of all students in second grade. But, as noted above, it took them two additional years to catch up with the grade average. Moreover, while close, they are still behind the group that is promoted after the two years.

We can also go beyond the means and look at the distribution of performance. Figures 1 and 2 give us an idea of the distribution of the achievement of the two groups. The distributions were calculated using z-scores (standard deviations from the mean), based on the means and standard deviations for all second grade students in 1983.

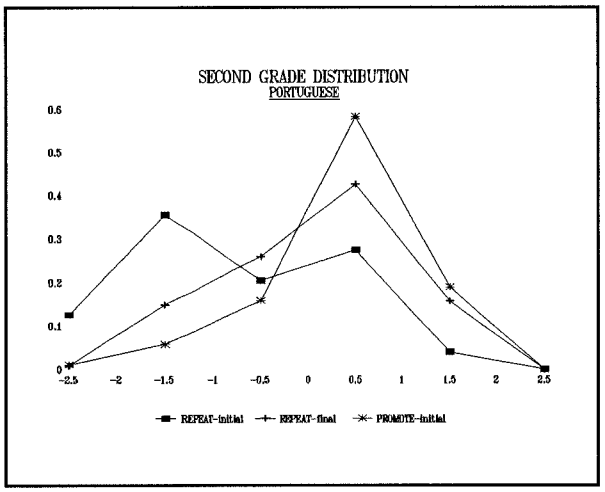

Figure 1

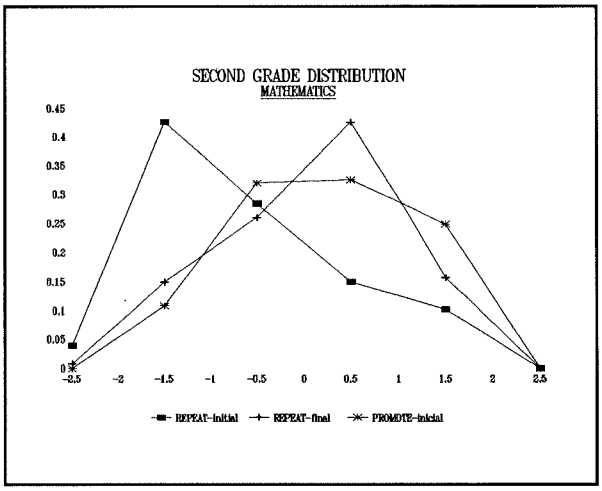

Figure 2

These figures show clearly how grade repetition shifts the distribution of student performance. But, importantly, they also show that the distributions of performance for repeaters and those promoted overlap to a significant extent. This suggests that one crude analytical approach would be to project fourth grade achievement on the basis of where each child falls in the distribution of those promoted. (For those promoted the distribution of fourth grade scores is known.) Such projections

clearly make very strong assumptions. Significantly, they assume that the previous achievement is the only thing that influences promotion and subsequent fourth grade student achievement. Such assumptions are almost certainly false, but this approach gives us some notion of an upper bound on achievement under a mandatory promotion policy.

We estimate the achievement or, at least, a range where achievement in the fourth grade will lie, if each student currently repeating the second grade were promoted. We begin by splitting the initial and final distribution into six subgroups: Z-score $<=$ -2; -2 $<$ Z-score $<=$ -1; -1 $<$ Z-score $<=$ 0; 0 $<$ Z-score $<=$ 1; 1 $<$ Z-score $<=$ 2; and Z-score $>$ 2. We then calculate transition probabilities based on the experiences of the promoted students. Finally, we apply these transition probabilities to the distribution of second grade scores for the repeaters. In this latter estimation we actually employ both the pre- and post-repeating score for the students. In other words, the use of the pre-repeating scores relate to a pure "mandatory promotion" policy.[23] The post-repeating scores relate to a modified plan of a fixed number of years in each grade.

Table 8 displays the transition probability matrices used for Portuguese and mathematics performance. These come directly from the matched sample of on-time promoted students.

[23] This must actually be qualified, since we are not sure that matched repeaters have just entered the second grade as opposed to already having been in the second grade for some period.

