SEMINÁRIO SOBRE EDUCAÇÃO, CRESCIMENTO E DESIGUALDADE NO BRASIL

Rio de Janeiro - 24-27 mar. 1991

SUMÁRIO

Doc. 11 - Good schools for poor students. Cláudio de Moura Castro; Sônia Dantas Guimarães & João Batista Araújo e Oliveira. 23p.

Doc. 12 - Do the returns to schooling vary across industries? Evidence from São Paulo, Brazil manufacturing. Donald Robbins & Mari Minowa. 30p.

Doc. 13 - Eucation, mobility and growth. Irma Adelman; Samuel Morley; Christoph Schenzler & Steven Vogel. 44p.

Doc.s/n - Wages, schooling and background: Investments in men and women in urban Brazil. John Strauss & Duncan Thomas. 49p.

Doc.s/n - Dealing with poor students. Cláudio de Moura Castro; Sônia Dantas Guimarães; João Batista Araújo e Oliveira & Sérgio Costa Ribeiro. 18p.

Doc.s/n - A note on Brazilian schooling investments in an international perspective: where does Brazil fit? Jere R. Behrman & Ryan Schneider. 13p.

Doc.s/n - A pós-graduação no Brasil: uma análise do período 1970-90. Ricardo C. de Rezende Martins. 30p.

Doc.s/n - Education and growth in Brazil: some cross-sectional evidence. Lawrence J. Lau; Dean T. Jamison; Shu-Cheng Liu & Steven Rivkin. 25p.

IPEA
Instituto de Pesquisa
Econômica Aplicada

Banco Mundial

PNUD
Programa das Nações
Unidas para o Desenvolvimento

SEMINÁRIO SOBRE EDUCAÇÃO, CRESCIMENTO E DESIGUALDADE NO BRASIL

SEMINAR ON EDUCATION, GROWTH AND INEQUALITY IN BRAZIL

Rio de Janeiro, Brasil

24 a 27 de Março de 1991

GRADE REPETITION, WASTAGE, AND EDUCATIONAL POLICY

João Batista Gomes Neto
Eric A. Hanushek

documento 10

Preliminary Draft

Grade Repetition, Wastage, and Educational Policy

by Joao Batista Gomes-Neto[1] and Eric A. Hanushek[2]

March 1991

"Education, Growth, and Inequality in Brazil"
Rio de Janeiro
March 25-26, 1991

[1] Universidade Federal do Ceará

[2] University of Rochester

Grade Repetition, Wastage, and Educational Policy

by Joao Batista Gomes-Neto and Eric A. Hanushek

Policy toward primary and secondary schooling in developing countries has often been regarded as involving directly trade-offs among school quality, equity and efficiency. The motivation flows from financial constraints on the school system, a common reality of public schooling in these countries. There are, in fact, several different, but related, formulations of the issues. One form begins with there being a shortage of available positions or "slots" for school. Therefore, efforts to provide secondary schooling or more expensive schooling comes at a reduction in schooling more generally available. Another form attributes the largest problems to high rates of wastage--grade repetition and drop out behavior. Because the schools are clogged with children who will never progress, others are denied access, and costs are excessive. Depending on the version employed and the underlying goals of the policy maker, a number of policy implications and proposals flow from such concerns.

Most of this discussion, however, is based on conceptual arguments and beliefs about what goes on in the schools, not on solid evidence about the actual operations of the schools. Very little empirical evidence about the interactions among costs, promotion patterns, and quality exist. This in turn makes the development of policy extremely difficult.

This paper presents direct evidence on these matters for Brazil. It begins with an overview of enrollment patterns in Brazil. It then moves on to present two types of information. First, an econometric analysis of the interaction between promotion and student achievement is used to assess the magnitude of any quality-equity trade-offs. Second, an investigation of the specific learning from grade repetition is used to provide information about qualitative aspects of the current system of repetition.

1. An Overview of Enrollment Patterns in Brazil

The Brazilian Constitution specifies mandatory school enrollment for every child between 7 and 14 years old. The school system does not, however, provide space for all children. A significant number of Brazilian children never enter the school system, and still more will drop out long before the legally fixed age is reached. According to Ministry of Education the total enrollment in Primary School is only 93.5% of school age children. This percentage is only an approximation of the capacity of the system, however, and it potentially provides a misleading picture of true attendance because there are many students older than 14 years (the denominator in the calculation). Simply modifying this to account for just the student enrolled at school between 7 and 14 year drops this ratio to 77.1%. Moreover, there are significant differences across the various regions of Brazil. According to Fletcher and Ribeiro,[3] the access to school ranges from 68% in the rural Northeast, which corresponds to 17% of the total Brazilian population, to 97% in urban areas of Southeast, the most developed area in Brazil.

These enrollment rates do not, however, tell the entire story. The system does not assure that those entering the system conclude even primary school, let alone secondary school. Actually, many of them do not conclude even the 1st grade. The repetition rates of the first grades could be as much as 74% in poor rural areas of the Northeast. But, even if we look at the richer urban areas in the South where repetition rate are the lowest, we still find 32 percent of the student in the first grade are repeaters. These figures point out the wastage in the Brazilian educational system.

These huge repetition rates drive to a enormous effort to have one graduate. For example, the effort to have one child in the 5th grade of Primary School could be as much as 36 student-years in rural Northeast and 8 student-years in the urban areas of the Southeast. The effort required to have a child in high school rockets to 136 and 23 student-years, respectively. According to the Ministry of Education,

---

[3]Philip Fletcher and Sérgio Costa Ribeiro developed a model, called by PROFLUXO, to estimate promotion, repetition and drop-out rates for the students in a school system using only census data. The applied their model to the 1982 PNAD data and estimated all these rates for different region and situations (as rural and urban) in Brazil. The summary statistics in this section, when not otherwise cited, rely upon their estimates.

Table 1 - Enrollment in Primary School and Population by Region, 1987

Region	Enrollment (1,000)		Population Age 7-14 (1,000)	Enrollment rate (percent)	
	Age 7-14	All		Net	Gross
North	1407	1770	1880	74.8%	92.0%
Northeast	6191	8010	8883	69.7%	90.2%
Southeast	8939	10409	10847	82.4%	96.0%
South	3188	3721	4092	77.9%	90.9%
Center West	1641	2067	2028	80.9%	101.9%
BRAZIL Total	20366	25937	27729	77.1%	93.5%

Source: Computed from data in SEEC/MEC: Estatísticas Educacionais: Brasil 1985-1988, Brasília, Jan, 1988, 2nd ed., and Sinopse Estatística do Ensino de 1° Grau 1977, Brasília, 1990.

only 15 percent of the children between 15 and 19 years old are in High School. This figure masks considerable disparity across regions: only 10 percent in the Northeast are in high school, while about 20 percent are in the Southeast.

Of course, the situation draw above is due the high repetition and drop-out rates. Many studies have been made to understand and explain this situation in Brazil.

Estimating the Repetition and Drop-out Rates

At the outset, it is necessary to estimate the repetition and drop-out rates. This, it turns out, is difficult, because different approaches yield quite different estimates of the situation. The Ministry of Education, through his statistical arm (SEEC), have come out with something around 25 percent for drop-out and 30 percent for repetition rates in 1st grade of Primary School. These figures have been argued by some authors as far from the true rates (Fletcher and Castro [1986], Fletcher and Ribeiro [1988]). Even the IBGE, another government source, come out with other rates, as we can see from the PNAD-82 where these rates were estimated in 5 percent and 34 percent, respectively. A longitudinal study made in Goias, although the sample was not a random sample, wherein 1008 students were followed for 8 years came out with 5 percent for drop-out and 53 percent for repetition rates in the 1st grade of Primary School. Our sample from Ceara in 1987 also give us an estimate of 121 out of 618 student sampled, roughly 20 percent, drop-out from 2nd to 4th grade in 2 years, while using enrollment reported by SEEC we can estimate a drop-out rate bigger than 50 percent in rural Northeast. To give a better idea of how different are the estimates from three sources (MEC, IBGE and PROFLUXO) we now display the three transition matrix for the first 4 (four) years of education.

Causes of School Failures

In order to make policy about wastage in the schools, it is essential first that we understand the underlying causes. The causes of failures in schools identified in past work suggest a division into two main factors: in-school and out-of-school factors.

A variety of social and economic problems have been identified as the main factors of out-of-school causes for school failures.[4] High direct costs--for example, for buying uniforms, writing materials, and text books, and the like--and sensitivity to opportunity costs to attend school are more likely to strike the children with low social and economic status. Other authors describe the malnutrition, which is clearly related to the social and economic status, as one of the causes of the school failures (Cunha [1981], Carvalho [1983]), although in our previous studies we could not find a strong relationship between malnutrition or receiving school lunch and student test achievements (Armitage et al. [1985], Harbison and Hanushek [1990]).

The in-school explanations concentrate on specific resource constraints and the general low quality of some schools. Low level of education, low salary and motivation, and attitudes and expectation of the teachers are pointed out by many researches as causes for school failures due to the teacher (Melo [1982], Brandao [1983], Armitage et al. [1985]). Lack of providing writing material and text book to the students, insufficient human and material resources and too little time in school are also described as school factor that cause high repetition and drop-out rates (Melo [1982], World Bank [Finance Primary

[4]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 social order (Popovic [1980], Oliveira [1981], Garcia [1982]).

Table 2 - Alternative Estimates of Repetition and Drop-out Rates by MEC, IBGE and PROFLUXO model

Grade	MEC		IBGE		PROFLUXO	
	repetition	drop-out	repetition	drop-out	repetition	drop-out
1st	29.6%	25.5%	33.8%	5.1%	54.7%	1.9%
2nd	20.7%	9.0%	15.1%	7.1%	36.0%	4.1%
3rd	16.9%	9.3%	11.6%	10.6%	28.3%	6.9%
4th	13.4%	4.8%	9.3%	18.5%	19.7%	18.5%

Source: P.R. Fletcher and S.C. Ribeiro - A Educação na Estatística Nacional, In: PNADs em Foco: Anos 80, D.O. Sawyer, org., pp 11-33, 1988.

Education-1986]). These arguments are frequently bolstered by data on aggregate expenditures. Per pupil spending ranges from US$24 to US$227 (World Bank [Finance of Primary Education-1986], Xavier and Marques[1984], Armitage et al.[1986].) These conclusions are also reinforced 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.

A primary objective of the research reported here is to identify various factors that enter into the enrollment patterns in primary schools. A related objective is the consideration of how alternative policies might affect enrollment, wastage, and efficiency in the schools.

2. The Brazilian Setting and Data Base

2.1 Rural Northeast Brazil and the EDURURAL Program

Beginning in the early 1980s, growing realization of the importance of education for development and the persistent welfare gap between the northeast and the rest of the country led the Federal and the concerned State Governments of Brazil to increase investment in schooling in the nine northeastern states. An important component of intensified Brazilian educational effort in the 1980s was the Northeast Rural Basic Education Project (EDURURAL). Planned in 1978-79 and launched in 1980, EDURURAL involved total incremental investment costs of US$92 million, of which US$32 million was financed with a loan from the World Bank. In 18% of the counties (municipios) of the nine states of the Brazilian northeast,[5] EDURURAL was designed to expand children's access to primary schooling, to reduce

[5] Brazil's Northeast Region, as officially designated by the IBGE (Instituto Brasileiro de Geografia e Estadistica, the national statistical office) comprises nine states and the Federal Territory of Fernando (continued...)

wastage of educational resources inherent in grade repetition and dropout as children progress through the system, and to increase achievement through improving the quality of instruction. Further, a hierarchical relationship was assumed among these three objectives. Improving learning achievement would decrease wastage (reduce repetition and dropouts), which in turn would make possible enrolling additional students in extant schools.[6]

The 218 rural counties which received EDURURAL assistance were selected (supposedly) because they were thought to be the least developed areas (especially in educational terms) in their respective states, and were not receiving special educational attention through other programs. The underlying principle was to concentrate sufficient resources in the most disadvantaged areas to make a real difference. In these selected rural counties, only schools outside the county seat -- often itself only a small town -- received the full range of EDURURAL inputs. Those incremental investments in instructional quality included: construction and refurbishment of schools, provision of furniture, training of teachers, development of curricula especially adapted to the poor rural environment, provision of textbooks and other student learning materials, and the strengthening of the county school administrative apparatus (OME - Orgao Municipal de Educacao).

During the period (1981-1987) when the EDURURAL project was implemented in 218 counties, a range of other educational improvement programs with similar objectives were underway elsewhere in the rural northeast. These other efforts typically sought at least one of the same three objectives and involved some, and occasionally all, of the same general kinds of inputs. But they differed among

[5](...continued)
de Naronha. The nine states, in alphabetical order, are: Alagoas, Bahia, Ceara, Maranhao, Paraiba, Pernambuco, Piaui, Rio Grande do Norte, and Sergipe. The combined land area is some 1.5 million square kilometers, roughly 18 percent of Brazil's total land mass. All nine states were included in the EDURURAL project.

[6] Access to education would also be attacked directly, through building new schools for unserved populations.

themselves and from EDURURAL in potentially important ways. The precise nature and mix of inputs varied from one project to another in different areas. One school might benefit from participating in the school lunch program. The teacher(s) in another might receive a salary supplement designed to reward their qualifications or simply to enhance their dedication to their jobs. Or teacher(s) might receive either general training in the form of academic upgrading and pedagogical techniques, or specific training on a newly developed curriculum. Still other schools might receive textbooks, or other instructional material, or furniture, or rehabilitation of the physical plant.

By contrast, EDURURAL sought within given counties in each state to provide to the selected schools a reasonably integrated and concentrated package of all essential inputs in a planned and rational manner. To meet the enormous managerial challenge of doing so, EDURURAL supported strengthening at federal, state, and county levels of the agencies involved in delivering public education. Where appropriate institutions did not exist, EDURURAL encouraged their establishment. In contrast, for example, to the primary schooling components of integrated rural development projects executed in several areas (frequently with separate World Bank support), EDURURAL was characterized by a certain focus and coherence as a single-purpose educational program. This, and the substantial participation of the World Bank, ensured that EDURURAL enjoyed a relative abundance of financial resources and other forms of special attention. Once EDURURAL was launched, counties not included because they were participating in some other program, eagerly sought incorporation into EDURURAL. In short, EDURURAL was the premier program among several lower grade education development programs being implemented simultaneously in the region. Together these programs offered an unusually attractive natural laboratory for learning how to improve educational performance among the rural poor.

2.2 EDURURAL Data Base.

Given the program's size and importance, EDURURAL's sponsors -- the Federal Ministry of Education in Brasilia, the nine Northeastern State Secretariats of Education, and the World Bank -- agreed upon an unusually comprehensive program of data collection and analysis to assess whether EDURURAL was meeting its objectives. The research began in 1980 when an unprecedented impact evaluation scheme was incorporated within the World Bank support for primary education in rural northeast Brazil. The evaluation plan involved extensive primary data collection to be conducted over a seven year period, paid for by the Brazilian Government with loan proceeds. The data collection from the very start was geared toward answering the straightforward impact questions: did the changes resulting from the increased funds improve student achievement, access to schools, and efficiency of schools? But from the beginning it was recognized that these deceptively simple impact questions could only be addressed within a far more elaborate research framework that would employ sophisticated analytical methods to delve into the fundamentals of how students learn and what factors affect student behavior.

The data collection supporting this research effort included preparing sampling frames for longitudinal data gathering, the testing of children (using special instruments produced by the project), and the surveying of an extensive sample of schools and teachers. The four panels of data referring to education in one of the poorest areas of the world are unparalleled in richness and in usefulness for addressing fundamental questions of educational performance.

The sampling frame was based on a random sample of schools found in both EDURURAL counties and a comparison set of counties (OTHER) in three of the participating northeast states: Ceara, Piaui, and Pernambuco. Beginning in 1981, data were collected on a random sample of up to ten second graders and five fourth graders in each school. Again in 1983 and 1985, if possible, the same schools were again sampled. The sampling was based on schools, with replacement of the large number of schools which were closed between sample years. Finally, in 1987, a special sample of previously sampled students in selected Ceara schools were located and sampled.

Table 3 displays the resulting samples. The underlying data collection involved an elaborate survey effort to obtain information on the background of students, the characteristics of schools and their teachers, and the performance of students on specially designed tests of Portuguese and mathematics.

The EDURURAL data set is superior to virtually all educational data sets for LDCs and, indeed, most for developed countries. The detailed longitudinal information allows a particularly productive line of inquiry into school operations. The framework for this inquiry, which exploits the panel aspects of the data, is described in the following section.

3. Specification and Estimation of Achievement and Promotion Models[7]

The analytical core of the existing research involves statistical investigations of what determines the availability of schools, the promotion of students, and the achievement of students. This is then linked to evaluation of the specific EDURURAL program along with more general considerations of cost-effectiveness of alternative policies for improving educational performance.

The completed analysis actually involves three separate investigations. First, because a number of schools disappeared between each successive sample, one line of inquiry involved modeling school survival. Second, many students did not get promoted from the second to the fourth grade, leading to an investigation of drop-out and promotion patterns. Third, and central to this proposal, the determinants of educational performance were analyzed. This section describes the promotion and achievement

[7]This section and the ensuing one on overall results is based on the results in Harbison and Hanushek[1990].

Table 3: SIZE OF SAMPLES

	1981		1983		1985		1987[Ceará]	
	EDUR	OTHERS	EDUR	OTHERS	EDUR	OTHERS	EDUR	OTHERS
Counties	30	30	30	30	30	30	6	5
Schools	397	189	404	195	477	195	48	32
Teachers	514	254	499	278	606	291	na	na
Students								
2nd grade	3037	1681	2619	1350	2950	1418	25	7
3rd grade							68	44
4th grade	1075	639	977	580	1273	631	103	107
Dropouts							22	4

modeling--the most important factors for the overall study.[8]

a. Promotion Models. The primary difficulty in analyzing school completion and promotion patterns has been a general lack of detailed data describing the paths of students through school and explaining the factors that influence decisions. In simplest terms, neither aggregate data nor data about a cross-section of students can support the kind of analyses that are required for policy purposes. The EDURURAL data set, while not explicitly designed for this purpose, goes some distance toward remedying previous data inadequacies.

The EDURURAL data collection was based upon repeated sampling from the student bodies in a set of schools drawn randomly within EDURURAL and control (OTHER) counties. The schools were observed at four different times (1981, 1983, 1985, and 1987), and during each observation a random sample of second and fourth graders was surveyed and tested.[9] This data collection design, in which interviewers returned to the same school every two years, offered an opportunity to observe individual students repeatedly. Most important, there was a group of students -- initially in the second grade -- who were progressing at the expected pace so that they were in the fourth grade in the follow-up sampling.[10] Whether or not a student was actually observed in subsequent data collection was a function of many intervening factors including purely random sampling chances.