Table 8 - Transition Probabilities: Portuguese and Mathematics Achievement, 1983/1985

Initial Achievement (second grade) z-score	Follow-up Achievement (fourth grade) z-score					
	Less or equal to -2	Between -2 and -1	Between -1 and 0	Between 0 and 1	Between 1 and 2	Greater than 2
Portuguese						
Less or equal to -2						
Between -2 and -1	0.25	0.50	0.25	0.00	0.00	0.00
Between -1 and 0	0.23	0.32	0.23	0.18	0.05	0.00
Between 0 and 1	0.05	0.48	0.23	0.17	0.07	0.00
Between 1 and 2	0.01	0.12	0.38	0.35	0.13	0.01
Greater than 2	0.00	0.01	0.18	0.57	0.21	0.03
Mathematics						
Less or equal to -2						
Between -2 and -1	0.06	0.44	0.42	0.04	0.04	0.00
Between -1 and 0	0.03	0.32	0.38	0.16	0.11	0.00
Between 0 and 1	0.00	0.09	0.39	0.37	0.15	0.01
Between 1 and 2	0.00	0.05	0.14	0.45	0.36	0.00
Greater than 2						

Figures 3 and 4 display the results of this estimation. These estimated distributions show two major things. First, the "current promotion" group--those promoted normally by the standards of the schools--do better than the repeaters. This is not particularly surprising. On the other hand, the mandatory promotion distribution, derived from inferring the fourth grade performance of those repeating based on their initial second grade score distribution, looks reasonably close to that obtained for delayed promotion (i.e., after repeating for two years). Since the delayed promotion is very costly--the full cost of two years of schooling, mandatory promotion may be an effective alternative to the current system. This is, it must be emphasized, just a second best policy. The first best policy is to improve the quality of primary schools so that student achievement is increased directly.

One group of repeating students--those performing well on both the Portuguese and

mathematics tests--are of special interest. In our sample, 14 percent of the repeating students were above the mean performance on both tests when they initially took the tests. When we investigated their circumstances, however, we found that 13 of the 18 students were in schools that did not offer instruction past the second grade. This again underscores the room for alternative, quality improving policies.[24]

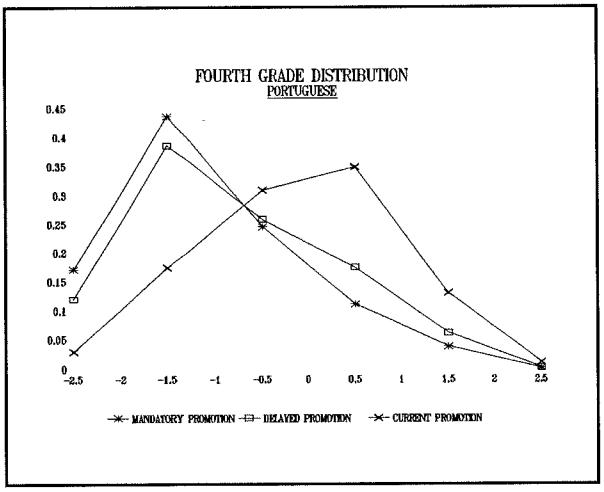

Figure 3

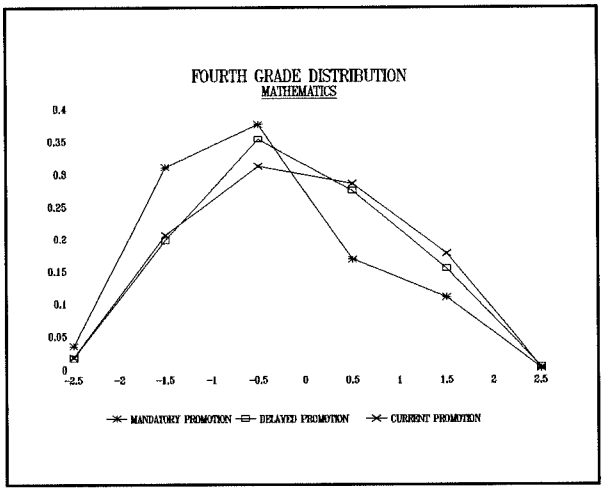

Figure 4

All of these findings must, of course, be highly qualified. It is quite likely that promotion involves other factors, observed by the teachers but not measured by the tests, which affect the learning of students. Therefore, inferring that the repeaters could acquire the third and fourth grade material at the same rate as those promoted on time is undoubtedly an overstatement.

[24]The previous analysis of predicted fourth grade performance (figures 3 and 4) is not substantially affected by limitations on grades offered. A total of 42 students in the sample of repeating students were found in schools ending at the second grade, and these students were distributed across the performance distribution. Therefore, when we duplicated the mandatory promotion analysis with the grade limited students eliminated, we obtained the same qualitative results.