---

[8]The third, school survival, is reported in Harbison and Hanushek[1990]. The results of this estimation are, as noted below, used in the achievement model estimation to correct for possible selection bias effects.

[9] The sampling in 1987 differed significantly from that in prior years. It can be described more as a student based sample design in which the primary factor driving the data collection was finding students who had been sampled in 1985. In this section, both the samples and the analytical models pertain directly to the 1981, 1983, and 1985 data. Harbison and Hanushek[1990] report analyses capitalizing on the unique aspects of the 1987 data.

[10] Additionally, there were two groups of students stuck in their initial grades -- second graders who had not been promoted, and fourth graders who had not been promoted, in the two year period between data collections.

The ability to use the EDURURAL samples to analyze questions about the quantity of schooling depends crucially on understanding the dynamics of the samples and utilizing the panel data on both schools and students. For analytical purposes, it is convenient to think in terms of probability models and to link the conditional probabilities of a series of basic events to their determinants. The difficulty in this analysis is that the observations of events are incomplete.

Two important linkages of the data across years can be identified. First, from the repeated sampling of the same schools, it is possible to identify whether or not a given school continues to serve its students over time. Contrasting schools that survive to those that do not offers insight into the prevalence of schools in the research area. Second, for those schools that both survive for the two year period and have a fourth grade, it is possible to find some second grade students who are promoted to the fourth grade by the subsequent data collection. Comparing promoted students with others allows some insights into the determinants of progression in school.

Whether individual student performance is related to promotion 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 promotion is only slightly related to actual student performance -- that is, the people being left behind or dropping out are about as good academically as those being promoted -- 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 other hand, if promotions are highly related to student quality, increasing the rates of promotion reduces wastage by continuing students with lower performance; the benefits of an external intervention program of lowering wastage would be much less. (A direct analysis of alternative promotion policies is contained in the last section of this paper).

The promotion probability models are estimated by probit techniques.[11] Because of the random sampling of students in the schools in each year, it is possible for an individual to be promoted on time but not to be included in the promotion sample. To deal directly with this, the probit models include 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 is significantly negative in the probit models, reflecting this sampling within schools.[12]

The most interesting part of the model is the relationship between second grade test scores and promotion probabilities.[13] Higher test scores consistently lead to greater promotion probabilities; this suggests that promotion has some basis in merit. Each 10 points on the Portuguese test, which has a standard deviation of approximately 25 points, increases promotion probabilities by about 1.5 to 2.5 percent. This implies that a student going from the 25th percentile to the 75th percentile on the test has

---

[11] Estimates of bivariate probit models which allow for correlations of the errors in the school survival and promotion equations were also done, but the correlation of errors was never over 0.001. Therefore, the results reported here are based on simple probit estimates for each equation. The complete probit models of promotion are displayed in Harbison and Hanushek (1990), table A4.3, along with means and standard deviations of the variables.

[12] If school size has an independent effect on promotion possibilities, this is mixed with the sampling effects. There is no clear reason, however, why size per se would affect promotion.

[13] A variety of other personal factors also entered into the promotion estimates. While interesting, they are not as central to the analysis in this paper. Other things being equal, females are over 3 percent more likely than males to stay in school and be promoted on time. Since the models incorporate differences in abilities, this reflects a lower opportunity cost of school attendance for girls; their value on the farms is less so they are less likely to quit school to work. Not surprisingly, promotion probabilities dip with age. The older a student is when sampled in the second grade, the more likely the student has already repeated grades or dropped out for some period. Therefore, it is less likely that the student will be promoted to the fourth grade on time. In the earlier period, each additional year of age lowers the probability of promotion by 0.7 percent, while in the later period this estimate rises to 1.7 percent. Since the mean promotion probabilities are respectively about 9 percent and 14 percent for the two periods, this effect of age is substantial.

The education level of a student's mother is positively related to promotion. This reflects both family "tastes" for education and direct aid in education at the home. The education level of the father was tested in the models, but had no additional independent impact, perhaps reflecting the conventional wisdom that the mother, not the father, is strongest educational influence on the child. The lasting impact of low education levels is seen from the intergenerational nature of the transmission of human capital from mothers to children; low attainment of this generation hurts not just this generation but also future generations.

5-9 percent higher promotion probabilities. Between the 10th and 90th percentile, promotion probabilities rises by 9-17 percent. Since the mean observed promotion rate in the sample is only 9 percent in 1983 and 14 percent in 1985, these are significant differences due to merit. Performance on the mathematics test does not have as strong an influence on promotion. It is statistically insignificant in the 1981-83 period and has about one third the effect of the Portuguese test in 1983-85. (The standard deviation of the mathematics test score is approximately equal to that for the Portuguese test.)

This factor is central to the later analysis of alternative programs and their costs. Specifically, this demonstrates the linkage between school quality and repetition.

b. Achievement Models. The overall framework for analysis follows a quite standard input-output specification for the educational process. The achievement of a given student at time t (A_t) is assumed to be related to current and past educational inputs from a variety of sources -- the home, the school, and the community. To highlight some of the important features, we consider a general conceptual model such as:

$$(1) \quad A_t = f(\mathbf{F}^{(t)}, \mathbf{S}^{(t)}, \mathbf{O}^{(t)}, \epsilon_t)$$

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

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

$\mathbf{O}^{(t)}$ = 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 of achievement at time t.

This conceptual model is described explicitly in a regression like format which incorporates a stochastic component, ϵ_t, reflecting the fact that we can never observe all of the factors that affect achievement. The approach is to measure the different possible inputs into education and to estimate the influence of each of the factors on student achievement. This emulates the approach adopted in the

Table 4: **DISTRIBUTION OF INITIAL YEAR STUDENT BY FOLLOW-UP YEAR STATUS: 1981/1983 AND 1983/1985**

STUDENT DISTRIBUTION	INITIAL YEAR	
	1981	1983
TOTAL SECOND GRADERS	4677	3918
School Nonexistent - Follow-up Year	1936	1141
School Survived - Follow-up Year	2741	277
Initial/Follow-up Total Matched		
Second Grade - Second Grade Match	n.a.	126
Second Grade - Fourth Grade Match	227	349
Fourth Grade - Second Grade Match	39	41

Hanushek[1990]--are too extensive to report in detail. We instead concentrate on one set of inputs, material inputs provided in schools. This set is particularly important for our discussions because it relates directly to how varying resource policies might affect overall performance of the system.

It should also be noted that this basic analysis confirms the central importance of quality teachers. Variations in teacher quality, while not captured well by traditional measures such as teacher education or experience levels, are nonetheless central to much of the performance differences of students. It is not the central part of this analysis because the supply function for high quality teachers is not understood. Specifically, it is quite unclear what it would cost to obtain a higher proportion of better teachers. Therefore, the kind of cost-benefit analysis done below is not possible with respect to teachers.

A primary thrust of the policy initiative of the EDURURAL program was improving the learning environment for the students. This had multiple attributes: insuring that adequate teaching materials were available, improving the quality of teaching, providing administrative support, and so forth. This section concentrates on measurable differences among schools, and the next delves into teaching quality.

Two categories of school factors are considered, roughly hardware and software. The analysis explicitly investigated how an aggregate measure of school facilities (hardware) and measures of both writing materials and textbooks (software) are related to performance. These factors are particularly important, because they are readily adjusted through governmental policies. They are also inputs that have entered significantly into previous investigations of the educational process in LDCs.

Improved facilities are systematically beneficial to student learning.[15] The results for fourth grade indicate that supplying all measured components of the facilities index to a school which previously had none of them could increase student achievement by a 9-13 points. While somewhat more ambiguous for the second grade, the overall picture is that quality of the physical plant is positively related to student

---

[15] The components of the hardware index include the availability of specific kinds of physical plant (more than one classroom, kitchen, sanitary facilities, storage space, offices), of specific items of furniture (desks and chairs for pupils, table for teacher, bookcases), of water, and of electricity. The value of the index varies from 0 (representing the absence of all component of the index) to 1 (representing the presence of all components).

performance.

Past research has quite generally found that the availability of writing materials and texts is important in schooling for LDCs. The results here reinforce that view, although again the results, particularly for textbooks, are not as statistically significant as one would like in order for strong policy statements to be made.

The 1985 value-added findings (and the second grade results for 1983 and 1985) support the importance of adequate writing materials for the students. The size of the coefficients suggests that achievement gains of roughly a third to one half of a standard deviation may be acquired by supplying a package of writing materials to all students.

Textbook effects are more ambiguous. The results for the second grade consistently support their importance. The results in the value-added models are estimated with large errors and have the wrong sign in 1983.

Combined with the strength of previous findings in other studies, these results support the call for improving the availability of materials to poor rural schools.

4. Cost-Benefit Analysis

A fundamental question for the educational policymaker is which specific school inputs are most efficient in raising student achievement scores, given an available level of resources. Often, policy is made solely on the basis of analyses like that reported in the previous section. That work provides direct estimates of the benefits of altering the different inputs to the educational process. It thus allows the inputs to be ranked on the basis of their effectiveness. Yet, while this is better than having no empirical basis at all for policy determination, it is perilous. Specifically, it does not incorporate anything about the costs of such alterations. This section integrates the previous analysis of marginal educational effects with estimates of the costs of making different input adjustments. Efficiency rather than effectiveness becomes the focus of analysis.

Efficiency calculations -- the appropriate basis for assessing different policies to educate a given student population -- involve the joint consideration of outputs and the costs of inputs required to implement any policy. In the best of all situations, the outputs can be valued in monetary terms so that costs of inputs can be compared directly to the resulting outputs, that is, cost-benefit analysis. In our situation, output is measured in terms of academic achievement which cannot be translated into monetary terms. Therefore, we concentrate on the closely related cost effectiveness analysis (see Lockheed and Hanushek 1988, 1990.)

This section concentrates on dynamic gains in efficiency of schools. The dynamic gains are conceptually straightforward even if seldom empirically demonstrated. When students learn more because of more or better inputs to their schooling experience, they are more likely to be promoted at each point in time (as reported in the previous section). This reduces the total time they spend in the system in order to reach any given grade level. Increasing the flow through the system implies cost savings, since fewer student-years of schooling services have to be provided on average for a student to reach the given level. These savings offset the costs of instituting the original policy change.

The cost reductions attributable to improving flow through schools can be quite substantial. The levels of repetition displayed previously imply that the average student arriving in the fourth grade in rural northeast Brazil will already have spent 7.6 years in school, instead of the three required by steady on-time progression.[16] Of course, this vastly understates the overall economic cost of attaining that level of schooling, since it ignores the resources expended on students who enter school but never reach fourth grade. In fact, in rural northeast Brazil an average of 15.2 student-years of schooling services is provided for each

[16] The numbers would change a little, but not the overall conclusions, if four years, not three, were deemed the standard time to reach fourth grade. This would make nominal provision (as Brazilian law does not) for the de facto kindergarten, represented by primary enrollees in the ano de alfabetizacao. Full data on progression in schools is available in Harbison and Hanushek (1990), tables A3.3 and A3.4.

student who reaches fourth grade. Repetition and dropout multiply the cost of a "graduate" fivefold.[17] Since wastage is so high, even small improvements in promotion probabilities (which, of course, imply decreases in rates of repetition or dropout) can result in significant savings.

Thus, this final analytical section calculates the offsets to gross program costs arising from improved student flows. The result is estimates of net cost effectiveness. While this is the appropriate criterion for considering policies, we are aware of only one other attempt to consider such feedback effects (Jamison 1978).

The results of this exercise are extraordinary. A wide range of investments made to improve educational quality actually "make money." The savings from improved flow efficiency are often larger than the original costs of providing improved inputs in the schooling process. The finding of net cost gains through improved efficiency generally holds even when allowance is made for uncertainty in the estimates. The normally postulated "trade off" between quality and quantity of schooling appears to be quite the opposite in circumstances of severe educational deprivation: instead, there is a positive interaction wherein enhanced quality engenders increased quantity.

a. Methodology. The empirical work sketched above provides the necessary ingredients for calculating what we call "partial benefit cost ratios," which are designed to net out the savings that accrue from better flows through the system. The achievement models provide estimates of the gains in achievement from changing inputs. When combined with cost estimates for providing the inputs, one can calculate expected achievement per dollar of expenditure on a given input. These are available for models of second and fourth grade achievement in both Portuguese and mathematics in the different years. For our purposes, we consider the different sets of parameter estimates of the achievement models for each subject, by year and

[17] We use the term "graduate" loosely here to denote a student who arrives in fourth grade. The true flow efficiency of the system for each completer of fourth grade is even less than indicated, because both repetition and dropout occur during the fourth grade year as well as earlier in the primary cycle.

grade, to be alternative estimates of the same fundamental underlying relationships of the educational process. Similarly, the promotion probabilities associated with different achievement levels are available for 1981, 1983, and 1985 from the on-time probit promotion models described above (and presented in Harbison and Hanushek[1990]).[18]

The expected number of student-years that accumulate before a person reaches any given grade level are directly related to the promotion and dropout probabilities at each grade. The lower the promotion probability, the slower students will progress through the system and thus the larger will be the number of years that go into producing a primary school graduate. For evaluation purposes, we base our calculation on estimated transition probabilities derived from the experience in various regions of Brazil in 1982.[19]

Finally, any savings in student years must be transformed into dollar values. Using information obtained directly from our survey data on teacher salaries and input costs (see Armitage et al.(1985)), we obtained US$29.57 as our estimate of the cost per student year of primary schools in the rural Northeast.[20]

[18] The estimated marginal probability associated with any input investment will vary, depending on where the probit models are evaluated. The marginal probability associated with a change in a given input is the probit coefficient times the ordinate of the normal distribution evaluated at the initial probability. For these calculations, we evaluate the probit models at the initial probabilities estimated for the region by the Profluxo model.

[19] The correspondence between promotion or dropout probabilities and years expended on schooling can be calculated by following a cohort through school until everybody either has been promoted or has dropped out. This can be derived mathematically, and the formula is used throughout these calculations.

The promotion and dropout probabilities used for different regions come directly from the PROFLUXO model of P. Fletcher and S. Costa Ribeiro (1986), which employs data from the 1982 Brazilian household survey (PNAD82). The basic transition probabilities (and the subsequent calculation of student flows) are based on individual grades. Our probit models, however, indicate on-time promotion between second and fourth grade. To apply these probit estimates, we assume: (i) that any change in promotion probability is evenly distributed between second and third grade; and (ii) that changes in the total promotion probability are proportional to the estimated change in the on-time promotion probability.

[20] This figure combines the mean teacher salary from our sample data with the complete hardware and software packages (identified in Harbison and Hanushek (1990), annex 2) costed out at the values in chapter 6 of that document.

Table 5: FLOW IMPROVEMENTS AND PARTIAL BENEFIT-COST RATIOS FOR SELECTED INVESTIMENTS IN LOW-INCOME RURAL NORTHEAST BRAZIL

	Mean estimates		Maximum Estimates
	Second Grade	Fourth Grade	
A. Student years saved per dollar invested in:			
Software	0.2316	0.1342	0.4206
Hardware	0.0465	0.0796	0.1011
B. Dollars saved per dollar invested in:[a/]			
Software	6.95	4.03	12.62
Hardware	1.39	2.39	3.03

a/ Years saved valued at US$30 per student year.
Source: Harbison and Hanushek, tables A6.5 and A6.6

The analogous figure from the best available Brazilian study is US$31.50;[21] for rural schools in the interior of the Center West states, the figure calculated by the same authors is US$33. Given the consistency of these three separate estimates, we have used a round figure of US$30 as the cost per student year when evaluating the value of time saved.[22]

b. Results. Table 5 displays both the years saved and the dollars saved per dollar invested from the key quality-enhancing resource inputs to schooling of hardware and software. These calculations rely on promotion and drop out probabilities for "low income rural northeast Brazil," the combination of geography and income status that most nearly approximates the areas in which our surveys were conducted.[23] These inputs were selected for analysis because they are often -- and were in the EDURURAL project -- the chosen instruments of public policy aiming to improve the quality of primary schooling.[24] The figures in table 5 are the mean and maximum of the estimates from the alternative models of promotion and achievement.[25] In the calculations underlying the mean benefit-cost ratios, the point estimates of all positive coefficients were employed without regard for statistical significance; all negative coefficients were treated as zero, or as having no relationship.

---

[21] See Xavier and Marques (1984). Their equivalent work on schools in the Center West states is reported in Winkler (1986), p. 29, table 10.

[22] These cost calculations assume that marginal costs -- the incremental dollar savings that would result from one fewer student-year -- are the same as average costs.

[23] Using data from the 1982 household sample survey (PNAD82), P. Fletcher and S.Costa Ribeiro, in elaborating their PROFLUXO model, further subdivide the normal geographical classification of the Instituto Brasileiro de Geografia e Estadistica (IBGE) into income groups, defined standardly across all Brazilian households included in the 1982 survey on the basis of regionally determined minimum wages and selected socioeconomic characteristics of the family.

[24] As indicated, a variety of teacher factors are not included here. For example, the teacher's command of the subject matter, revealed in earlier analyses to be among the most important determinants of the learning achievement of children, is not among the inputs to schooling selected for partial benefit cost analysis. In the absence of a fully specified supply function for teachers, it is not feasible to calculate the cost of providing teachers with some incremental amount of subject matter knowledge.

[25] More extensive results, disaggregated by grade and year, are available in Harbison and Hanushek (1990), tables A6.5 and A6.6.

Since the underlying achievement models are so different in analytical perspective, we report second and fourth grade results separately. Because the second grade models, based on cross-sectional or level models of performance, have not eliminated various potential sources of bias that could contaminate results, we rely most heavily on the fourth grade results and tend to treat the second grade results as simply reasonably plausible upper bounds. For purposes of this discussion, we assume that the methodologically stronger estimates at fourth grade are proxies for what would have been obtained by a similarly rigorous approach at second grade -- that is, the "exaggerated" second grade findings are attributable to the methodological deficiencies of those estimates rather than to any underlying differences in the educational production process between second and fourth grades.

The results are stunning. The direct material inputs -- hardware and software -- produce much more than the original investment in dollars saved from increased flow efficiency. In other words, by investing in known quality-enhancing resources, it is possible to produce the same number of fourth graders, although fourth graders of higher quality, with no true additional costs, just savings.

Further, the magnitude of these net benefits can be breathtaking. The partial benefit-cost ratios can be greater than 2.0, signifying that twice the original cost of the investment is returned quickly in savings resulting from increased flow-efficiency brought about by investing in inputs which engender achievement gains. At least in the severely deprived environment of rural northeast Brazil, investment in school quality is a real money machine.

Are these estimates important only in the extreme conditions of the rural Northeast? What do the results say about investments outside the low-income areas of the rural northeast?