5 - Conclusions

It is impossible to ignore the problems of grade repetition in developing countries. The consistent pattern of students' being stuck in primary grades with the concomitant demands on scarce educational resources commands the attention of policy makers in most developing countries. Yet, for its importance, there is extremely little known about either the causes or effects of repetition.

This paper provides a systematic investigation of grade repetition in rural northeast Brazil. Employing a unique data set that allows observation of the same students over time, it is possible to estimate the determinants of repetition. Further, the educational effects of repetition are open to analysis.

The results are straightforward. Two factors are most important in determining repetition. First, student achievement levels are very important. Low performance, and not other less educationally relevant factors, is a key element. Second, governmental policy as evidenced by supplying advanced grade levels in the schools is central. Simply put, if there is no place to go, students will stay where they are, repeating primary grades.

Repetition also has a direct impact on achievement. Repeating the second grade over a two year period moves students from between half and one standard deviation below the mean to a position close to the mean in achievement. But, this is an expensive policy, and there are quite likely to be alternative and less costly ways to improve achievement.

Mandatory promotion policies would produce lower achievement in later grades (because there is learning that goes on through repetition). On the other hand, while mandatory promotion appears undesirable to a policy of improving school quality, it does seem superior to the current unguided repetition policies.

These results are, nonetheless, based on rather small and less than perfect samples. The dearth of information about the entire process of promotion, repetition, and dropping out behavior implies that informed decision making is extremely difficult.

REFERENCES

Armitage, J., Gomes-Neto, J.B., Harbison, R.W., Holsinger, D.B., and Leite, R.H., School Quality and Achievement in Rural Brazil. World Bank Education and Training Series, No. EDT25. Washington, D.C., 1986.

Brandão, Z. et al, Evasão e repetência no Brasil: a escola em questão. Rio de Janeiro: Achiame, 1983.

Carvalho, I.M.M., Escolarização em famílias de classe trabalhadora. Cadernos do CEAS. Salvador, 83:44-59, jan/fev, 1983.

Coleman, J.S., Hoffer, T., and Kilgore, S., High school Achievement: Public, Catholic, and Private Schools Compared. New York, N.Y.: Basic Books, 1982.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfield, F.D., and York, R.L., Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Cunha, L.A., Educação e desenvolvimento social no Brasil. Rio de Janeiro, Francisco Alves, 1981.

Fletcher, P.R. and Castro, C.M., "Os mitos, as estratégias e as prioridades para o ensino de 1° grau," Educação e Realidade, Porto Alegre, 11(1):35-42, jan/jun, 1986.

Fletcher, P.R. and Ribeiro, S.C., "A Educação na estatística nacional," in:PNADs em Foco: Anos 80, D.O. Sawyer, org., pp 11-33, 1988.

Fletcher, P.R. and Ribeiro, S.C., "Modelling Education System Performance with Demographic Data: An Introduction to the PROFLUXO Model," Mimeo: Brasília, March 8, 1989.

Fuller, B., Raising School Quality in Developing Countries: What Investments Boost Learning?. World Bank Education and Training Series, No. EDT7, Washington, D.C., 1985.

Garcia, R.L., "A qualidade comprometida e o comportamento da qualidade," ANDE, São Paulo, 1(3):51-55, 1982.

Gatti, B.A., "A Reprovação na 1ª série do 1° grau: um estudo de caso," Cadernos de Pesquisa, São Paulo, 38:3-13, ago, 1981.

Hanushek, E.A., "The Economics of Schooling: Production and Efficiency in Public Schools," Journal of Economic Literature 24(3):1141-77, 1986.

Harbison, R.W. and Hanushek, E.A., Educational Performance of the Poor: Lessons from Rural Northeast Brazil, *forthcoming*.

Heckman, J.S., "Sample Selection Bias as a Specification Error," Econometrica 47: 153-61, 1979.

Jamison, D.T., "Radio Education and Student Failure in Nicaragua: A Further Note," in Jamesine Friend, Barbara Searle, and Patrick Suppes (Ed.), Radio Mathematics in Nicaragua. Stanford, Ca.: Institute for Mathematical Studies in the Social Sciences, Stanford University. p. 225-36, 1980.

Kafuri, R. et al., Pesquisa sobre evasão, repetência e fatores condicionantes. Goiânia: Universidade Federal de Goiás, 1985.

Lockheed, M.E. and Hanushek, E.A., "Improving Educational Efficiency in Developing Countries: What Do We Know?," Compare 18(1):21-38, 1988.