As one might expect, the partial benefit-cost ratios are highly sensitive to the underlying transition matrices for movements from grade to grade. The benchmark, repeated in Table 6, is that an investment in software in low-income rural northeast Brazil will return about US$4.00 for each dollar it costs (US$6.95

Table 6: PARTIAL BENEFIT-COST RATIOS FOR SELECTED INVESTMENTS IN VARIOUS REGIONS OF BRAZIL

	Mean estimates					
	Brazil		Rural northeast		Urban southeast	
	All	Low-income	All	Low-income	All	Low-income
2nd grade estimates						
Software	1.40	4.93	5.38	6.95	1.03	0.90
Hardware	0.28	0.99	1.08	1.39	0.20	0.18
4th grade estimates						
Software	0.81	2.86	3.12	4.02	0.60	0.52
Hardware	0.47	1.69	1.84	2.39	0.35	0.30

Note: Years saved valued at US$30 per student year.
Source: Harbison and Hanushek (1990), table A6.9

if the second grade cross-section models are used rather than the fourth grade value added specification). But if the level of educational wastage began at that prevailing in low income Brazil generally, the payoff would be only about US$2.90. While the decline is substantial, this is still a remarkable figure. If the sample areas of the rural northeast started at the further reduced repetition and dropout levels prevailing in the most advantaged areas of the country (that is, high-income urban southeast), the offset to investment costs, while still a considerable US$0.52 per dollar of investment, would no longer exceed initial costs.

An alternative interpretation of the data of table 6 puts these calculations into an overall development perspective. Suppose it is assumed that the underlying education production function is roughly the same in all primary schools (with variations in the quantity and quality of inputs explaining the known differences in outcomes) and that relative costs of inputs are the same throughout the country. While it could be argued that these are strong assumptions if comparing the very worst areas with the very best, it is much more plausible when not dealing with the polar extremes. In these circumstances, the partial benefit-cost ratios broken down by geographical area are reasonable indicators of the results to be had from investments in quality-enhancing inputs outside the rural northeast. Given these assumptions, the data demonstrate that, for most combinations of geography and income in Brazil, educational wastage remains high enough that investments in at least some, and often several, quality-enhancing inputs have partial benefit-cost ratios greater than one -- that is, they pay back in monetary savings more than the cost of the investment. This conclusion, again, <u>ignores</u> the value of higher achieving students and cumulative impacts higher up in the educational pyramid.

These results effectively rank order the efficiency gains to be had in educational investments in various parts of the country.

4. Direct Evidence from Grade Repetition

The previous analysis concentrated on the interaction between school quality and promotion. We believe that that evidence is most persuasive and that tackling the problem of quality improvement should be the first order of business. The problem of wastage is, nevertheless, pervasive and, realistically, unlikely to be solved in the near future. Therefore, we also present some preliminary information about gross learning effects of repetition.

This potentially ties into another way of tackling the wastage problem that come from high repetition and drop-out rate: mandatory promotion. Indeed, if promotion 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, where the promotion was strongly related with the student achievement. In this latter case, we expect mandatory promotion to lower the grade effective level of achievement, damaging overall school quality.

On the other hand, the incredibly high repetition and drop out rates existing in the Brazil school system, especially in the primary school, increase the cost of getting a graduate--since money are 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 exactly what the nature of the trade-offs is.

A central question is how student achievement is affected by repetition and, inferentially, by mandatory promotion. While far from ideal, our data do allow some preliminary analyses of these issues.

Two special groups of sampled students can be used to do this. First, we will examine the matched students in 1983/1985 who were in second grade in both years. We are going to compare them with those who were promoted on-time. This comparison is, however, somewhat complicated because the repeater group was given the second grade tests both in 1983 and in 1985 while the promoted group was given the second grade tests in 1983 and the fourth grade tests in 1985. Second, we will examine the students in the special sample for 1987. This sample, for a subset of Ceara schools, does allow direct comparisons of repeaters and those promoted on-time on the basis of second grade tests. As described in Harbison and Hanushek (1990), all students took the same tests in 1987 regardless of whether they were in second or fourth grade.

Matched 1983/1985 Analyses

A total of 3944 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 ones were found in the fourth grade. The means and standard deviations of the Portuguese and mathematics achievement scores in the second grade for these groups are found in the table 7.

As we can see from table 7, in 1983 the mean achievement in Portuguese as well as in mathematics of the repeater group is below the respective means for all second graders. By 1985, however, the means of the repeat group are slightly above the means of all students in second grade. In other words, students learn something when they repeat a grade. But it took them two year more to catch up with the other students. Their initial and later performance can also be contrasted with the mean achievement of the on-time promotion group which scores a quarter to half of a standard deviation above the mean.

TABLE 7 - Mean and Standard Deviation of the Portuguese and mathematics achievement for the matched students and for the entire sample in second grade.

		Portuguese	Mathematics
second grade			
Achievement in 1983	mean	58.7	51.2
	s.d.	23.6	24.9
	n	3944	3944
On-time promote			
Achievement in 1983	mean	68.6	56.8
	s.d.	17.9	22.4
	n	379	379
Repeat			
Achievement in 1983	mean	40.2	35.7
	s.d.	25.1	25.3
	n	127	127
Achievement in 1985	mean	61.1	52.4
	s.d.	22.7	25.1
	n	127	127

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 repeater and promoted groups. The distributions were calculated using z-scores, where the means and standard deviations relied on scores for all second grade students in 1983.

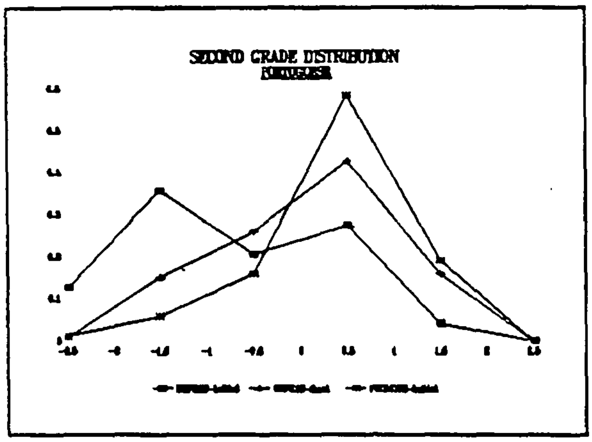


Figure 1

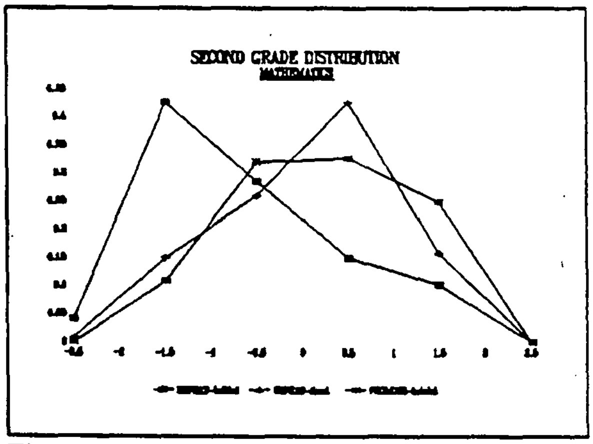


Figure 2

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. Importantly, 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.[1] The post-repeating scores relate to a modified plan of a fixed number of years in each grade.

Tables 8 and 9 display the transition probability matrices used for Portuguese and mathematics performance. These come directly from the matched sample of promoted students.

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

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

costly--the full cost of two years of schooling, mandatory promotion may be an effective alternative to the current system.

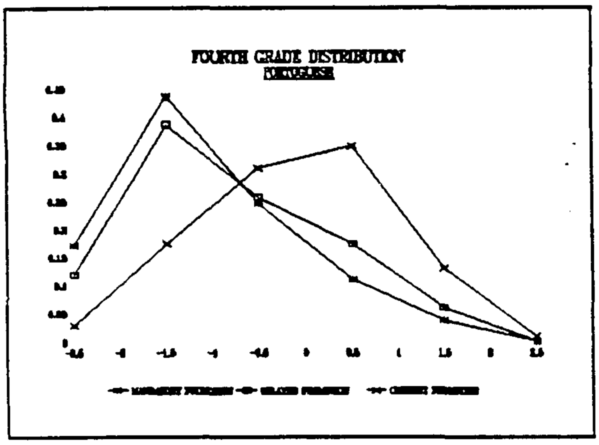


Figure 3

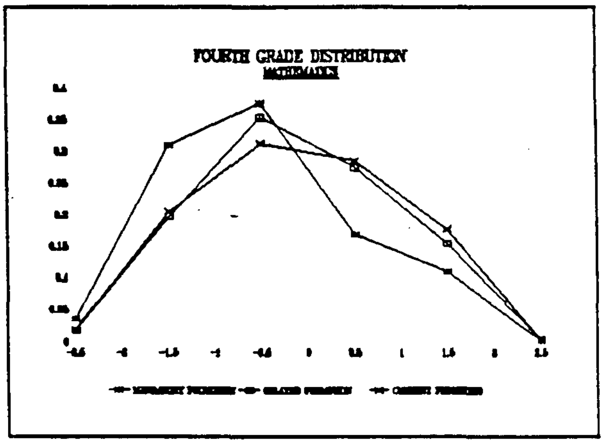


Figure 4

Matched 1985/1987 Analyses

An alternative approach is to investigate the special sample of matched students from Ceara in 1987. While this is not a representative sample, it was possible to trace the progress of 352 students who were in the second grade in 1985. Of these 32 still were in the second grade and 209 were in the fourth grade. They took the same second grade tests they had taken in 1985, and the means and standard deviations from these students are shown in table 10.

We repeated the estimation of the effects of mandatory promotion by comparing repeaters to those who were promoted. Tables 11 and 12 give the transition probabilities, calculated from those promoted on-time. Figures 5 and 6 give the initial (1985) distributions of scores. Figures 7 and 8 give the 1987 distributions of scores estimated under different conditions. Again, while these

TABLE 8 - Estimates of the conditional probabilities -1983/1985

Previous Achievement in Portuguese Z-score	Follow-up Achievement in Portuguese Z-score <= -2	between -2 and -1	between -1 and 0	between 0 and 1	between 1 and 2	>= 2
<= -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
> 2	0.00	0.01	0.18	0.57	0.21	0.03

TABLE 9 - Estimates of the conditional probabilities - 1983/1985

Previous Achievement in Mathematics Z-score	Follow-up Achievement in Mathematics Z-score <= -2	between -2 and -1	between -1 and 0	between 0 and 1	between 1 and 2	>= 2
<= -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
> 2						

TABLE 10 - Means and Standard Deviation of the Portuguese and mathematics achievement for the matched students and for the entire sample in second grade (1985/1987)

		Portuguese	Mathematics
second grade			
Achievement in 1985	mean	59.6	49.2
	s.d.	25.2	25.0
	n	4321	4321
On-time promote			
Achievement in 1985	mean	75.7	61.2
	s.d.	15.0	21.6
	n	209	209
Achievement in 1987	mean	89.3	77.2
	s.d.	9.9	16.4
	n	209	209
Repeat			
Achievement in 1985	mean	50.0	46.3
	s.d.	27.2	22.6
	n	32	32
Achievement in 1987	mean	77.6	69.8
	s.d.	22.2	25.3
	n	32	32

TABLE 11 - Estimates of the conditional probabilities - 1985/1987

Previous Achievement in Portuguese Z-score	Follow-up Achievement in Portuguese Z-score					
	<= -2	between -2 and -1	between -1 and 0	between 0 and 1	between 1 and 2	>= 2
<+ -2	0.00	0.00	0.00	0.67	0.33	0.00
between -2 and -1	0.00	0.03	0.00	0.39	0.58	0.00
between -1 and 0	0.00	0.00	0.01	0.20	0.79	0.00
between 1 and 2	0.00	0.00	0.00	0.09	0.91	0.00
> 2						

TABLE 12 - Estimates of the conditional probabilities - 1985/1987

Previous Achievement in Mathematics Z-score	Follow-up Achievement in Mathematics Z-score					
	<= -2	between -2 and -1	between -1 and 0	between 0 and 1	between 1 and 2	>= 2
<= -2	0.00	0.00	0.46	0.38	0.15	0.00
between -2 and -1	0.00	0.00	0.02	0.45	0.48	0.05
between -1 and 0	0.00	0.02	0.04	0.23	0.65	0.06
between 1 and 2	0.00	0.00	0.08	0.15	0.71	0.06
> 2	0.00	0.00	0.00	0.00	1.00	0.00

estimates are very crude, they suggest an intermediate strategy that employed some more rapid promotion policies might be in order. There is some loss in overall quality (i.e., the fourth grade distribution is shifted to the left), but the rough estimates do not indicate an enormous shift.

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.

On the other hand, this evidence reinforces a conclusion of the more detailed previous analyses. The dearth of information about the entire process of promotion, repetition, and dropping out behavior makes informed decision making extremely difficult.

6. Conclusions

The one conclusion that we reach with some certainty is that investing in quality schooling is a good investment. The gains in system efficiency, by our estimates, more than compensate for the costs of certain productive investments.

The level of information about the related issues of repetition, wastage, and achievement is, nonetheless, disastrously low. There is no doubt that increased research directed at these issues would pay enormous dividends to Brazil.

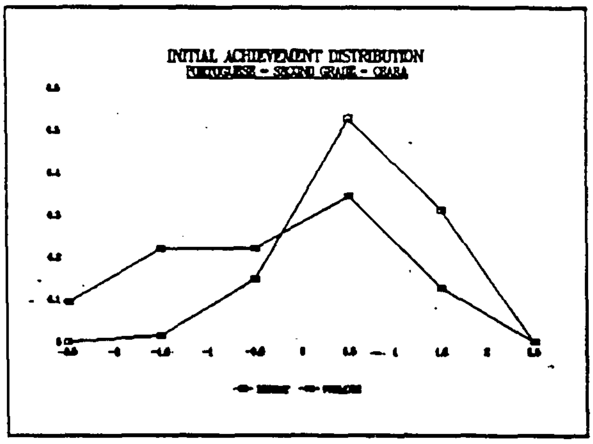


Figure 5

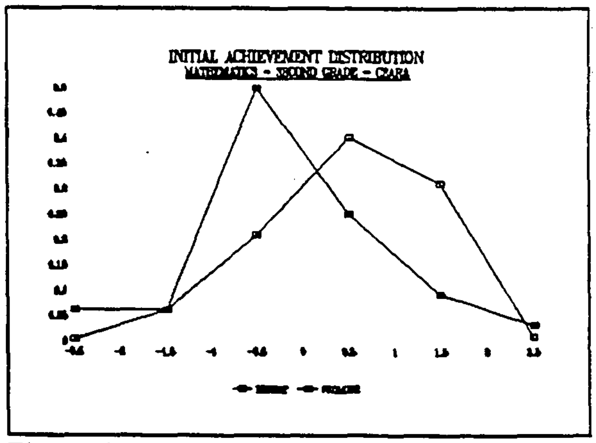


Figure 6

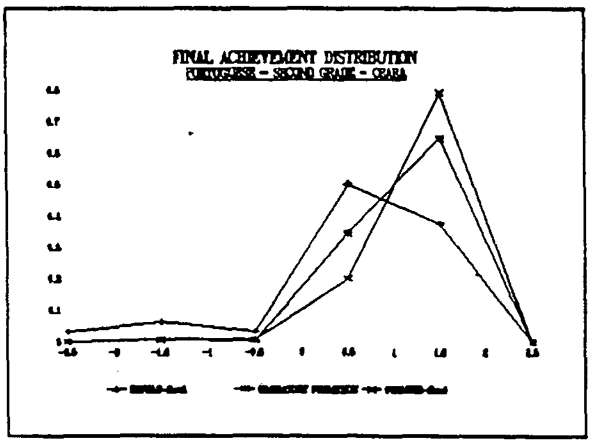


Figure 7

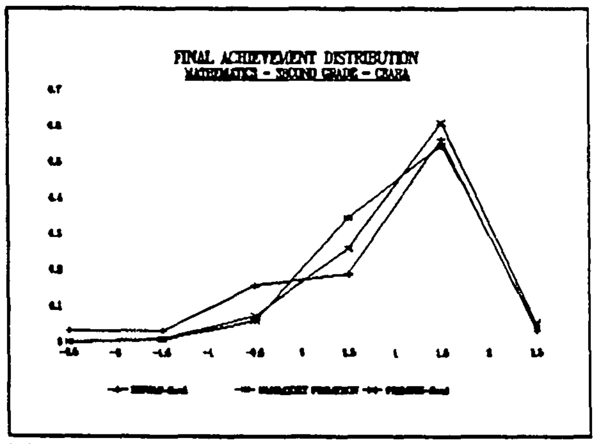


Figure 8

References

Armitage, Jane, Joao B.F. Gomes Neto, Ralph W. Harbison, Donald B. Holsinger, and Raimundo H. Leite. 1986. "School Quality and Achievement in Rural Brazil," World Bank Education and Training Series, No. EDT25. Washington, D.C.

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, jan/fev 1983: 44-59.

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

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfield, and Robert L. York. 1966. Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office.

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 1º grau. Educação e Realidade. Porto Alegre, 11 (1), jan/jun 1986: 35-42.

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.

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

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

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

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

Harbison, R.W. and E.A.Hanushek. 1990. Educational Performance of the Poor: Lessons from Rural Northeast Brazil. Washington, D.C.: The World Bank (mimeo).

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

Jamison, Dean T. 1980. "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.

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

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

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

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

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, jan 1981: 49-60.

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

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.

IPEA
Instituto de Pesquisa
Econômica Aplicada

Banco Mundial

PNUD
Programa das Nações
Unidas para o Desenvolvimento

SEMINÁRIO SOBRE EDUCAÇÃO, CRESCIMENTO E DESIGUALDADE NO BRASIL

SEMINAR ON EDUCATION, GROWTH AND INEQUALITY IN BRAZIL

Rio de Janeiro, Brasil

24 a 27 de Março de 1991

GOOD SCHOOLS FOR POOR STUDENTS

Cláudio Moura Castro
Sônia Guimarães
João Batista Araújo e Oliveira

documento 11

Good Schools for Poor Students

Cláudio de Moura Castro, Sonia Guimarães e João Batista Araujo e Oliveira. Draft, Feb. 1991.

The son of the worker in the school of the boss

What is the cost of a typical primary public school in Brazil? How much would it cost to provide a school of good quality, comparable to that of the elite private schools in the big cities? How do costs and administrative procedures differ between these two types of schools?

Brazil lacks both traditions and models of quality education for the poor. The great majority of public schools which cater to the poor are clearly inadequate. Even when their intentions are good, there are structural difficulties which make sustaining high standards in the long-term very difficult. As a rule, the poor receive low quality education (Fundação Carlos Chagas....), and that makes it even more difficult for them to succeed in these schools. The country badly needs to identify initiatives which, in one way or another, could suggest ideas and models of good schools for poor students.

In the Brazilian educational scenario one can find poor schools for poor students, and rich schools for the rich. There are virtually no good and rich schools for the children from lower SES families. When they exist, they are exceptions to confirm the rule. Examples include the experimental schools (Colégios de Aplicação) the model schools like the Colégio Pedro II, or an outstanding principal or teacher here and there, who manage to make the difference - in spite of the system.