Lockheed, M.E. and Hanushek, E.A., "Concepts of Educational Efficiency and Effectiveness," International Encyclopedia of Education forthcoming 1990.

Maddala, G.S., Limited-Dependent and Qualitative Variables in Econometrics. Cambridge, Mass.: Cambridge University Press, 1983.

McGinn, N.F., Soto, M.C., Lopez, S., Loera, A., Cassidy, T., Schielfelbein, E., and Reimers, F. - "Asistir y aprender o repetir y desertar, Sintesis del Informe," BRIDGES, 1991.

Mello, G.N., Magistério de 1º grau: da competência técnica ao compromisso político. São Paulo: Cortez, 1982.

Oliveira, R.D., "Os Movimento sociais reinventam a educação," Educação e Sociedade. São Paulo, 8: 49-60, jan, 1981.

Popovic, A.M., "Enfrentando o fracasso escolar," ANDE, São Paulo, 1(2):17-22, 1981.

Schiefelbein, E., Repetition: the key issue in Latin American primary education. Washington DC: World Bank, LAC Technical Department, Human Resources Division, 1989.

Verhine, R.E. and Melo, A.M.P., "Causes of school failure: the case of the state of Bahia in Brazil," Prospects, Vol. XVIII, No. 4, 557-568, 1988.

Xavier, A.C. and Marques, A.E., Custo direto de funcionamento das escolas públicas de primeiro grau na região Centro-Oeste. Brasília, 1984.

World Bank, BRAZIL : finance of Primary Education (A World Bank country study). The World Bank, Washington,D.C., 1986.

ANNEX

In this annex contains the complete results of the statistics models used in this paper followed by the definitions of the all variables used in them.

Table A1 - Probit Model on Schools providing Second Grade as the highest grade, 1983-1985

	Coefficient	t-ratio
County Characteristics		
Percent selling crops	0.0123	1.55
Participation in Emergencia	0.0078	1.53
School Characteristics		
Number os students	-0.0044	-2.30
Hardware index	-1.0108	-2.80
Teacher's house	0.3835	1.87
OME index	0.3874	1.10
State		
Piaui	-0.5948	-1.36
Ceara	-0.1130	-0.29
Program state		
EDURURAL: Piaui	0.2212	0.63
EDURURAL: Ceara	0.7624	2.40
EDURURAL: Pernambuco	-0.2149	-0.78
Constant	-0.8279	-2.51
Sample size	489	
Mean probability	0.241	
Log likelihood	-215.59	

Table A2 - Probit Model of Student Repetition, 1983-1985

	Coefficient	t-ratio
Student Characteristics		
Female student	-0.0491	-0.53
Student's age	-0.0174	-0.93
Portuguese test - 1983	-0.0119	-5.06
Mathematics test - 1983	-0.0056	-2.32
Parent's Education		
Father's education (years)	-0.0416	-1.43
Mother's education (years)	0.0020	0.09
School Characteristics		
Number of students	-0.0043	-3.44
School not providing Advanced Grades	0.2741	2.41
OME index	-0.4930	-2.24
Socioeconomic index	0.9694	4.52
State		
Piaui	0.3230	2.44
Ceara	0.5591	4.17
Constant	-0.6857	-2.29
Sample size	3240	
Mean probability	0.039	
Log likelihood	-465.98	

Table A3 - Value-added Regressions - Portuguese and mathematics - Repeaters 1983/1985

Variables	Portuguese		Mathematics	
	(1)	(2)	(1)	(2)
State				
Ceara	1.156 (0.14)	7.038 (0.89)	9.448 (1.00)	13.265 (1.54)
Piaui	-11.886 (-1.36)	-11.745 (-1.40)	-11.474 (-1.19)	-11.248 (-1.23)
Program states				
EDURURAL: Pernambuco	-2.791 (-0.40)	-0.445 (-0.07)	3.855 (0.51)	5.457 (0.73)
EDURURAL: Ceara	3.828 (0.56)	4.997 (0.75)	4.404 (0.58)	5.364 (0.74)
EDURURAL: Piaui	9.295 (1.13)	12.373 (1.68)	11.805 (1.30)	14.598 (1.82)
Personal characteristics				
Female student	4.660 (1.23)	4.026 (1.11)	-9.491 (-2.28)	-10.009 (-2.54)
Student's age	-1.690 (-2.28)	-1.589 (-2.18)	-1.619 (-1.98)	-1.518 (-1.92)
Parents' Education				
Mother's Education	1.024 (1.05)	0.603 (0.64)	0.044 (0.04)	-0.167 (-0.16)
Father's Education	-0.521 (-0.50)	-0.596 (-0.58)	1.218 (1.06)	1.164 (1.05)
Joint characteristics: pupil and school				
Portuguese test score, 1983	0.450 (4.18)	0.442 (4.15)	0.218 (1.84)	0.211 (1.83)
Mathematics test score, 1983	-0.002 (-0.02)	-0.001 (-0.01)	0.235 (2.29)	0.243 (2.44)
School characteristics				
Graded class	-7.216 (-1.59)	-5.761 (-1.40)	-3.576 (-0.71)	-2.271 (-0.51)
Pupil-teacher ratio	-0.047 (-0.23)	-0.119 (-0.61)	-0.210 (-0.94)	-0.258 (-1.22)
School hardware index	-3.901 (-0.43)	-3.868 (-0.43)	-0.655 (-0.07)	-1.133 (-0.12)
School software index	12.297 (1.30)	15.485 (1.69)	25.004 (2.41)	28.124 (2.82)
Teacher characteristics				
Years teacher's education	-0.017 (-0.23)		0.024 (0.03)	
Years teacher's experience	-0.278 (-0.86)		-0.170 (-0.48)	
Logos II -- teacher training	-1.243 (-0.24)		2.238 (0.40)	
Qualificação -- teacher training	4.867 (1.04)		4.979 (0.97)	
Teacher's Portuguese test score	-0.251 (-1.50)		-0.216 (-1.17)	
Teacher's mathematics test score	0.154 (1.26)		0.128 (0.95)	
Teacher's salary		0.019 (0.51)		0.013 (0.31)
Constant	65.581 (3.82)	53.842 (4.09)	47.648 (2.52)	39.297 (2.75)
Adjusted R^2	0.422	0.419	0.419	0.432
N (number of cases)	113	113	113	113
Statistic F	4.897	6.048	4.853	6.322

Table A4 - Cross-Section Regressions - Second Grade - Portuguese and Mathematics: 1983 and 1985