This paper describes the costs of a set of schools very peculiar and rare in Brazil. This is a group of schools for the children of the workers from big enterprises or subcontractors operating in some far- away regions of the country. We do not assume that these schools are good, or that they make any

difference for these students. We also do not attempt to defend the private school system, particularly because it would be hard to define as private a school financed by a public enterprise. Rather this paper attempts to examine these schools, how much they cost, what they are, and whether they make any difference for the children of the poor.

The set of schools analysed are all managed by the same private educational institution. There is an explicit attempt to keep the same standards of quality of the elite school they operate in a major capital city. This is reflected by the general norms, criteria for personnel selection, and staffing patterns. Each school, however, is managed under contract with a different enterprise. Each has specific costs, rules, and almost total local managerial autonomy. This paper also analyses the only school owned and directly operated by the institution in the capital city, which is considered one of the best three, and which will serve as a technical parameter for the comparisons.

The costs of these schools - henceforth called private - are compared with those of the state and municipal public schools in the same sites or towns where they are located. These are the public schools where the children of the workers would go - if it were not for the availability of these private schools. Since these private schools offer a quality similar to that of the school in the capital, it becomes possible to calculate the costs of offering a rich school to poor students.

In another parallel study the authors examine the educational characteristics and outcomes of these two types of schools and attempt to perform some cost-effectiveness analysis.

Data and Methods

This is a semi-experimental study, which is methodologically warranted when the examination of the average or typical cases does not lead to an understanding of the phenomenon under analysis. Throughout the country we looked for circumstances or

accidents which would reveal a reality normally denied by the majority of the schools. This search led us to find some schools in which the children of the workers went to the same school of those of their bosses.

A previous study by one of the authors (Castro, 1976) showed that the schools in the satellite cities around Brasilia used to offer similar education to that available in the schools of the Plano Piloto where bureaucrats live. This is a similar case of a decent school for the poor.

In these types of studies, there is no way to talk about representative sampling. On the contrary, the sample was based on the opposite criterion, in order to reveal one aspect of the reality which would be impossible to detect in the regular schools.

We selected 3 out of a possible 13 schools operated by the educational institution, a major criterion being the availability of other public schools in the area with which they could be compared. With this restriction, and others which appeared in the sampling process, we end up with the three schools finally chosen.

The public schools were selected on the basis of proximity to the locale where most of the workers live. In other words, we tried to select the schools where they would normally send their children. In each case we selected one state and one municipal school, ending up with 3 private and 6 public schools. We also included the school in the major capital - to serve as a parameter of a typical elite school.

More descriptive data on the teachers, schools, students, and academic performance are presented elsewhere (Castro, Guimarães, Oliveira, Ribeiro, 1991). This paper discusses the costs indicators and the organizational issues related to allocation and management of resources.

Cost data were collected using a modified version of the questionnaires of the ECIEL studies (ref.). The responses to the

phenomena of a more general nature. For example, one of the schools éxamined received an entire library with over 20,000 volumes as a donation. Adding the value of these books to the costs of the school would be a distortion.

By the same token, very large or under-utilised schools would be excessively expensive, and their costs would not necessarily help explain differences in performance. On the other hand, the costs of sophisticated pedagogic equipment have been included, since they reflect the intention of providing a different type of teaching environment. In the case of large samples, the law of large numbers would cancel out these accidental factors, and corrections would not be recommended. But this was not the situation in our limited and differentiated sample.

What we did, in practice, was estimate the average area of the schools (square meters per student), and we did it separately for the public and the private schools. These averages were multiplied by the average cost of construction (which does not vary significantly in the two cases). Ten per cent of these costs were added to correspond to the capital costs, as described before.

Costs of land can also be distorted in many ways. There is no reason to compute the total cost of a lot when a mayor decides to build a school in a 20,000 square meter lot. To make it simple, we increased the costs of construction by 30% to take care of land costs.

Furniture and equipment are important inputs for the operation of the schools. They were exhaustively listed, and their costs estimated by one or more qualified persons, in each place. Costing these components did not present major difficulties, and they represent very little in the total costs.

In principle we intended to include operational costs (water, electricity, telephone,etc). Unfortunately it was impossible to obtain such costs, and in one of the few cases in which the data were available, they did not make any sense, and had to be

more than 50 questions of this questionnaire were obtained in the field by one of the authors, and allowed her to map out all the factors of education production and their related costs. There was a built-in redundancy in the questionnaire, which allowed the same question to be examined from different angles. The data on the questionnaire were completed with interviews with the principals, local authorities and *in loco* observation. In the case of some of the private schools it was also possible to collect data directly from the financial department of the sponsoring enterprise.

The purpose of the questionnaire was to determine the costs of the inputs in the school, such as building and facilities, equipment, pedagogic materials, as well as the costs of maintenance, operations, personnel (teachers, technical and administrative staff), and other costs associated with school lunches, transportation and school uniform, the latter two typically borne by the families.

Capital costs were analyzed in a peculiar way. The majority of cost studies usually consider only recurrent costs, obtained in the balance sheets or financial statements. It is obvious that the schools have to pay their bills, and need this kind of information. However, in a discussion about resource utilization, this is not enough. It is essential to include the alternative uses of the resources (opportunity costs) mobilized in the construction of buildings or in the acquisition of equipment. To make a complex discussion simple (Castro & Assis, ...), we used shadow-prices as indicators of capital costs. This is the rental cost of similar facilities, which was added to the operational costs. We also added to the shadow price a capital cost of 10% a.a. (including depreciation and maintenance).

Costs reflected in accounting documents are defined by administrative and legal norms. In a cost study, definitions derive from the purpose of the research. In this case, we wanted to compare different schools, in order to understand why some function better than others, or why some are more expensive. But costs also reflect accidental factors which, if included, could hide

discarded. As it was observed in this and other similar studies. these costs have very minimal impact on total costs (between 1 and 5%). and we decided to ignore them.

Personnel costs are the most important data in any study of this kind. In this case. the data were sufficiently reliable. Personnel costs were obtained from the pay-roll. Inflation was dealt with by using the value of the official dollar in the 15th day of the month of the pay-roll. Obviously the dollar is not a perfect deflator and salary adjustments occur by quantum leaps. We had no way to account for these two limitations. We also decided to exclude indirect costs associated with personnel. In the case of private firms - within which our private schools are included - these social security and benefits can represent 50% and more of the pay-roll. Even though such costs do not exist for the public schools. there are still some social security benefits which are covered by higher levels of administration. Thus. we decided to exclude these costs in both systems.

The cost of instructional materials used by students was based on the actual cost of the list prepared by the schools - independently of the fact that they were either bought by the students. provided by the schools or not bought at all. This procedure overestimates the costs in the public schools. where most students did not buy or did not have the books and materials listed. In the private schools. the list is either bought by the school - and deducted from the pay-roll. provided by the PTA or bought directly by the students.

How schools and costs differ

As it is the case in most cost studies. we are primarily interested in trends. orders or magnitude and comparisons. Given the nature of the data. the limitations associated with data gathering. and the corrections and adjustments made. the actual absolute figures must be analysed with much caution. Relative values are less risky and more instructive. since the biases tend to

cancel out each other. Other studies (Xavier,...), however, tend to arrive at figures similar to the ones found in the present study.

Table I presents the annual student/cost ratio (US$). Direct costs (column 1) include personnel and instructional materials. Column 2 includes capital costs.

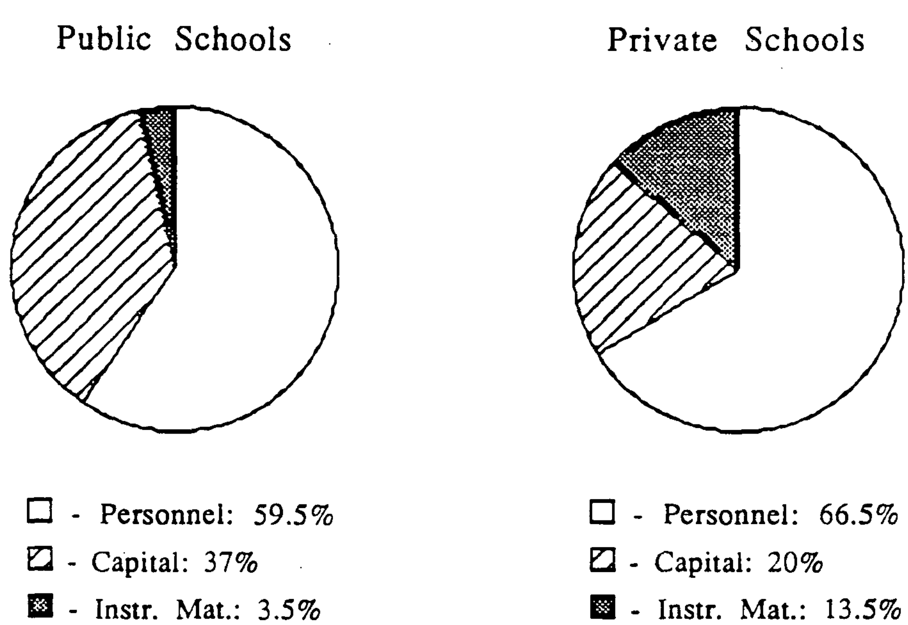
TABLE I

Direct and Totals Cost (US$)

	Direct Costs	Total Costs
Public	102	136
Private	376	488

Figure I, below, illustrates the composition of these costs in the two systems, broken down into three major categories of cost.

FIGURE I
Breakdown of major costs



The data from Figure I illustrates a different pattern of resource allocation between the two systems, which is analysed below. As expected, personnel represents the majority of the costs, 59.5% in the public schools, and 66.5% in the private. These costs are relatively low when compared with other cost studies which do not include capital costs of the expenditures borne by students and their families.

If the private schools spend more on personnel than the public ones, how is this money spent? In the public schools, teachers represent 65% of the total personnel expenditures, and 72% of the total costs of the school (not including social costs). These proportions are respectively 58% and 76% in the private schools. This means that private schools spend more on personnel, but they spend even more on non-teaching personnel. These are the technical staff, which are practically non-existent in the public schools, but which represent more than 18% of the pay-roll of private schools (For details please refer to Annex 1, Summary Cost data).

There is a striking difference in the size and allocation of administrative costs. Private schools pay fairly high salaries to their principals - usually 5 to 10 times that of the teachers in these same schools. In the public system these salaries are virtually indistinguishable. However, the majority of the administrative costs of public schools - which represent 28% of personnel costs - are related to cleaners, security guards, and other non-qualified personnel. Their salaries, however, are very similar to the low salaries of the teachers. It is timely to remark that there is no correlation between the number of janitors and the degree of school cleanliness or the quality of school maintenance. In fact, the opposite was probably true.

Equipment expenditures refer to two cost categories: (i) school equipment, i.e. desks for teachers and students; and (ii) other materials, which include shelves, office furniture, audio-visual equipment, and miscellanea, such as stoves, refrigerators

and fans. The differences between the two systems again are striking. In the public schools, school equipment (i) represents 65% of these resources, and in at least two cases it was insufficient or in very bad shape. The majority of the public schools in the sample did not have any kind of material resources for administrative or pedagogic support. All schools, however, had stoves, an essential item for the school lunch program.

The private schools offered a variety and diversity of materials, particularly those directly related to instruction. All private schools had a library with at least 500 books - usually much more - and a monthly record of 600 individual requests. No public school had a library.

Expenditures on instructional materials represent a rather modest sum, in both systems. In the case of the public school, there are two complications. Two of the six schools provided lists of materials (books and supplies). In these two cases, the costs of the items on the list were computed, regardless of their availability. In the public schools, the availability of textbooks was attributed to donations by the FAE (Fundação de Assistência ao Educando) in previous years. In no school in the sample, books arrived in 1990, and in no schools were there books available for every student in the third grade. By contrast, in the private schools the books and materials are compulsory, and when it is not bought by the parents (directly or through the school), it is provided by the PTA.

Limited as they might be, these data suggest some trends and differences among these 10 schools. Two clear-cut groups emerge. On the one hand, four private schools, each operating with large degrees of autonomy, are very similar in the way they spend their resources. On the other hand, the public schools, state and municipal ones, display an enormous variance in the way they allocate their resources. Indeed, not only the salaries and expenditures are different, but the variance in the patterns of resource allocation within the public schools reveals an enormous differentiation. In other words, in the public schools there are no

rules or patterns for resource allocation, personnel policies or other expenditures with maintenance, books or instructional materials.

An apparent distortion exists in the capital costs, which represent 37% in the public schools, and 20% in the private. In the private systems, students have 3 times more space than in the private one. Yet, since total personnel costs are lower in the public system, construction costs represent a higher proportion.

At the same time, it must be noted that school equipment represents a minimum fraction of total costs. Even the best equipped school does not spend more than 1% of total costs on equipment, and the worse school, less than half of one per cent. We will come later to the implications of such a pattern of expenditure.

Finally, it must be observed that the major difference between these two systems is reflected in the expenditures on teaching materials as a percentage of total costs: 3.5% in the public, as opposed to 13.5% in the private. And one must be cautious because the cost in the public schools is probably overestimated.

It is reasonable to assume that pedagogic equipment, teachers and books make a difference in the quality of teaching, and that the buildings have little or no influence. Yet, one observes that the major differences between these two systems reveal exactly the opposite trend, as far as public schools are concerned: they have relatively high expenditures on buildings, which make no difference, and almost nil expenditures on instructional materials, which directly affect quality.

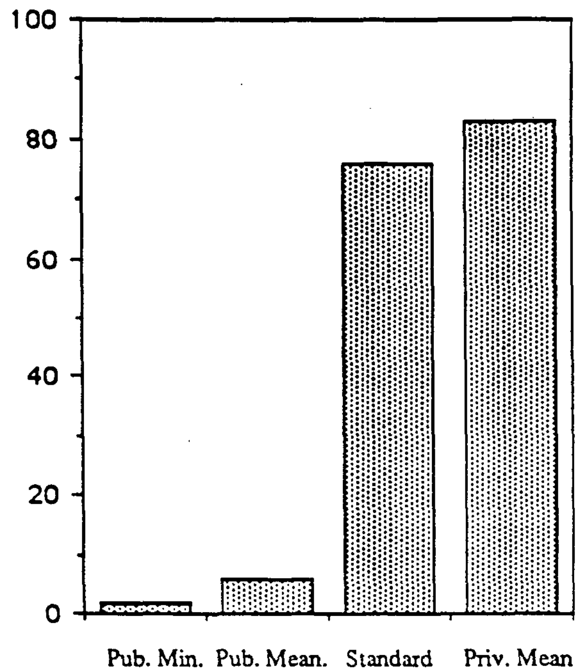
Good and bad expenditures

A further analysis of the data presented in the previous section will concentrate on three differences between the two

systems: instructional materials, teacher salaries and management.

Instructional materials. Figure II illustrates the striking differences in the allocation of resources for instructional materials: public schools, in average, spend 11 times less than private schools. The difference between the extremes is 52 times (lower, public and higher, private).

FIGURE II
Instructional Materials



These results are important for a number of reasons. First, the proportion 1:52 is virtually the same found by other authors (Farrell and Heyneman, 1990) between OECD and developing countries. In other works, private schools in Brazil have a level of expenditure compatible with that of OECD countries, and public schools, with that of developing countries.

Second, total expenditures with instructional materials, and textbooks in particular, represent a minimum percentage of total costs, even in the public schools, where the over-estimated

average is US$ 6.50. In spite of the abundance of these materials in the private schools, only in one case do they reach US$ 80. It is appropriate to notice that the ECIEL studies in 1975 revealed comparable results for Latin America, including Brazil. Yet, the literature is very clear about the importance of instructional materials, and even more so for students from lower SES (Heyneman & Loxley, 1983; Lockheed and Verspoor, 1990).

Third, there is a clear difference on how these two systems deal with the issue. In the private schools, books and materials are compulsory, there is no schooling without them. They are a problem to be dealt with by schools, the parents or the PTA. In the public school, they are a problem of the federal government (FAE). It is up to FAE to send or not to send the books, this is not a problem of the school - or for the school. And even if they send the books, this does not mean that the schools wants them, that they are even used (Oliveira & Guimarães, 1985).

Available data reveals very neatly the under-investment in books and instructional materials - and this is more difficult to explain given the relatively low weight of this item on total costs. This is probably one of the most cost-effective investment that public schools systems could make.

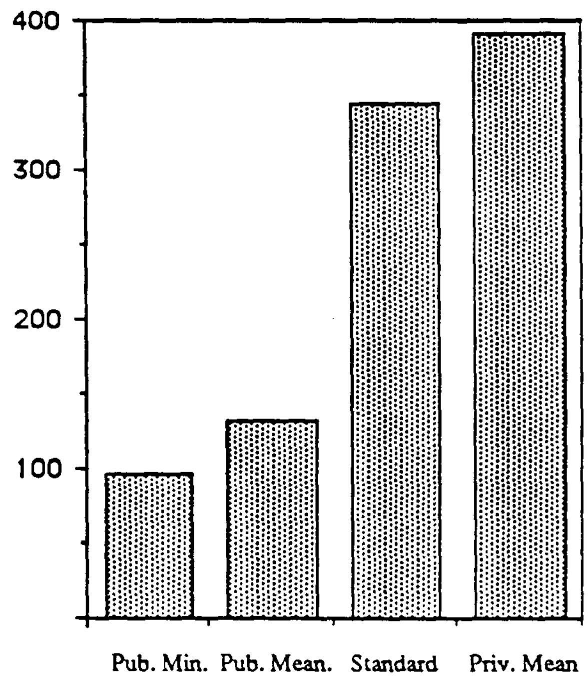
Teacher salaries. It is not a new finding that teachers get low salaries. It is also not new that public school teachers get even less than their colleagues in the private sector. But our data present some news.

Figure III shows the average salaries for a typical 20-hours/week contract for a certified secondary school teacher (escola normal or equivalent). In the standard elite private school, teachers receive, in average, 2.6 times more than those in the public schools.

This difference, however, tends to be significantly reduced when corrected for effective teaching hours. Dividing the total number of available teaching hours in a school (the total of hours

contracted and paid for) by the number of hours of operation (shifts x number of classes) yields a proportion of 1.77 in the public schools, and 1.28 in the private. This represents an actual

Figure III
Teacher Salaries



cost of US$ 2.30 per effective hour in the public system, as opposed to $4.90 in the private, a difference of 2.1 times. This means that in the public schools the contractual arrangements contribute to increase the level of inefficiency: teachers receive salaries equivalent to 20 hours of work, but, in average, they teach little more than 10 hours. If this inefficiency were equally shared, it could be interpreted as an indirect form of remuneration. In other words, all teachers would be paid for 20 hours, but they would all teach less than that. However, in practice this mechanism only reveals additional inequalities within public schools: some teachers work the contracted hours, while others teach much less, depending on the subject matter or on the number of classes. Similar results were observed by Castro and alia. in public schools in the state of Rio de Janeiro (Castro et alia....).

This disparity is less accentuated among the primary school teachers (grades 1-4), where there is one teacher for the whole class. But this only reveals an even greater disparity between these teachers and the ones teaching in grades 5-8, where distortions concentrate, and salaries tend to be even higher. In other words, in the public school system there seems to be an enormous remuneration gap between those teachers who effectively work as stated in their contract - and these are the ones that receive less and work more - and the others who teach in more advanced grades, work much less, and hence receive much more per worked hour.

In addition, the hourly salary for the public schools has been overestimated, given the fact that we assumed that all classes would be operating during all shifts - which evidently is not the case, particularly in the evening. As a result, the actual cost per hour is even higher, and the cost differences with the private systems, even narrower, representing less the double.

A previous study by Castor (1988) attributes the widespread existence of this problem to imperfections in the statuses regulating teacher careers, work contracts and collective bargaining. On the basis of a diagnostic undertaken by the State Secretariat of Education in Paraná, it became possible to renegotiate rules of allocation of teachers to classes on the basis of more objective criteria. Without salary losses, and with a slight reduction in teaching staff, it became possible to enrol an additional 260,000 new students in the public school system. The study mentioned above not only confirms the existence of the problem, the inefficiency of the work contracts and the inequities in the allocation of benefits, but it also indicates the possibility of an effective and equitable intervention by the public authorities.

Management. The idea of management itself reveals a major gap between the two systems observed. In the private sector, the salaries of principals are totally different from those of teachers, revealing a preoccupation with management. More than salaries, the functions of the principals in these two systems are

radically different, and reveal much more than the monetary values of' their salaries.

In the public school there is virtually nothing to be managed: books come (or do not come) from FAE, based on criteria that change every year, or with each new Minister of Education; teachers are hired by the state or the municipality, and the principal has no say in these matters; the school has no autonomy of any type. Moreover, the parents, particularly those from lower SES, tend to have low expectations of their children and the schools. Even when they have higher expectations, they have no channels for expression, and little or no power over the school.

Thus, without any internal or external pressure, and being accountable to no one, the principal is left with the administration of daily chores. Indeed, this confirms the findings of a previous of a study by the World Bank (Armitage, ...) in which the author reveals that in rural schools of Northeast Brazil where educational performance was close to zero, the bureaucracy and the paperwork were actually in order. Our field observations confirm that the principal in these public schools perceives himself as a minor civil servant, similar to any other bureaucratic civil servant. It is noticeable that they feel impotent in dealing with personnel or educational matters. All principals refer to the inefficiency of the personnel, but emphasize that there are no conditions to do anything about it.

The most important managerial function would seem to command the battalions of janitors jammed in the pay-roll of these generally decrepit schools. Yet, not even this seems to be the case. As previously mentioned, there is no relation between the number of the these employees and the appearance of the school, or its size. In one of the schools, the vice-principal said that she is not even able to ensure that washrooms are properly kept. In emergencies, one has to use the lavatory of a neighbouring private home!

In the private schools, by contrast, the principal is accountable to the sponsoring firm, which, in turn, receives pressure from the employees to provide a decent school. Such pressure comes particularly from the more high qualified workers, for whom a good school is part of their contractual requirements. Besides, the director relies on a qualified technical staff and manages the budget and staff as would any other manager responsible for results. Instructional materials, as previously mentioned, are a school problem, and are considered an integral part of the basic inputs - not an uncertain gift from the authorities in Brasília.

In short, the previous analysis shows, on one side, a public school in which the government acts as if it were offering education, as if it were paying for services, while the teachers feign that they teach, and the students make believe that they are learning. This is not a judgement based on the merits - sometimes overwhelming - of dedicated individuals. But, as a whole, it is an overall system of make-belief, in a country in which reality is more brutal than fiction. Even though the educational systems represent an enormous financial effort, it is insufficiently funded, and the scarce resources are inefficiently spent. Wrong decisions, insufficiencies, inefficiencies and equivocal signalling are observed in the policies and practices of educational financing, in both state and municipal schools.

Conclusions

Even though our sample is not necessarily representative and certainly limited, the main conclusions of the present study both illustrate and reflect those of many other studies (Edurural/WB; ECIEL; Xavier; Castro and Fletcher, 1985, among others).

It is not possible to state, on the basis of this paper, whether or not the costs of the private schools contribute to worthy academic results. Nonetheless some conclusions and inferences are possible. First, it would be possible for a public school to offer

the minimum essential conditions typical of a private school (teacher salaries and instructional materials) with a relatively limited additional financial effort. Independently of learning gains, it cannot be denied that the quality of the inputs in the two systems is incomparable - and not proportional to the costs. In other words, the public resources wasted on non-essential activities are extremely high, and probably unjustified. To put it another way, the public school is what it is, and the teachers are unfairly remunerated not necessarily because of a lack of resources, but mostly due to resource misallocation.

It cannot be concluded from this cost analysis that the private schools are good, or that they reach the desired or expected results. But if the cost study leaves doubts about that, one does not need to go too far to perceive some interesting indicators about their merits. There is a public school on the same street where the elite school is located. While the fee paying private school rejects applicants, the tuition-free public school has empty places. In one of the three private schools of the sample, about 30% of students come from families of non-employees, who decided to pay its costs, rather than to send their children to the local schools. In these schools, the dismissal of an employee creates a tremendous shock for the student who must leave the school and has to return to the local public school. In one of the schools visited, this is almost a daily occurrence, and the panic is widely perceived. Good or bad, these schools tend to be preferred by both those who can and those who cannot afford their costs.

The present study highlights the margins of manoeuvre which could be conquered through more effective and equalitarian public policies. Three clear directions for action seem to emerge.

The first one refers to school management. Education occurs in schools. Schools need administrative and financial autonomy to become efficient. Otherwise, the school will be perceived as a typical bureaucracy - or a mere deposit of janitors and teachers -

a victim of patronage politics. Yet, the experience of the private sector demonstrates that autonomy by itself is not enough. To be effective, autonomy must be coupled with accountability. The environment of public schools creates little or no pressure, and this is even more critical when there are no public exams or other external mechanisms of quality control or incentives for quality. Without these ingredients, autonomy and decentralization can only contribute to increase inequalities (McGinn, 1985; Oliveira, 1991). Quality control, examination systems and decentralized management can only be efficient if they co-exist. What we saw was a school system completely centralized on bureaucratic and administrative matters, but totally free to offer education of any imaginable (or unimaginable) quality.

It is never enough to insist on the importance of educational materials. In Brazil, the Ministry of Education has proven its capacity to operate a rather complex system of school lunches (Castro & Fletcher, 1985). Yet, it has never been able to provide schools with basic instructional materials, such as blackboards, chalk, textbooks, and other instructional materials which cost much less and which are much simpler to deliver. By centralizing the problem, the central government waved local authorities from any responsibility over these matters - and yet, it failed to provide the solution. Nonetheless, books and instructional materials - which seldom represent more than 5% of educational expenditures in OECD countries - may have a substantial impact on education quality. This is an area for simple and relatively cheap solutions, but that it has not yet received adequate attention from government. In other words, the government demonstrates its capacity to administer a complex operation such as the school lunch program, but at the same time demonstrates its incapacity to deal with the issues of instructional materials.

Finally, teacher remuneration policies need to be scrutinized. Our data reinforces the thesis that educational resources are indeed insufficient. Education is costly, and quality education, even costlier. Additional costs of quality, however, are well justified, and could be obtained through better resource allocation,

before new resources are pumped in. Public schools, in general, display a typical pattern of under-financing, in which a threshold has yet to be obtained, before quality education can be delivered. As it is, the present system generates enormous inefficiencies, reflected in the low teacher salaries, the low levels of performance and the high attrition rates of students. Two alternatives exist to cope with the problem of teacher remuneration: the creation of attractive careers, or the adoption of market criteria. Or a combination of both. In either case, more resources would be needed for a lasting solution.

However, our data reveals that, to a great extent, resources which would be sufficient to double teacher salaries already exist - thus roughly enough to equating the salaries of public teachers with those presently paid by the elite, private schools. A competent management of the educational sector could greatly contribute to reduce the rates of repetition, thus reducing the demand for new school places and correcting student flows. Some states have already started revising the criteria for staffing schools, thus unfreezing additional resources and making the teaching careers more efficient and equitable. Administrative costs at the state and municipal secretariats sometimes represent up to 25% of total education budgets - and these could be significantly reduced without damage to the schools (Xavier, ...). Even in a state like São Paulo, the present level of administrative expenditures (about 7%) could be reduced without any pedagogical harm. Finally there are also pockets of inefficiency within the schools themselves: 15 to 20% of resources for personnel are spent on activities which are either unnecessary or totally irrelevant for educational quality. Savings in these various categories could generate resources sufficient to double the salary of teachers.

As stated in the introduction of this paper, it was not our purpose to boast the virtues of private schools. In fact, we do not even know if the schools in our sample could indeed be considered as private. Rather, our analysis attempted to

illustrate what it is possible to do with other management practices and with other ideas.

Implications for public policies

The analysis of the cost patterns of the public education system reveals a fundamental problem of institutional design: the logic of the system, as a whole, precludes schools from performing adequately. Schools are under-financed; there are neither the resources nor the autonomy for local management; there is no quality control; there is no pressure for quality; the millions spent on the structures of administration and technical staff in the central bureaucracies do not reach the schools in any visible way. There are tremendous inequalities in the allocation and use of resources. All these factors illustrate the lack of an appropriate institutional model.

What would be the characteristics of an appropriate model? It is at this point that the analysis of the private school system would help. What makes a good school? Based on the evidence presented above, here is an outline:

• schools as the center of decisions. Resources and power are allocated at the school level. The principal is responsible - and accountable - for what occurs and for what fails to occur. All aspects of the school are under local responsibility - including the provision of instructional materials;

• salaries are defined by the market - i.e., salaries which are effectively paid are those necessary to attract manpower with the characteristics desired by the school;

• resources are efficiently allocated. This includes the resources for technical staff within the school;

• mechanisms of control and evaluation. The principal is evaluated by the sponsors - and is under both constant direct and indirect pressure of the parents.

The public school system examined shows exactly the opposite characteristics - and its results are also known. The challenge consists of defining a new institutional logic that would contribute to the reversal of this situation. If the behavior of the private schools studied above could serve as a source of inspiration, here are some suggestions for thinking about an effective model for public schools:

• to define a minimum and realistic level of *per capita* resources, to be allocated to the schools. There are strong indications, including those based on the costs of the private schools studied above, that an initial allocation of U$ 300 per capita would be a reasonable starting point. This roughly represents double what is presently allocated in public school systems such as those studied here. Such resources would be enough to pay salaries comparable to those presently received by the private sector - if other inefficiencies were simultaneously abolished. Even if this level is not the ideal, it could represent a good start.

Doubling resources alone would be neither sufficient nor efficient for three major reasons:

(1) present patterns of allocation of resources would need to be modified; otherwise new resources would be simply misallocated;

(2) changing present resource allocation mechanisms can significantly improve the efficiency of schools, particularly by diminishing rates of repetition, now draining about 30% or more of the existing resources. The elimination of this bottle-neck would generate additional resources, thus diminishing the pressure for new investments on new buildings;

(3) new resources could only be efficiently used if coupled with other institutional changes, such as:

• increasing the autonomy of the schools to manage their resources, including human resources, in more efficient ways;

• developing attractive teaching careers based on market or market-like criteria and rules; moreover, such a system would need to create incentives and balance individual career security with local authority and management discretion;

• creating public mechanisms of school evaluation. This is particularly critical for the public schools, where parents of lower SES usually have no conditions to pressure the school to improve its quality;

• reducing the costs of central administration, and redirecting technical staff and resources directly to the schools.

The problem is systemic and one of logic: additional resources without major institutional changes do not alter the existing logic, and may lead to inefficient results. The origins of the present situation and the failure to address it correctly suggest that these considerations have to be considered against a broader, political framework. The idea of public goods has been lost - if it ever existed. The present public school system is only public in its name - and it carries a very negative connotation.

Economic authorities, nonetheless, can no longer ignore the fact that the underinvestment in education and the misallocation of resources cannot be justified on the grounds that the system is public: there is nothing inherent to the public sector which would prevent it from adopting more efficient management practices. This does not mean that the adoption of such measures would settle all education matters once and for all. But it means that it is only by adopting a logic which allows for an effective and equitable operation of the school system that the government will be able to create the conditions for the public school system to start operating *de facto*.

Resumo dez Escolas

	Pitágoras B.Horizonte	%	Pitagoras Teofilillandia	%	Pitágoras Niquelandia	%	Pitágoras Tucurui	%	Estadual Teofilandia	%	Municipal Teofilandia	%	Municipal Teofilandia	%	Estadual Niquelandia	%	Estadual Tucurui	%	Municipal Tucurui	%
número de alunos	3707,00		532,00		554,00		2398,00		512,00		828,00		846,00		1258,00		1775,00		902,00	
CUSTOS CAPITAL																				
custo prédio	3990214,00		572644,00		596325,00		2581207,00		165324,00		267361,00		273190,00		406208,00		569918,00		291255,00	
custo equipamento	62385,00	0,003	20811,00	0,01	21828,00	0,01	82593,00	0,01	7192,00	0,01	10926,00	0,01	7624,00	0,006	12359,00	0	13450,00	0,01	9605,00	
manutenção			1956,00						648,00								540,00			
10% préd+equipam	405259,90		59345,50		61815,30		266380,00		17251,60		27828,70		28081,40		41856,70		58336,80		30086,00	
total serv. capital	405259,90		61301,50		61815,30		266380,00		17899,60		27828,70		28081,40		41856,70		58876,80		30086,00	
total/aluno	109,32	0,19	115,23	0,18	111,58	0,19	111,08	0,24	34,96	0,29	33,61	0,17	33,19	0,24	33,27	0,16	33,17	0,37	33,35	0,32
MÃO DE OBRA																				
docente	77605,00		10954,00		8598,00		34524,00		2670,00		7183,00		4000,00		10938,00		4892,00		3024,00	
técnico	13923,00		5659,00		2811,00		8362,00						263,00		3247,00				951,00	
administrativo	26426,00		3643,00		7046,00		10789,00		911,00		4010,00		2172,00		2987,00		2724,00		1046,00	
total ano	1415448,00		243072,00		221460,00		644100,00		42972,00		134316,00		77220,00		206064,00		91392,00		60252,00	
total /aluno	381,83	0,68	456,90	0,70	399,75	0,69	268,60	0,58	83,93	0,70	162,22	0,82	91,28	0,66	163,80	0,77	51,49	0,58	66,80	0,64
GASTOS DIRETOS ALUNO																				
mat. didático/ano	73,00		83,00		67,00		80,00		1,60		1,60		13,00		15,00		4,70		3,00	
outros							6,00						1,30		2,00				1,00	
total	73,00	0,13	83,00	0,13	67,00		86,00	0,18	1,60	0,01	1,60	0,01	14,30	0,10	17,00	0,08	4,70	0,05	4,00	0,04
TOTAL CUSTOS DIRETC	381,83		456,90		399,75		268,60		83,93		162,22		91,28		163,80		51,49		66,80	
TOTAL CUSTOS ESCOL (=DIRETOS+CAPITAL)	491,15		572,13		511,33		379,68		118,89		195,83		124,47		197,08		84,66		100,15	
TOTAL CUSTOS SOCIA (=DIR+CAP+ALUN)	564,15	100	655,13	100	578,33	100	465,68	100	120,49	100	197,43	100	138,77		214,08	100	89,36	100	104,15	100
Docentes / Pessoal	0,66		0,54		0,47		0,64		0,75		0,64		0,62		0,64		0,64		0,60	
Admin. / Pessoal	0,12		0,28		0,15		0,16		0,00		0,00		0,04		0,19		0,00		0,19	
Técnico / Pessoal	0,22		0,18		0,38		0,20		0,25		0,36		0,34		0,17		0,36		0,21	

IPEA
Instituto de Pesquisa
Econômica Aplicada

Banco Mundial

PNUD
Programa das Nações
Unidas para o Desenvolvimento

SEMINÁRIO SOBRE EDUCAÇÃO, CRESCIMENTO E DESIGUALDADE NO BRASIL

SEMINAR ON EDUCATION, GROWTH AND INEQUALITY IN BRAZIL

Rio de Janeiro, Brasil

24 a 27 de Março de 1991

DO THE RETURNS TO SCHOOLING VARY ACROSS INDUSTRIES? EVIDENCE FROM SÃO PAULO, BRAZIL MANUFACTURING

Donald Robbins
Mari Minowa

documento 12

Do the Returns to Schooling Vary Across Industries?

Evidence from São Paulo, Brazil Manufacturing

February, 1991

Donald Robbins
Harvard Institute For International Development

and

Mari Minowa
Cornell University

Introduction

This paper examines evidence on variation in the earnings functions across industries for 1977 in São Paulo, Brazil. We focus on whether the gross returns to schooling vary across industries, and explore the structure of this variation. Building upon other work done for Brazil showing large, statistically significant interindustry wage differentials that are correlated to wage differentials in the United States(see Robbins 1989a), we enquire whether high-wage industries exhibit higher gross returns to schooling. Lastly, we examine whether gross returns to schooling vary systematically with industry characteristics. This work provides a potential linkage from the distribution of firms of different types, to the dispersion in the gross returns to schooling, to the distribution of income.

Research on the returns to schooling often incorporates implicit, but debatable, assumptions regarding the relationship between the structure of wages - by which we mean the pattern of the variation in earnings functions' parameters across firms - and the characteristics of firms in which workers are employed. Economists beliefs about these assumptions have changed through time. Before the nineteen sixties, economists such as Schlichter (1950), Reynolds (1951), Lester (1948) and Lewis (1956) believed that firm characteristics affected the structure of earnings. During the 1960s,in the aftermath of work by Gary Becker (1964) and again in 1974 with the work of Jacob Mincer, long-run labor supply became the principle focus of empirical labor market research, while firms were regarded as price-takers and demand-side (firm and industry) characteristics were largely ignored. Thus, economists typically assumed that the equilibrium wage structure is constant across firms. More recently, theory and evidence has increasingly emphasized the potential variation of earnings functions across firms (Lazear 1979, Yellen 1984, Stiglitz 1986, Katz and Summers 1989). However, despite a growing belief in the importance of variation in earnings functions across firms, little attention has been directed to the possibility that gross returns to schooling vary over firms. The possibility that earnings functions and the gross returns to schooling vary over firms merits closer examination.