	Portuguese		Mathematics	
	1983	1985	1983	1985
Student is repeating a grade	0.552 (0.60)	2.632 (2.81)	1.574 (1.61)	4.149 (4.41)
Student Characteristics				
Female student	-0.62 (-0.02)	9.000 (2.85)	-9.816 (-3.28)	-4.489 (-1.41)
Student's age	0.375 (2.37)	0.600 (3.53)	0.878 (5.17)	1.024 (5.99)
Pupil works	-1.155 (-1.04)	-0.641 (-0.38)	0.311 (0.26)	1.810 (1.08)
Family Characteristics				
Mother's education	0.599 (3.35)	0.680 (3.45)	0.533 (2.78)	0.496 (2.51)
Father's education	0.681 (3.16)	0.255 (1.14)	0.931 (4.03)	0.649 (2.87)
Family size	-0.272 (-2.06)	-0.396 (-2.88)	0.030 (0.21)	-0.120 (-0.87)
Peer Influence				
Percent families not farming	3.736 (1.67)	10.269 (3.92)	-0.819 (-0.34)	1.214 (0.46)
Relatively large landholders	0.061 (3.22)	0.064 (3.00)	0.042 (2.06)	0.103 (4.82)
Percent female classmates	-4.466 (-1.46)	4.654 (1.41)	-0.551 (-0.17)	3.174 (0.96)
Female classmates when female student	5.586 (1.44)	2.116 (0.50)	1.028 (0.25)	1.800 (0.42)
Join Characteristics: pupil and school				
Homework	3.427 (4.82)	3.430 (4.67)	2.590 (3.40)	2.071 (2.80)
School lunch some day	-4.594 (-3.05)	-11.267 (-3.05)	-4.795 (-2.97)	-10.005 (-2.69)
School lunch every day	-4.958 (-3.02)	-7.945 (-2.13)	-5.863 (-3.34)	-5.565 (-1.48)
Male teacher/male student	0.503 (0.26)	8.259 (3.44)	2.285 (1.08)	5.805 (2.40)
Female teacher/female student	1.553 (0.87)	-4.122 (-2.02)	1.384 (0.73)	-1.364 (-0.67)
Percent seek 9 or more years of school		7.907 (5.04)		7.841 (4.97)
School Characteristics				
Graded class	-4.174 (-4.72)	0.332 (0.39)	-2.405 (-2.54)	-1.205 (-1.41)
Pupil-teacher ratio	-0.064 (-2.19)	0.074 (2.10)	-0.038 (-1.22)	-0.008 (-0.23)
School hardware index	9.201 (5.45)	-2.243 (-1.22)	6.740 (3.72)	0.825 (0.45)
School software index	5.645 (3.12)	9.770 (4.65)	3.299 (1.70)	6.942 (3.28)
Teacher Characteristics				
Teacher's education	0.793 (5.54)	0.029 (0.20)	1.228 (8.00)	0.546 (3.68)
Teacher's experience	-0.006 (-0.10)	0.000 (0.00)	0.102 (1.62)	0.052 (0.87)
LOGOS II -- teacher training	3.365 (3.11)	2.021 (1.98)	2.225 (1.92)	2.111 (2.06)
Qualificação -- teacher training	-0.426 (-0.41)	0.607 (0.60)	-3.912 (-3.49)	1.494 (1.47)
Teacher activity index	5.848 (2.94)	-4.475 (-2.06)	5.315 (2.50)	0.866 (0.40)
Teacher material index	1.888 (1.09)	0.929 (0.54)	2.461 (1.33)	-0.229 (-0.13)
Teacher's Portuguese test score		-0.089 (-2.90)		-0.159 (-5.17)
Teacher's mathematics test score		0.138 (5.88)		0.123 (5.18)
State				
Piaui	11.230 (6.10)	0.298 (0.14)	-3.751 (-1.90)	-12.607 (-5.84)
Ceara	13.923 (8.87)	14.409 (7.78)	7.082 (4.21)	6.258 (3.36)
State Program				
EDURURAL: Piaui	-1.973 (-1.36)	0.700 (0.42)	5.284 (3.39)	11.137 (6.71)
EDURURAL: Ceara	11.102 (7.74)	0.316 (0.23)	11.099 (7.22)	-0.697 (-0.51)
EDURURAL: Pernambuco	1.941 (1.34)	-3.866 (-2.20)	-0.981 (-0.63)	-9.139 (-5.16)
OME index	-1.323 (-0.69)	-2.306 (-1.12)	-6.620 (-3.23)	-8.987 (-4.32)
School Control				
State operated	2.813 (2.20)	-0.216 (-0.15)	2.268 (1.65)	-1.919 (-1.35)
Federally operated	3.614 (0.51)	13.360 (2.51)	-3.068 (-0.40)	-3.657 (-0.68)
Privately operated	8.434 (2.33)	5.904 (1.27)	5.792 (1.49)	3.459 (0.74)
Constant	26.486 (6.66)	30.970 (5.36)	23.188 (5.44)	31.008 (5.33)
Adjusted R^2	0.143	0.179	0.126	0.161
Number of cases	3744	3739	3744	3739
F statistics	18.887	22.470	16.358	19.807
Mean of dependent variable	58.766	59.630	51.095	49.025

Table A5 - Cross-Section Regressions - Fourth Grade - Portuguese and Mathematics: 1983 and 1985