There are two major classes of theories leading to varying earnings functions across firms. The first class consists of competitive theories and includes models of specific human capital (Becker 1964, Hashimoto 1981, Carmichael 1983), and principal-agent models of incentive schemes (Lazear 1979). These theories describe how some firms may find it profit maximizing to pay rising wages which start below the competitive rate, then rise with seniority to eventually exceed the competitive wage. Firms would adopt these strategies in the absence of credible contracts to lower turnover or increase effort. These wage schedules are market-clearing, since they lead to constant average wages across firms. The second class consists of non-competitive theories, such as Efficiency Wages, rent-sharing, Insider-Outsider theories, [Yellen 1984, Stiglitz 1986, Katz and Summers 1989). In these non-competitive theories some firms find it profit maximizing to pay higher wages than the market clearing wage. Wages in excess of workers' opportunity costs are paid for varied reasons; these reasons include reducing shirking or turnover and raising the quality of the applicant pool. Matching theories < as introduced by Roy (1951) and more recently discussed by Heckman and Sedlacek (1990) offer a third approach. In this paper, we shall concentrate upon the first two broad classes of theories: competitive and non-competitive theories.

Empirical research on non-competitive theories has principally focused on testing for the existence of wage differentials across firms. These models have been empirically specified as simple modifications of the traditional semi-log specification. The modification is simply to let the intercept term vary across firms. In this specification, the coefficients on worker characteristics are modeled as constant over firms or industries. However, it is possible that estimated wage differentials arise from varying returns to human capital variables, such as schooling, rather than exclusively from shifts in the intercept terms of the log-wage earnings functions. To our knowledge, the possibility that variation in the gross returns to schooling might contribute to the observed differentials has not been studied.

This paper uses a large cross-sectional data set of workers in private, São Paulo manufacturing firms to empirically explore whether earnings functions vary across industries. We explore variation at the industry level, to keep sample sizes large and sampling error in estimated parameters small. Analysis at the industry level also relates directly to the aforementioned work on Brazilian interindustry wage differentials. We focus upon whether there is statistically and economically significant variation in the gross returns to schooling across industries, and whether there is structure to that variation. Specifically, we examine whether gross returns to schooling vary consistently with interindustry wage differentials, and in relation to industry characteristics.

The rest of the paper is organized as follows. In section I, we examine the specification of the earnings function. We argue that non-competitive theories of variable earnings functions across firms may be mis-specified by the simple semi-log earnings function with variable intercepts. Moreover, we conclude that non-competitive theories suggest that the gross returns to schooling may vary across firms and wage differentials may arise principally from this variation, rather than simply through the variation in shift coefficients. These considerations lead to our use of a flexible specification, where earnings functions may freely vary over firms.

Section II presents the data and key results. We first examine the evidence for variation in earnings functions across industries. Where log wages are specified as a function of schooling alone, or schooling and schooling-squared, we find considerable variation in estimated parameters, with homogeneity of earnings functions very strongly rejected for both the conventional F-test and for the more rigorous Swamy test. Second, we examine variation in estimated gross returns to schooling across industries, finding this variation to be large. Third, we examine the structure of variation in earnings functions over industries. The evidence supports the hypothesis that interindustry wage differentials arise from variation in gross returns to schooling, rather than from intercept shifts. Evidence also suggests that firms have seniority wage structures which are largely orthogonal to interindustry wage differentials and, by themselves, are market clearing. Finally, there is weak support that seniority wage structures may sometimes partly solve the agency problem in eliciting effort, allowing firms to avoid wage premia. Fourthly, we briefly examine the relationship between gross returns to schooling and industry characteristics. We find that highly multinational industries pay higher returns to schooling, while concentrated industries pay lower returns to schooling.

Section III summarizes and discusses the findings.

I. Variation in Earnings Functions and Gross Returns to Schooling Across Firms

In the 1970's it became common to estimate a hedonic wage equation where the log of an individual worker's wage is a linear function of observed personal characteristics (see Becker 1964, Mincer 1974):

(1) $\ln W_i = \alpha + PC_i'\beta + e_i.$

The parameters of this equation are regarded as varying over time and space, but constant across firms in one place and moment(see Griliches 1977). This constancy of the parameters across firms reflected the focus on long-run labor supply and the belief that firms were price-takers. Thus, in this formulation, the vector or parameters, β, is constant over firms.

As discussed above, however, some competitive theories of wage determination - including the specific human capital model - and non-competitive theories lead to specifications where the parameters of earnings functions vary across firms, in one place and time. The competitive models lead to an alternative to specification (2) where returns to time in the firm, tenure, vary over firms. At the same time, the average wage over a worker's spell in the firm is constant across firms. This leads to firm-varying returns to tenure and a corresponding constraint on the intercept:

$$(2) \qquad \ln W_i = \alpha_f + \beta_S \cdot S_i + \beta_{Tf} \cdot T_i + Z_i'\delta + v_i,$$

where we have broken the vector PC of (1) into schooling, S, tenure, T, and other personal characteristics, Z. The parameters β_{Tf} and α_f in (3) will vary over firms; also, they will negatively covary to maintain a constant average wage over the worker's spell, ceteris paribus. Thus, while the specification (2) conveys a break in the strict link between marginal productivity and the spot wage, it retains constancy of the schooling parameter, β.

More recently, non-competitive models have also led to firm-varying specifications of the earnings function. These theoretical models of non-competitive labor market are typically formulated in terms of homogeneous workers, and aim to show that wage levels may differ across firms. It seems intuitive, therefore, that empirical testing of non-competitive theories commonly have adopted a simple modification of equation (1) that changes the intercept from being constant over firms to being a firm-specific shift effect:

$$(3) \qquad \ln W_i = (\alpha + \delta_f) + PC_i'\beta + e_i.$$

Here α is the intercept for competitive firms and $(\alpha + \delta_t)$, $\delta_t > 0$, is the intercept for non-competitive firms.

In many respects the specification in (3) is a logical generalization of the theory of non-competitive wage determination, formulated in terms of homogeneous workers, to the empirical world of heterogeneous workers. To see this, note that perhaps the simplest model of worker heterogeneity uses the Becker-Mincer concept of human capital. Here workers posses an underlying homogeneous human capital, HC (see Willis 1986) for discussion of homogeneous versus heterogeneous human capital). In equilibrium one price is paid for this human capital, so that in equation (4) the wage is the rental price of human capital, A, times the level of human capital. In equation (5) we model a production function of human capital that is exponential in inputs, X. Equations (4) and (5) combining to give us equation (6):

$$(4) \qquad W = A \cdot HC.$$

$$(5) \qquad HC = e^{\Sigma X \alpha},$$

$$\text{or} \quad (6) \qquad W = A \cdot e^{\Sigma X \alpha}.$$

Where 'A' = $\exp(\alpha)$. By adding an exponential stochastic term and taking logs, (6) leads to equation (1), above.

Since non-competitive theories of wage determination are really theories of firm heterogeneity, where competitive firms coexist with non-competitive firms, their empirical specifications try to model a diversity of firms: both competitive and non-competitive. For example, in the shirking model of Efficiency Wages, firms differ in their technology, and therefore in their monitoring costs and the level of wages they must pay to induce worker effort. To specify a labor market with a both non-competitive and competitive firms, leading to equation (3), equation (6) can be modified so the rental price of human capital varies over firms. Competitive firms pay the competitive rental costs of human capital, $A = \exp(\alpha)$, while others pay a higher rate, $A_t = \exp(\alpha + \delta_t)$. Taking logs leads to the specification typically used to study non-competitive wage determination, equation (3).

Equation (3) appears to satisfies a basic property we would expect of non-competitive firms: that the wage premia paid by a non-competitive firm would be proportional to the worker's opportunity costs. To achieve this, while the coefficients on the vector of personal characteristics is constant in (3), the shifting intercept implies changing derivatives of the level of the wage with respect to individual personal characteristic variables, across firms.

Thus, specification (3) seems to be a logical extension of the theoretic models formulated in terms of homogeneous workers, and appears to maintain proportionality between the wage premia and workers' opportunity costs. However, it may not be the best specification. First, it is not the only way for wage premia to be proportional to workers' opportunity costs. An alternative would be to allow the slope coefficients to vary. Second, despite appearances, equation (3) may imply that wage premia are not proportional to workers' opportunity costs. This inconsistency with non-competitive theory arises because the vector of worker characteristics, PC, in (3) may include general and specific human capital, while workers' opportunity costs are proportional only to general human capital. Non-competitive firms pay wage premia according to workers' opportunity costs. But specification (3) erroneously implies that wage premia increase in equal proportion to both general and specific human capital.

A useful alternative specification reflecting this asymmetry between specific human capital and general human capital for the non-competitive models might be to allow slope coefficients on the PC variables to vary over firms:

$$(7)\ \ln W = \alpha + gHC_i'\beta_f + sHC_i \cdot \gamma + e_i,$$

where gHC represents general human capital and sHC represents specific human capital. The specification in (7) reflects the idea that a non-competitive firm would pay premia in proportion to the general human capital of the workers in the firm. A more realistic empirical specification allowing for seniority wage profiles may require variation in the parameters of the specific human capital variables, as well. Seniority profiles associated with specific human capital are potentially compatible with non-competitive wage premia. The non-competitive firm would first set a schedule of wages, setting wages as mark-ups over the opportunity cots of individual workers. Then, from this schedule of wages marked-up over the opportunity costs of workers, the firm could vary the wages of workers with the same level of general human capital according to their seniority. As we discussed above, specific human capital theory suggests that seniority wage profiles vary across firms, that these profiles tend to be positively sloped, and that the intercept decreases as the slope rises to keep a given worker's average expected wage within the firm unchanged. On average over a worker's completed spell, the worker would simply receive the non-competitive mark-up.

To accommodate non-competitive models and seniority wage profiles, we broaden equation (7), allowing all parameters to vary across firms. This is a hybrid specification of (2) and (7):

$$\text{(8) } \ln W = \alpha_f + gHC_i'\beta_f + sHC_i \cdot \gamma_f + e_i.$$

Note that, here, non-competitive high-wage firms would pay higher gross returns to general human capital.

Equation (8) allows seniority profiles to vary separately from the wage premia of non-competitive firms.

Structure of Earnings Functions Variation

The foregoing discussion suggests that returns to general human capital may vary across firms(industries), and that they will positively covary with firm(industry) level predicted wages. Thus, unlike standard specifications of inter-firm wage differentials, the non-competitive wage differentials will arise from changing gross returns to general human capital. Since the clearest measure of general human capital is schooling, this suggests a positive correlation between gross returns to schooling and the wage level for identical workers. To the extent that experience also reflects general human capital, gross returns to experience would also positively covary with the wage level for identical workers. However, since experience is generally proxied by age - schooling - 6, the experience variable may capture other effects than general human capital.

Our focus is the variation and pattern of gross returns to schooling over firms, but some remarks on seniority profiles are useful. It is not clear what relationship we should expect between high-wage non-competitive firms and seniority wage profiles. On the other hand, while Akerlof and Katz (1986) have shown that seniority wage profiles of the sort discussed by Lazear (1979) cannot completely substitute for up-front bonds to clear the market, these profiles can partially substitute for such bonds. Hence, these seniority profiles can lower the wage premium needed to induce effort. And, to the extent that observed seniority wage profiles are partial solutions to the principal-agency effort elicitation problem, we might expect seniority profiles to be steeper in low-wage firms. Since in practice we proxy specific human capital by tenure, we might expect that the gross returns to tenure will negatively covary with the wage premium.

Below we summarize the expected wage structure. Predicted wages, $\hat{W}$, are the wages for identical workers in a given firm, and correlations of estimated parameters with predicted wages give us information about the patterns of wage structure leading to wage premia. First, we expect a positive covariance between the derivative of log(wage) with respect to schooling and the predicted log wage, and a weak positive covariance between the derivative of log(wage) with respect to experience and the predicted log wage:

$$\rho(\partial \ln W/\partial S, \ln\hat{W}) > 0$$

$$\rho(\partial \ln W/\partial X, \ln\hat{W}) \geq 0$$

Another way to think of these relations is that the inter-firm wage differentials arise largely from returns to schooling that vary across firms, rather than from vertical shifts of the earnings function. Second, seniority profiles may negatively covary with predicted wages:

$$\rho(\partial \ln W/\partial T, \ln\hat{W}) \leq 0.$$

II. Data and Empirical Results

The data used to estimate earnings functions is a sample of workers in private manufacturing firms in greater São Paulo, from the 1977 RAIS (Relação Anual de Informações Sociais), a firm-reported census of workers in formal sector firms. Firms with more than 50 workers were selected and workers were randomly sampled from within these firms. We obtained 77,691 observations in the final sample of male workers. Means and standard deviations for workers across all industries of key variables are reported in Appendix B. The mean sample size for individual industries was 3884. Except for the leather industry, the industry with the fewest observations was furniture and had 639 workers, while the industry with the most observations was construction, with 28,631 workers. The leather industry had only 158 workers, and was dropped from most of analysis.

Equation (8) in Section I specifies the earnings function as varying freely across firms. We estimated equation (8) at the two-digit industry level, rather than the firm level for comparability with interindustry wage differentials results and to minimize sampling error in the parameter estimates. The following variables were included as regressors in all estimated earnings functions, where the dependant variable was log(wages):

S: years of schooling

SSQ: years of schooling squared [This term was used only in the specification of quadratic schooling.]

X: years of experience, proxied by Age - schooling - 6

XSQ: years of experience squared

T: years of tenure in the firm

TSQ: years of tenure squared

T*S: the interaction term between tenure and schooling

X*S: the interaction term between experience and schooling

T*X: the interaction term between tenure and experience

NODEPS: number of dependents

DMIG1: equals 1 if a recent migrant in two stages, 0 otherwise

DMIG2: equals 1 if worker is a recent migrant in one stage, 0 otherwise

DMIG3: equals 1 if worker is a non-recent migrant, 0 otherwise

IMIG: equals 1 if worker is an immigrant, 0 otherwise

DCIVIL1: equals 1 if married, 0 otherwise

LOGNOCU: log of the number of jobs in the firm in Greater São Paulo

LOGNOBR: log of the number of jobs in the firm throughout Brazil

Here we focus upon gross returns to three variables: years of schooling, S, years of experience, X, and years of tenure in the firm, T.

A preliminary examination of estimated earnings functions that were segmented by years of schooling suggested that estimating the earnings functions with schooling squared was superior to the specification of log(wages) as linear function of schooling. The results of the wage equations that were quadratic in schooling closely correspond to the results from the segmented regressions that were linear in schooling. Appendix C presents the segmented regressions and Appendix D presents the industry level regressions that are linear in schooling. Below we present selected results from the "linear in schooling" specification, focusing on the results from the "quadratic in schooling" specification. First we discuss the evidence on the heterogeneity of the earnings functions and then turn to the structure of the earnings functions.

Heterogeneity of the Wage Equation

There is large variation in the estimated coefficients across industries for most variables. Estimated parameters for key variables in the industry level earnings functions with linear schooling are presented in Appendix D. We can see this dispersion in Figure 1, where the estimated coefficients on the schooling variable are plotted against their t-statistics. Note the high t-statistics and that higher t-statistics do not lead to convergence in the estimated parameters. Table II.1 reports the corresponding estimated derivatives of log wage with respect to schooling, all statistically significant. The mean of the estimated derivatives is .172, and the standard deviation is .025. The smallest gross returns to schooling, .11, is for the tobacco industry, while the largest, .219 is for the pharmaceutical industry. Thus, gross estimated returns to schooling vary nearly one hundred percent from the smallest to largest estimated derivative. Eliminating the tobacco industry, the largest estimated derivative is still fifty percent larger than the smallest derivative. The wide spread in estimated gross returns to schooling is seen in Figure 2, where the estimated derivatives are plotted by industry.

The F statistic for differences in earnings functions across industries is 13, strongly rejecting the null hypothesis of homogeneity according to the standard F test. For large sample sizes, however, Leamer (1978) argues that the conventional critical values for the F test are too low, and should grow with the sample size. We calculated Leamer's Bayesian critical value of the F test. The resulting value is a little over 11, still rejecting the null hypothesis of homogeneity. Later, in the specification that is quadratic in schooling, we will also present results for an alternative test of homogeneity.

Turning to the results for the specification of the earnings function with quadratic schooling, panel A of Table II.2 presents the averages and observed dispersions of the estimated coefficients for key variables and their interactions. The ratio of the standard deviation to the mean for the estimated parameters varies from roughly .25 to 2. This ratio is highest for the coefficients on schooling, the interaction between tenure and schooling, and the interaction between tenure and experience. For comparison we also list, in panel B, similar statistics for four other variables - which are not regression coefficients.

Table II.3 presents estimated gross returns to schooling, tenure and experience - the derivatives of the estimated log wage equation with respect to schooling, tenure and experience - with their means and standard deviations of these estimates at the bottom. The dispersion in estimated gross returns is large. For all three variables the ratio of largest to smallest derivative is roughly four. Another measure of dispersion is the ratio of the standard deviations to the average estimates. This ratio is roughly one third for all three estimated derivatives . In Figure 3, which plots the estimated derivatives of log wage with respect to schooling by industry, this considerable dispersion is apparent. In panel C of Table II.2 we see that while the mean and standard deviation of the observed log wage are 7.67 and 0.38 respectively, the corresponding values for the estimated log wage evaluated at the mean worker characteristics are 7.64 and 0.21. Thus, roughly fifty-five percent of the variation in the log wage across industries may be attributed to variation in the coefficients of wage equation, rather than the variation in mean worker characteristics.

The F statistic for the test of homogeneity of earnings functions across industries equals 13.2. As in the linear specification in schooling, the null-hypothesis of homogeneity is strongly rejected, using the conventional critical value as well as Leamer's Bayesian critical value.

The Swamy test is an alternative, more stringent test of homogeneity, related to the Empirical Bayes techniques in that it reflects the notion that individual industry estimates may not be drawn from truly independent populations. (Swamy 1970, 1971) This statistic essentially compares weighted estimates of the industry earnings functions parameters to their overall mean. The weighted parameter estimates are standard-error weighted averages between the group mean and the unconstrained industry level estimates, resulting in a "shrinkage" of industry parameter estimates to the overall mean that increases with the sampling error of the particular industry parameter estimates. This shrinkage necessarily leads to less disperse estimates. The Swamy test of homogeneity asks whether these less disperse parameter estimates are indeed different from their overall mean. The Swamy statistic for our data is highly statistically significant[1], thus strongly rejecting the null hypothesis of homogeneity of the earnings functions.

In conclusion, we find that earnings functions differ over industries. This variation differs from the common specification used in estimating interindustry wage differentials, where only the intercept term in the log wage earnings functions varies over industries. Instead, there is wide dispersion of the estimated slope coefficients across industries. In particular, the dispersion in gross returns to schooling appear to be economically quite large and statistically significant.