	Portuguese		Mathematics	
	1983	1985	1983	1985
Student is repeating a grade	4.271 (3.42)	5.562 (4.85)	4.136 (2.60)	4.385 (2.92)
Student Characteristics				
Female student	10.415 (3.24)	10.745 (3.57)	-2.284 (-0.56)	-6.017 (-1.53)
Student's age	-0.583 (-2.81)	-0.962 (-5.02)	-0.768 (-2.90)	-0.798 (-3.17)
Pupil works	-5.990 (-3.70)	-6.665 (-2.90)	-3.235 (-1.57)	-5.682 (-1.89)
Family Characteristics				
Mother's education	0.215 (1.00)	0.332 (1.54)	-0.044 (-0.16)	0.241 (0.85)
Father's education	-0.129 (-0.50)	-0.455 (-1.83)	0.351 (1.08)	0.049 (0.15)
Family size	-0.392 (-2.46)	-0.014 (-0.10)	-0.062 (-0.31)	0.007 (0.03)
Peer Influence				
Percent families not farming	1.463 (0.62)	1.586 (0.67)	-2.157 (-0.72)	-1.810 (-0.59)
Relatively large landholders	0.042 (2.27)	-0.044 (-2.36)	0.037 (1.56)	-0.045 (-1.85)
Percent female classmates	8.592 (2.45)	5.527 (1.81)	9.911 (2.22)	3.746 (0.94)
Female classmates when female student	-7.688 (-1.77)	-5.236 (-1.32)	-12.672 (-2.29)	-0.569 (-0.11)
Percent seek 9 or more years school		7.237 (5.04)		9.168 (4.87)
Join Characteristics: pupil and school				
Homework	3.577 (3.70)	1.685 (1.91)	3.978 (3.23)	4.305 (3.72)
School lunch some day	-5.040 (-2.36)	9.666 (1.66)	-4.122 (-1.52)	11.416 (1.50)
School lunch every day	-5.907 (-2.59)	10.764 (1.84)	-3.042 (-1.05)	12.763 (1.67)
Male teacher/male student	1.364 (0.58)	2.406 (1.04)	1.937 (0.65)	5.022 (1.65)
Female teacher/female student	-1.170 (-0.57)	-3.602 (-1.92)	-0.499 (-0.19)	-2.535 (-1.03)
School Characteristics				
Graded class	-0.362 (-0.30)	2.705 (2.62)	-2.417 (-1.57)	-0.461 (-0.34)
Pupil-teacher ratio	0.086 (2.40)	0.099 (2.67)	0.150 (3.29)	0.115 (2.35)
School hardware index	8.433 (3.97)	-2.046 (-0.98)	7.639 (2.82)	3.015 (1.11)
School software index	1.518 (0.72)	4.554 (2.13)	4.613 (1.71)	6.654 (2.37)
Teacher Characteristics				
Teacher's education	0.543 (3.09)	-0.313 (-1.83)	0.816 (3.64)	-0.514 (-2.29)
Teacher's experience	0.002 (0.03)	0.041 (0.61)	0.043 (0.56)	0.240 (2.73)
LOGOS II -- teacher training	0.863 (0.67)	-0.124 (-0.11)	2.192 (1.34)	-0.549 (-0.36)
Qualificação -- teacher training	0.542 (0.40)	-0.662 (-0.56)	-0.046 (-0.03)	-3.129 (-2.02)
Teacher activity index	7.389 (2.70)	1.928 (0.77)	2.848 (0.82)	-5.809 (-1.78)
Teacher material index	-0.128 (-0.06)	1.073 (0.53)	0.428 (0.15)	3.378 (1.27)
Teacher's Portuguese test score		0.081 (2.06)		-0.059 (-1.15)
Teacher's mathematics test score		0.083 (1.95)		0.203 (3.64)
State				
Piaui	7.416 (3.58)	6.369 (2.79)	0.107 (0.04)	-0.777 (-0.26)
Ceara	11.825 (5.75)	13.051 (6.45)	9.383 (3.58)	7.832 (2.95)
State Program				
EDURURAL: Piaui	-3.486 (-2.03)	1.208 (0.67)	3.044 (1.39)	5.910 (2.49)
EDURURAL: Ceara	3.357 (1.63)	-0.141 (-0.09)	10.026 (3.81)	1.319 (0.63)
EDURURAL: Pernambuco	-1.223 (-0.71)	3.500 (1.92)	-1.859 (-0.85)	-1.847 (-0.77)
OME index	1.102 (0.44)	3.217 (1.40)	0.474 (0.15)	-2.586 (-0.86)
School Control				
State operated	-1.448 (-0.94)	-2.726 (-1.74)	2.998 (1.53)	-4.587 (-2.24)
Federally operated	-0.246 (-0.03)	-2.929 (-0.52)	-7.366 (-0.73)	3.325 (0.45)
Privately operated	14.587 (2.98)	-0.386 (-0.10)	3.943 (0.63)	-6.621 (-1.26)
Constant	37.042 (6.98)	20.036 (2.44)	35.698 (5.28)	27.396 (2.55)
Adjusted R^2	0.131	0.161	0.131	0.120
Number of cases	1448	1594	1448	1594
F statistics	7.236	9.023	7.255	6.693
Mean of dependent variable	52.019	48.682	47.813	50.142

VARIABLE DEFINITIONS

Student Characteristics

Age	Student Age (in Years)
Female Student	1 For Female Student
Pupil Works	1 If Student Works (Wording of Question Varied Slightly by Year)

Family Characteristics

Mother's Education fourth grade	Level of mother's formal education (For 1985, all responses greater than were coded as fifth grade)
Father's Education grade	Level of father's formal education (For 1985, all responses greater than fourth were coded as fifth grade)
Family Size	Number of persons living in the household

Peer Influence

Percent Families Not Farming	Proportion of families not farming (measured at school level)
Relatively Large of minimum Landholders	Proportion of Families own more than 35% of MODULO, a measure amount of land required to support a single family according to local land characteristics. MODULO is developed by IBGE.
% Sold Crops	Percentage of families who sell crops
Percent Female Classmates	Proportion of female classmates
Female Classmates When Female Student	Proportion of female classmates if the student is female. Otherwise it is 0.

Join Characteristics: pupil and school

Homework	1 if the student does homework always
School Lunch Every Day	1 if the school received lunch all year long.

School Lunch Some Days	1 if the school received lunch only some months a year.
Male Teacher/ Male Student	1 if the teacher and the student are both male
Female Teacher/ Female Student	1 if the teacher and the student are both female

School Characteristics

Graded Class	1 if it is a graded classroom
Pupil/Teacher Ratio	Number of students divided by number of teachers in school.
Hardware Index[1]	(FURN + FACIL + WATER + ELECT)/4
Software Index[2]	1983: (BOOKA/2 + BOOKB + WRMAT)/2 1985: (BOOKC/2 + BOOKD + WRMAT)/2
Teacher's House	1 if the school is in the teacher's house
Number of Students	Sum of the number of students in kindergarten through the fourth grade.

---

[1] FURN (DESK FOR TEACHER + BOOKCASES + SEAT FOR ALL STUDENT + PLACE TO WRITE FOR ALL STUDENT)/4, where each component of the index is a dummy variable

FACIL (TWO OR MORE CLASSROOM + MULTIPURPOSE ROOM + PRINCIPAL OR SECRETARY OFFICE + KITCHEN + BATHROOM + STORAGE ROOM)/6, where each component of the index is a dummy variable

WATER 1 if school has drinkable water

ELECT 1 if school has electricity

[2] BOOKA/C 1983: 1 if student uses the textebook some days a week
1985: 1 if student has book but uses it only at school

BOOKB/D 1983: 1 if student uses textbook everyday
1985: 1 if student has book and uses it at school and at home

WRMAT (CHALK + (NOTEBOOK + PENCIL + ERASER + COLORED PENCIL)/2)/5, where the first component of the index is a dummy which is equal to 1 if the school received chalk, and the other ones is 2 if the school received the material for everyone, 1 if the school received the material only for some students, and 0 if the school did not receive the material

Teacher Characteristics

Teacher Activity Index + TRIPS	(DRAMA + SINGING + MANUAL WORK + TELL STORIES + GAMES + GROUP STUDY + COMMEMORATIVES DAYS + CLEAN SCHOOL)/10, where each component of the index is a dummy variable.
Teacher Materials Index	(OTHER TEXTBOOK + MATERIAL WROTE BY TEACHER + MATERIAL WROTE BY STUDENTS + POSTER + MAPS)/9, where each component of the index is a dummy variable.
Teacher Salary	Teacher Salary as a Percentage of the Minimum Wage
Teacher's Mathematics Test Score	Teacher's score in mathematics -- same student's fourth grade test
Teacher's Portuguese Test Score	Teacher's score in Portuguese -- same student's fourth grade test
Years Teachers Education	Level of Teacher's Formal Education
Years Teachers Experience	Year of Teacher's Experience as a teacher
Logos II - Teacher Training	1 if teacher took LOGOS (in-service training)
Qualificacao Training	1983: 1 if teacher took Qualificacao training (inservice training) 1985: 1 if teacher took at least one "modulo" of Qualificacao training (inservice training)

School Control

State Operated	1 if State is a State School
Federally Operated	1 If it is a Federal School
Privately Operated	1 If it is a Private School

County Characteristics

% Emergencia	Percentage of families whose head of the household work in the "Emergencia" Program
OME	This is a more complex index, which take into account the number and qualification of the OME's personnel (See Armitage *et al.*[1986] and Harbison and Hanushek[fortcoming].)

SES	Six socioeconomic variables were used to built this index through a factor analysis (See Armitage *et al.*[1986].)
State - Program	
EDURURAL - Pernambuco	1 If the County is in Pernambuco and it is in Edurural Project
EDURURAL - Ceara	1 If the County is in Ceara and it is in Edurural Project
EDURURAL - Piaui	1 If the County is in Piaui and it is in Edurural Project
State	
Piaui	1 If State is Piaui
Ceara	1 If State is Ceara