The Structure of Variation in Earnings Functions

Having found that earnings functions differ across industries, including wide variation in estimated slope coefficients, now we examine whether there is consistent structure to this variation. In particular we ask whether wage differentials across industries arise from systematic changes in slope coefficients or, as in the interindustry wage differential literature, simply from shifts in the intercept of the log wage earnings functions. We find that the interindustry wage differentials arise largely from differing gross returns to schooling. The principal variables in the estimated earnings functions explaining log wages are schooling, experience, and tenure. Of these variables, only the derivative of log wage with respect to schooling is positively correlated with the predicted wage. This correlation is .49, and is reported in the first column in row one of Table II.4. Continuing to columns two and three of that table, we see that variation in the gross returns to experience or to tenure are not associated with wage premia: on the contrary, the derivative of log wage with respect to tenure is strongly negative, at (-.46), and with respect to experience is roughly zero, at (-.04). Finally, in contrast with the common specification of the non-competitive interindustry wage differentials where differentials arise from shifts in the intercept, the correlation of the predicted wage with the intercept is negative and large, equalling (-.42).

These results correspond to the predictions of Section I, where we argued that the common specification of non-competitive theories - with constant slope coefficients across firms and differing intercepts - was likely to be incorrect. High-wage firms would pay wage premia proportional to only general human capital of workers. Since years of schooling strongly reflects general human capital, the correlation between the gross returns to schooling and the predicted log wage should be positive. And since measured experience is more questionable measure of general human capital, the correlation between gross returns to experience and the predicted log wage ought to be roughly zero.

Further, the seniority profile generally will not correspond to general human capital. Rather, seniority profiles more likely correspond to market-clearing incentive schemes related to effort elicitation or specific human capital. Therefore we predicted that the correlation between the predicted log wage and the derivative of the log wage with respect to tenure would not be positive. This correlation could, however, be negative if seniority profiles generally constitute an implicit up front bond which partially offset the wage premia otherwise required to elicit effort.

To see if the structure of earnings functions varied importantly across workers with different levels of schooling, tenure and experience, we examined the correlations between the predicted log wage and the schooling, tenure and experience derivatives evaluated for differing levels of the arguments. Rows two through five of Table II.3 report the estimated derivatives of the log wage with respect to schooling, tenure and experience evaluated at mean worker characteristics for workers with two, four, eight and thirteen years of schooling. The overall pattern is very close to the correlations discussed above using the derivatives evaluated at mean characteristics for the entire sample. The correlation of the estimated gross returns to schooling with the predicted log wage is positive for all but the lowest level of schooling, and highest for average workers with eight years of schooling. This may be because education among unskilled workers matters less to their opportunity costs, and hence wage premia, than for more highly educated workers. The correlation of the experience derivative with the predicted log wage derivative increases with schooling level, though the corresponding t-statistics are not large and the result is difficult to interpret. This could be because job experience leads to greater general human capital and wage premia for more highly educated workers. However, this result could also be interpreted as high wage firms prefer younger workers in production jobs, and older workers in skilled white-collar jobs. As with the derivative evaluated at the mean characteristics for the entire sample, the correlation of the tenure derivative with the predicted log wage is negative for all schooling levels.

The principal finding here is that interindustry wage differentials arise largely from differing gross returns to schooling. To determine attributes of industries with high gross returns to schooling, we performed second-stage regressions where the estimated derivatives of log wage with respect to schooling were regressed onto industry concentration(Concentration), the percent of an industry's production by multinationals(Multinational), the percent of industry production bought by the top thirteen percent of the income distribution(Demand Concentration), and dummies for technology - Advanced and Traditional. This was done unweighted and weighted by the standard deviation of the estimated schooling derivatives. Between sixty-four and seventy-two percent of the variation in gross returns to schooling is explained by these variables. We find that gross returns to schooling are lower in concentrated industries, lower in industries with high Demand Concentration, and lower in Traditional industries. Gross returns are higher in multinational and Advanced industries. (see Appendix E). With the exception of Concentration, these results parallel other work of ours examining the industry correlates of interindustry wage differentials. (Robbins 1989b)

III. Discussion and Summary

We have found that earnings functions differ over industries and that this variation differs from the common specification used in estimating interindustry wage differentials, where only the intercept term in the log wage earnings functions varies over industries. Instead, there is wide dispersion of the estimated slope coefficients across industries. The dispersion in gross returns to schooling appear to be economically quite large and statistically significant. We find that interindustry wage differentials arise largely from differing gross returns to schooling, and that these gross returns to schooling are highest in industries with advanced technology, highly multinational, with high degrees of demand concentration, and with low levels of concentration.

The pattern of income distribution is affected by the level of gross returns to schooling as well as by wage differentials across firms or industries. For a given distribution of education, higher gross returns to schooling will tend to widen the dispersion of incomes. A wide dispersion in firm types that leads to a greater variance in both interindustry wage differentials and, relatedly, a greater variance in the gross returns to schooling, will therefore engender still greater variance in wage income.

This suggests that the emergence of industrial diversity, with the growth of industries high in technological sophistication, and increasingly multinational contributed to the observed widening of the Brazilian income distribution, beginning in the 1960s.

Policy options designed to mitigate these sources of widening income distribution would need to alter the distribution of firm types. Policies altering income distribution and therefore consumption patterns are candidates. However, since the non-competitive theories of differences in wage structure across industries derive from firms' concerns to maximize worker productivity, any policy designed to alter the mix of firm types would need to address the potential trade-offs between wage premia and returns to schooling versus efficiency. To date, only rudimentary progress has been made in our understanding of such trade-offs (Katz and Summers (1989), Bulow and Summers (1986)), and this is an important area for future research.

1. The chi-square distributed Swamy statistic with degree of freedom 323 is 24,700, versus a critical value of 366 for 95 percent and 384 for 99 percent confidence.

Table II.1 Estimated Gross Returns to Schooling
-For Regression with Linear Schooling Specification-

Industry	$\partial \ln W / \partial S$
Metallurgy	0.185
Machinery	0.196
Electronics	0.169
Transport	0.148
Wood	0.154
Furniture	0.165
Paper	0.186
Rubber	0.153
Leather	0.186
Chemicals	0.191
Pharmaceutical	0.219
Perfumes	0.204
Plastics	0.197
Textiles	0.148
Clothing	0.175
Food	0.173
Beverages	0.178
Tobacco	0.111
Editorial and Graphics	0.154
Civil Construction	0.151

Note: All derivatives are statistically significant.

Table II.2 Variability of the Estimated Coefficients and Derivatives across Industries
-For Regression with Quadratic Schooling Specification-

	Average	Observed Dispersion (Standard Deviation)
A. Variability of coefficients on the variables.		
Schooling	0.0365	0.0796
Schooling^2	0.0084	0.0047
Tenure	0.2188	0.0589
Tenure^2	-0.0067	0.0093
Experience	0.0666	0.0244
Experience^2	-0.0010	0.0005
Tenure*schooling	-0.0010	0.0026
Experience*schooling	0.0010	0.0011
Tenure*experience	-0.0027	0.0015
B. Variability of industry means of schooling, tenure, experience and log wage.		
Schooling	5.53	1.29
Tenure	2.41	0.73
Experience	14.10	1.70
Log wage	7.67	0.38
C. Variability of estimated log wage evaluated at the mean worker characteristics.		
Estimated log wage	7.64	0.21

Table II.3 Estimated Gross Returns to Schooling, Tenure, and Experience
- For Regression with Quadratic Schooling Specification -

Industry	$\partial \ln W / \partial S$	$\partial \ln W / \partial T$	$\partial \ln W / \partial X$	R^2	Sample Size
Metallurgy	0.096	0.138	0.042	0.5319	4946
Machinery	0.138	0.220	0.054	0.4969	4300
Electronics	0.115	0.119	0.046	0.5799	4212
Transport	0.130	0.129	0.043	0.5095	7885
Wood	0.098	0.155	0.055	0.5036	790
Furniture	0.138	0.159	0.064	0.6100	639
Paper	0.161	0.102	0.049	0.5785	1533
Rubber	0.102	0.060	0.038	0.4228	1713
Leather	0.150	0.141	0.097	0.6705	158
Chemicals	0.103	0.174	0.055	0.5422	3246
Pharmaceutical	0.230	0.145	0.067	0.5514	1878
Perfumes	0.178	0.117	0.049	0.7386	643
Plastics	0.160	0.174	0.054	0.5645	2933
Textiles	0.100	0.163	0.050	0.4757	3654
Clothing	0.103	0.138	0.034	0.5305	1196
Food	0.147	0.123	0.047	0.5708	4643
Beverages	0.134	0.130	0.030	0.5511	1499
Tobacco	0.120	0.154	0.025	0.4970	1034
Editorial and Graphics	0.055	0.200	0.056	0.4243	2160
Civil construction	0.119	0.140	0.034	0.2241	28631
Mean	0.129	0.144	0.050		
Standard Deviation	0.037	0.034	0.016		

Note: All derivatives are statistically significant.

Table II.4 Correlation between Estimated log Wages and Estimated Returns to Schooling, Tenure, and Experience -Quadratic Schooling Specification-

	$\partial \ln W/\partial S$	$\partial \ln W/\partial T$	$\partial \ln W/\partial X$
(A) Evaluated at the Mean Worker Characteristics for the Entire Sample			
Estimated log Wage	0.4886 (2.3759)	-0.4625 (-2.2132)	-0.0423 (-0.1796)
(B) Evaluated at mean worker characteristics for workers with 2, 4, 8 and 13 years of schooling			
Estimated log Wage for average Workers with:			
Schooling=2	-0.1642 (-0.7062)	-0.4989 (-2.4423)	-0.2664 (-1.1726)
Schooling=4	0.3483 (1.5764)	-0.4714 (-2.2678)	-0.1912 (-0.8264)
Schooling=8	0.5687 (2.9333)	-0.4120 (-1.9184)	0.0524 (0.2226)
Schooling=13	0.3651 (1.6638)	-0.1239 (-0.5298)	0.2496 (1.0936)

Note: Number in parentheses are t-statistics for the correlations. In calculating these correlation, industry 19 was dropped because of its very small sample size.

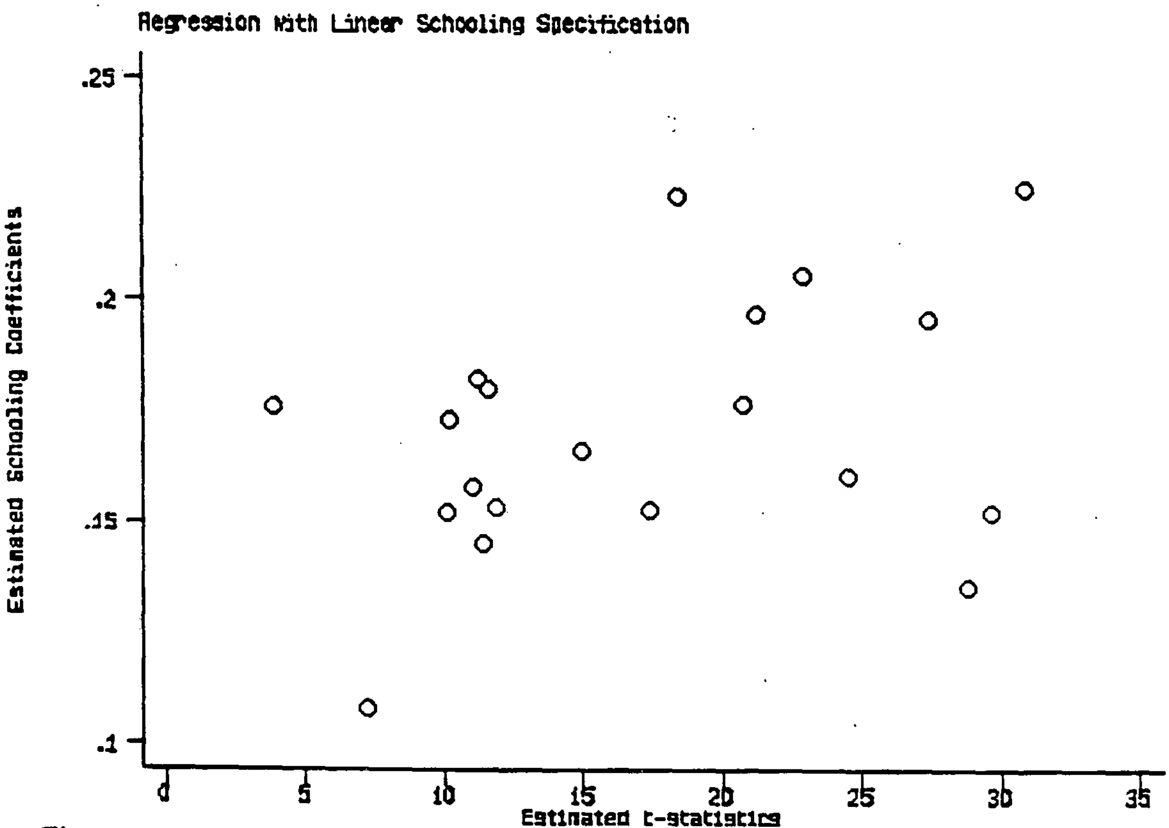


Figure 1. Estimated Schooling Coefficients versus Their t-statistics

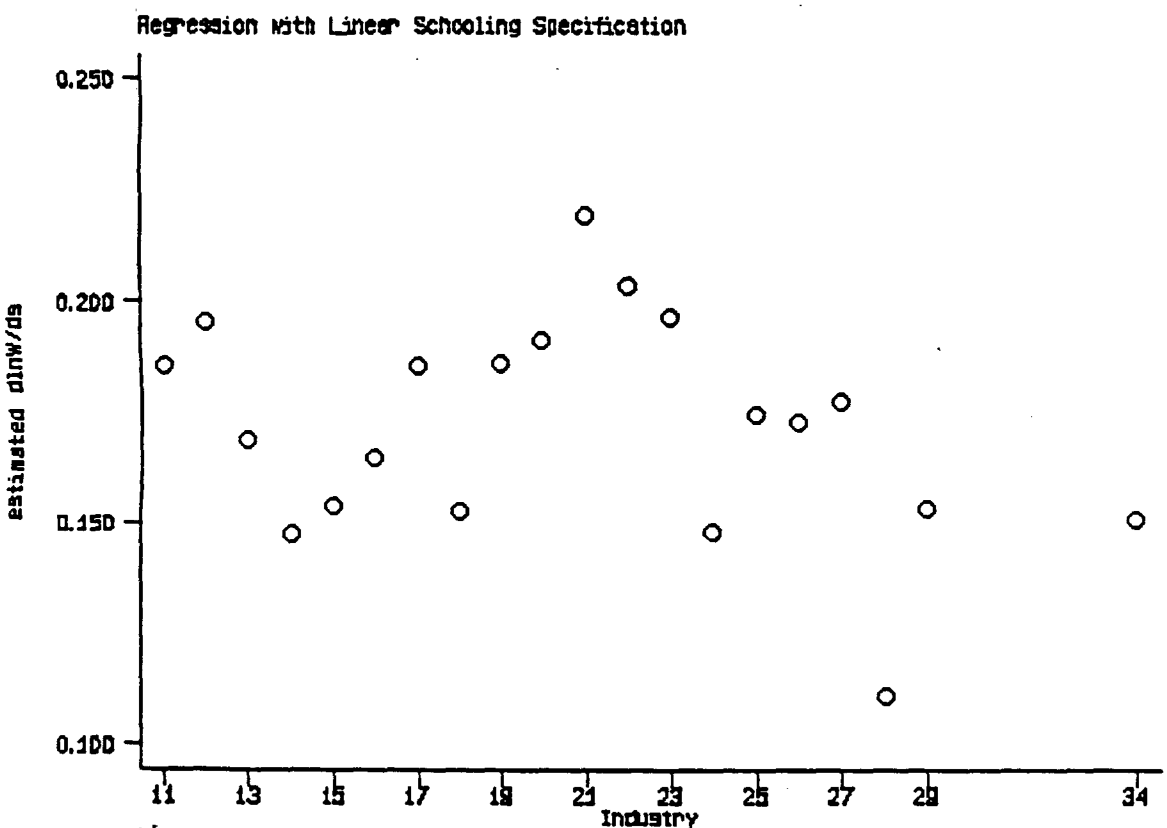


Figure 2. Estimated Derivatives of lnW with respect to Schooling

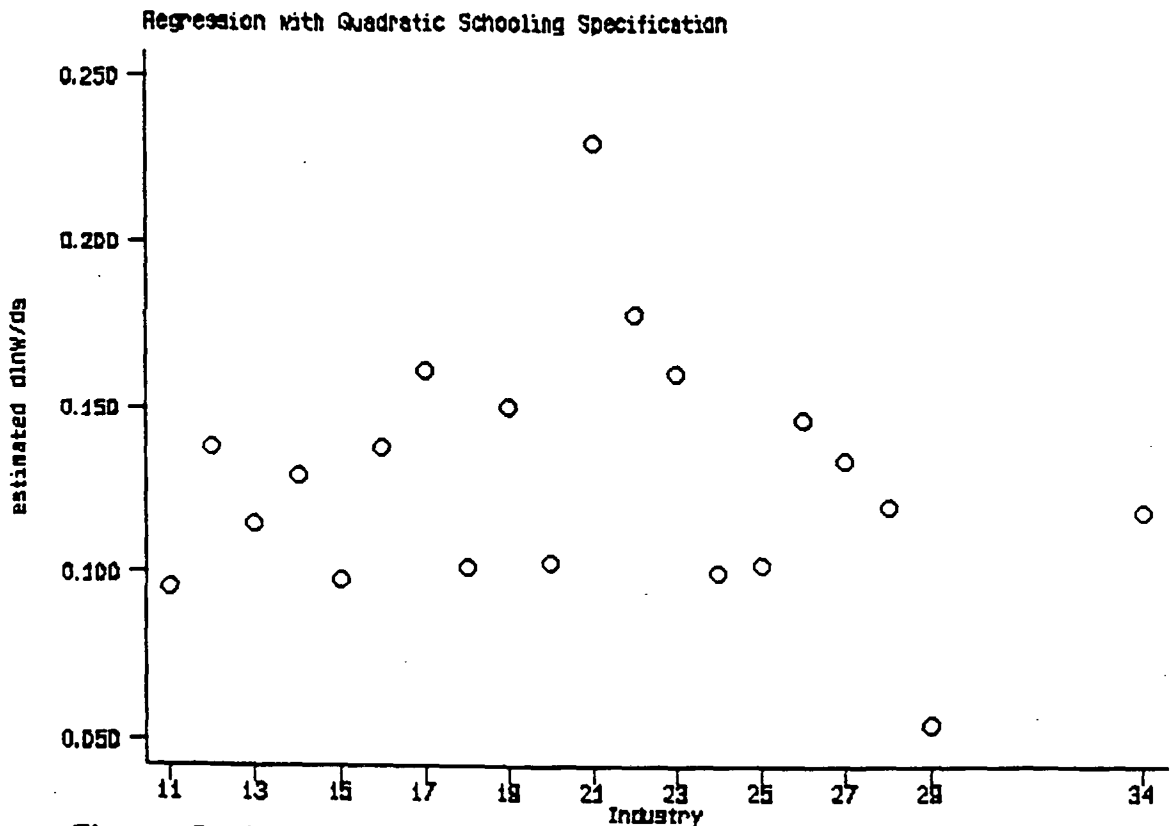


Figure 3. Estimated Derivatives of lnW with respect to Schooling

Appendix A List of Variables

lnW:	natural log of wage
S:	years of schooling
SSQ:	years of schooling squared
T:	years of tenure at the current firm
TSQ:	years of tenure squared
X:	years of labor market experience
XSQ:	years of experience squared
T*S:	interaction between tenure and schooling
X*S:	interaction between experience and schooling
T*X:	interaction between tenure and experience
NODEPS:	number of dependents
DMIG1:	1 if worker is a recent migrant in two stages 0 otherwise
DMIG2:	1 if worker is a recent migrant in one stage 0 otherwise
DMIG3:	1 if worker is a non-recent migrant 0 otherwise
IMIG:	1 if worker is an immigrant 0 otherwise
DCIVIL1:	1 if married 0 otherwise
LOGNOCU:	log of the number of jobs in the firm
LOGNOBR:	log of the number of jobs in the firm on level of Brazil

Appendix B Mean and Standard Deviation of Variables

Industry	lnW	S	T	X	NODEPS	DMIG1	DMIG2	DMIG3	IMIG	DCIVIL1	LOGNOCU	LOGNOBR
All	7.61	4.73	2.02	15.78	1.50	0.03	0.05	0.40	0.02	0.54	6.79	7.87
	(1.09)	(3.41)	(3.75)	(8.98)	(4.86)	(0.16)	(0.22)	(0.49)	(0.15)	(0.50)	(1.77)	(1.50)
11	7.74	5.85	3.59	13.03	4.41	0.04	0.03	0.32	0.03	0.53	5.98	7.88
	(1.19)	(3.62)	(6.30)	(8.11)	(17.67)	(0.20)	(0.16)	(0.47)	(0.16)	(0.50)	(1.76)	(1.53)
12	7.59	5.52	2.00	13.91	1.32	0.01	0.02	0.38	0.03	0.55	6.57	7.09
	(1.24)	(3.50)	(3.02)	(8.65)	(1.80)	(0.11)	(0.16)	(0.49)	(0.17)	(0.50)	(1.28)	(1.02)
13	8.11	6.71	3.05	13.16	1.45	0.01	0.03	0.33	0.06	0.50	6.84	8.08
	(1.13)	(3.80)	(3.97)	(8.48)	(1.70)	(0.12)	(0.16)	(0.47)	(0.24)	(0.50)	(1.75)	(1.31)
14	8.05	5.53	3.08	14.36	1.58	0.01	0.01	0.37	0.04	0.48	7.54	8.84
	(0.96)	(3.28)	(3.60)	(8.08)	(1.80)	(0.10)	(0.10)	(0.48)	(0.19)	(0.50)	(2.96)	(1.75)
15	7.41	4.94	1.96	13.36	1.18	0.00	0.01	0.37	0.02	0.57	5.83	6.84
	(0.96)	(3.25)	(3.24)	(8.41)	(1.72)	(0.06)	(0.11)	(0.48)	(0.13)	(0.49)	(1.32)	(1.09)
16	7.80	5.58	2.06	14.71	1.01	0.00	0.01	0.49	0.03	0.56	6.06	6.23
	(1.02)	(3.33)	(2.84)	(9.28)	(1.52)	(0.07)	(0.09)	(0.50)	(0.17)	(0.50)	(1.18)	(0.94)
17	7.71	5.13	2.29	14.62	1.30	0.01	0.02	0.37	0.03	0.51	5.30	6.38
	(1.03)	(3.56)	(3.90)	(8.65)	(1.72)	(0.10)	(0.13)	(0.48)	(0.17)	(0.50)	(1.35)	(1.10)
18	7.63	4.44	2.64	15.40	1.06	0.01	0.06	0.42	0.02	0.55	6.36	7.03
	(1.05)	(3.02)	(4.46)	(8.36)	(1.66)	(0.07)	(0.24)	(0.49)	(0.14)	(0.50)	(2.00)	(1.21)
19	6.71	4.89	1.11	9.79	0.89	0.00	0.01	0.23	0.02	0.70	4.61	5.04
	(1.10)	(2.87)	(1.64)	(9.68)	(1.72)	(0.00)	(0.08)	(0.42)	(0.14)	(0.46)	(1.00)	(0.41)
20	7.88	6.42	3.08	13.38	1.37	0.02	0.05	0.37	0.03	0.50	5.72	6.67
	(1.36)	(4.05)	(5.08)	(8.74)	(1.78)	(0.15)	(0.22)	(0.48)	(0.17)	(0.50)	(1.38)	(1.50)
21	8.56	9.09	3.27	13.07	1.73	0.03	0.06	0.33	0.06	0.38	5.72	6.25
	(1.30)	(3.86)	(4.85)	(8.53)	(1.74)	(0.17)	(0.23)	(0.47)	(0.23)	(0.48)	(2.04)	(1.90)
22	8.21	7.04	3.23	13.75	1.42	0.02	0.03	0.35	0.07	0.46	5.83	7.12
	(1.27)	(4.41)	(5.11)	(9.33)	(1.67)	(0.12)	(0.17)	(0.48)	(0.25)	(0.50)	(1.77)	(1.04)
23	7.52	4.72	1.89	14.57	1.23	0.01	0.02	0.36	0.01	0.57	6.18	6.64
	(1.05)	(3.33)	(2.92)	(8.30)	(1.73)	(0.11)	(0.15)	(0.48)	(0.11)	(0.49)	(0.91)	(0.92)
24	7.34	3.88	1.70	15.24	1.02	0.00	0.02	0.53	0.02	0.62	6.77	7.12
	(0.93)	(2.67)	(3.20)	(8.74)	(1.64)	(0.06)	(0.13)	(0.50)	(0.13)	(0.48)	(1.20)	(0.94)
25	7.52	5.41	1.95	13.79	1.20	0.02	0.01	0.37	0.03	0.57	6.80	7.86
	(1.16)	(3.35)	(3.05)	(8.25)	(1.76)	(0.14)	(0.11)	(0.48)	(0.16)	(0.50)	(2.42)	(1.58)
26	7.55	4.58	2.76	16.32	1.58	0.01	0.03	0.42	0.03	0.48	6.33	7.32
	(0.96)	(3.26)	(5.14)	(9.43)	(2.07)	(0.11)	(0.18)	(0.49)	(0.16)	(0.50)	(1.65)	(1.18)
27	7.59	5.37	2.81	14.13	1.24	0.01	0.03	0.33	0.02	0.53	5.89	7.85
	(1.07)	(3.49)	(5.88)	(7.88)	(1.67)	(0.09)	(0.17)	(0.47)	(0.15)	(0.50)	(1.84)	(1.52)
28	7.52	5.18	2.33	14.35	1.22	0.03	0.05	0.14	0.00	0.54	7.54	9.28
	(0.79)	(2.65)	(5.54)	(8.79)	(1.66)	(0.16)	(0.23)	(0.35)	(0.05)	(0.50)	(1.31)	(1.42)
29	7.55	7.12	2.48	12.44	1.13	0.01	0.01	0.34	0.02	0.58	6.46	6.72
	(1.34)	(3.70)	(5.10)	(8.99)	(1.64)	(0.09)	(0.09)	(0.47)	(0.15)	(0.49)	(1.32)	(1.26)
34	7.36	3.21	0.87	18.60	1.22	0.04	0.10	0.44	0.01	0.58	7.31	8.50
	(0.93)	(2.34)	(1.44)	(8.90)	(1.94)	(0.21)	(0.29)	(0.50)	(0.08)	(0.49)	(1.22)	(1.03)

Note: 1) See text for definition of variables. Standard deviations are in parentheses.
2) Two digit industry codes are taken from RAIS, which uses the IBGE classification.

11: Metallurgy | 12: Machinery | 13: Electronics | 14: Transport | 15: Wood
16: Furniture | 17: Paper | 18: Rubber | 19: Leather | 20: Chemicals
21: Pharmaceuticals | 22: Perfumes | 23: Plastics | 24: Textiles | 25: Clothing
26: Food | 27: Beverages | 28: Tobacco | 29: Editorial and Graphics | 34: Civil Construction

Appendix C. Regression of log Wage onto Human Capital Variables Segmented by Years of Schooling - Linear Specification in Schooling -

Variables	Unsegmented	Segmented by Years of Schooling (S)		
		S<6	6<= S< 11	11<=S
Schooling	0.1680 (85.9)	0.0342 (4.6)	0.1503 (18.6)	0.2622 (31.7)
Tenure	0.0173 (63.3)	0.0147 (53.7)	0.0173 (40.2)	0.0124 (27.5)
$Tenure^2$	-0.00003 (-57.1)	-0.00002 (-40.1)	-0.00003 (-27.1)	-0.00002 (-19.5)
Experience	0.0796 (48.9)	0.0493 (22.4)	0.0862 (16.5)	0.1318 (16.1)
$Experience^2$	-0.0012 (-42.4)	-0.0008 (-25.7)	-0.0014 (-20.1)	-0.0012 (-13.4)
Tenure*Schooling	-0.000011 (-0.6)	-	-	-
Experience*Schooling	0.00015 (1.4)	0.00262 (7.3)	0.00000 (1.2)	-0.00451 (8.0)
Tenure*Experience	-0.00024 (-26.2)	-0.00019 (-19.8)	-0.00024 (-12.6)	-0.00020 (-9.1)
Adjusted R-squared	0.44	0.24	0.45	0.51

Note: Regressions also controlled for demographic variables (number of dependents, migrant status, immigrant status, civil status) and log firm size at Greater Sao Paulo level and country-wide level.

APPENDIX D. ESTIMATED COEFFICIENTS FOR SELECTED VARIABLES
DEPENDENT VARIABLE: LOG WAGE

INDUSTRY*[1]	SCHOOLING		SCHOOLING x EXPERIENCE		SCHOOLING x TENURE		EXPER-IENCE(X)	X^2	TENURE	$TENURE^2$
11	0.225	(30.9)	-0.0021	(-4.9)	-0.00025	(-5.3)	0.094	-0.0013	0.019	-0.000026
12	0.202	(22.9)	-0.0006	(-1.1)	-0.00028	(-2.7)	0.100	-0.0014	0.029	-0.000059
13	0.195	(27.3)	-0.0014	(-3.6)	-0.00022	(-3.4)	0.086	-0.0010	0.017	-0.000022
14	0.159	(29.6)	0.0006	(2.0)	-0.00011	(-2.2)	0.069	-0.0009	0.018	-0.000032
15	0.170	(10.0)	-0.0017	(-1.6)	0.00042	(1.8)	0.108	-0.0017	0.016	-0.000028
16	0.160	(11.2)	-0.0019	(-2.1)	0.00052	(1.8)	0.110	-0.0016	0.016	-0.000055
17	0.144	(14.9)	0.0012	(2.0)	0.00002	(0.2)	0.069	-0.0011	0.012	-0.000018
18	0.158	(10.9)	-0.0003	(-0.4)	-0.00009	(-0.8)	0.073	-0.0009	0.013	-0.000012
19	0.186	(3.9)	-0.0021	(-0.9)	0.00095	(1.0)	0.192	-0.0037	0.027	-0.000367
20	0.197	(21.2)	-0.0005	(-0.9)	-0.00023	(-3.2)	0.095	-0.0015	0.020	-0.000030
21	0.223	(18.4)	-0.0008	(-1.3)	-0.00016	(-1.6)	0.092	-0.0012	0.017	-0.000029
22	0.279	(11.6)	0.0015	(1.8)	-0.00015	(-1.3)	0.067	-0.0012	0.013	-0.000017
23	0.176	(20.8)	0.0016	(3.5)	-0.00025	(-2.3)	0.079	-0.0012	0.023	-0.000040
24	0.152	(17.3)	0.0006	(1.5)	-0.00016	(-1.8)	0.076	-0.0011	0.021	-0.000036
25	0.152	(10.1)	0.0014	(1.8)	0.00016	(0.9)	0.046	-0.0006	0.020	-0.000040
26	0.160	(24.5)	0.0008	(2.3)	0.00007	(1.2)	0.068	-0.0010	0.014	-0.000025
27	0.144	(11.3)	0.0018	(2.4)	0.00019	(2.1)	0.048	-0.0009	0.013	-0.000027
28	0.111	(7.4)	0.0004	(0.4)	-0.00003	(-0.3)	0.040	-0.0009	0.014	-0.000022
29	0.153	(11.9)	0.0005	(0.7)	0.00011	(1.1)	0.101	-0.0016	0.026	-0.000048
34	0.135	(28.7)	0.0000	(0.0)	0.00064	(7.9)	0.050	-0.0008	0.014	-0.000060

Notes: Additional regressors for demographic variables (number of dependents, migrant status, immigrant status, civil status) and firm log firm size at Great Sao Paulo Level and country-wide level also included. Estimated coefficient for EXPERIENCE, $EXPERIENCE^2$, TENURE, $TENURE^2$ all significant at .0001 level. A summary of their t-statistics follows:

SUMMARY OF T-STATISTICS

	MEAN	STD.DEV	MIN	MAX
EXPERIENCE	9.2	4.0	3.3	17.1
$EXPERIENCE^2$	-7.5	3.2	-15.8	-2.2
TENURE	12.8	6.3	2.1	24.5
$TENURE^2$	-10.7	5.2	-19.5	-2.6

*1: Industry codes are: 11-metallurgy, 12-machinery, 13-electronics, 14-transport, 15-wood, 16-furniture, 17-paper, 18-rubber, 19-leather, 20-chemicals, 21-pharmaceuticals, 22-perfumes, 23-plastics, 24-textiles, 25-clothing, 26-food, 27-beverages, 28-tobacco, 29-editorial and graphics, 34-civil construction ppendix.E

Second Stage Regressions: Regressing Estimated Returns to Human Capital Variables onto Industry Characteristics (t-statistics in parentheses)

VARIABLES[*1]	*UNWEIGHTED*	*WEIGHTED*[*2]
DEPENDENT VARIABLE: Estimated returns to schooling		
Concentration	-0.00096 (-4.968)	-0.00100 (-3.760)
Multinational	0.00051 (2.717)	0.00040 (1.383)
Demand Concentration	-0.00114 (-2.186)	-0.00121 (-2.083)
Advanced-B	0.03850 (3.010)	0.04222 (2.900)
Traditional	-0.01192 (-1.074)	-0.01190 (-0.872)
Adjusted-R^2	0.7261	0.6411
DEPENDENT VARIABLE: Average returns to tenure (dlog(W)/dT)		
Concentration	-.0000079 (-1.54)	.
Multinational	.0000045 (0.09)	.
Demand Concentration	-.000065 (-.047)	.
Advanced-B	.0064 (1.90)	.
Traditional	.00565 (1.93)	.
Adjusted-R^2	.2095	.
DEPENDENT VARIABLE: Average returns to experience (dlog(W)/dX)		
Concentration	-.00050 (-3.58)	.
Multinational	.00017 (1.25)	
Demand Concentration	-.00013 (-0.35)	
Advanced-B	.0093 (0.992)	
Traditional	-.00060 (-.074)	
Adjusted-R^2	.4056	

Notes: (1) Concentration: four-firm industrial concentration ratios (A. Calabi. 1982)
Multinational: percent of industry product corresponding to multinational firms.
Demand Concentration: demand concentration, percent of final demand consumed by top 13% of population (from E.Sadoulet, 1985, drawn from IBGE tables).
Advanced-B: dummy variables for industries 11, 12, 13, 14. (Transportation, Electronics, Metallurgy, Machinery) as classified by Milton da Mata(1978)
DTraditional: dummy variables for industries 17, 18, 23, 24, 29. (Paper, Rubber, Plastics, Textiles, Graphics and Publishing)
(2)Model: $\beta_{[illegible]}$ = f(Concentration, Multinational, Demand Concentration, Advanced-B, Traditional)
Weighted regression was weighted by the t-statistics of schooling coefficients.

REFERENCES

- Akerlof, George A. and Lawrence F. Katz, "Do Deferred Wages Dominate Involuntary Unemployment as a Worker Discipline Device?" National Bureau of Economic Research, Working Paper No.2025, September 1986.
- Becker, Gary S., Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, 1964.
- Bulow, Jeremy I. and Lawrence H. Summers, "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment," Journal of Labor Economics, vol. 4, no. 3, 1986.
- Carmichael, Lorne, "Firm-Specific Human Capital and Promotion Ladders," The Bell Journal of Economics, Spring 1983.
- Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," Econometrica, Vol.45, No.1, January 1977.
- Hashimoto, Masanori, "Firm-Specific Human Capital as a Shared Investment," American Economic Review, Vol. 71, No.3, June 1981.
- Heckman, James J. and Guilherme L.Sedlacek, "Self-Selection and the Distribution of Hourly Wages," Journal of Labor Economics, vol.8, No.1, Part 2, 1990.
- Katz, Lawrence and Lawrence Summers, "Industry Rents: Evidence and Implications," Brookings Papers on Economic Activity: Microeconomics, 1989.
- Lazear, Edward P., "Why Is There Mandatory Retirement?" Journal of Political Economy, vol.87, No.6, 1979.
- Lazear, Edward P., "Agency, Earnings Profiles, Productivity, and Hours Restrictions," American Economic Review, September 1981.
- Leamer, E.E., Specification Searches: Ad Hoc Inference with Nonexperimental Data, John Wiley & Sons, New York, 1978.
- Lester, Richard A., Company Wage Policies, Industrial Relations Section Research Report 77, Princeton University, 1948.
- Lewis, L. Earl, "Wage Dispersion in Manufacturing Industries 1950-1955," Monthly Labor Review, July 1956,
- Mincer, Jacob, Schooling, Experience, and Earnings, Columbia University Press, New York, 1974.
- Reynolds, Lloyd G., The Structure of Labor Markets, New York, Harper & Brothers, 1951.
- Robbins, Donald J., "Inter-Industry Wage Differentials: Evidence for Brazil and Comparison to Results for the U.S.," University of California, Berkeley, mimeo, 1989.
- Roy, A.D., "Some Thoughts on the Distribution of Earnings," Oxford Economic Papers, Vol.3, June 1951.

- Slichter, Sumner, "Notes on the Structure of Wages," Review of Economics and Statistics, Vol.32, 1950.
- Stiglitz, Joseph E., "Theories of Wage Rigidity," Keynes' Economic Legacy, Butkiewicz, J.L. et al (ed.), Praeger, 1986.
- Swamy, P.A.V.B., "Efficient Inference in a Random Coefficient Regression Model," Econometrica, Vol.38, 1970.
- Swamy, P.A.V.B., Statistical Inference in Random Coefficient Regression Models, Springer-Verlag, New York, 1971.
- Willis, Robert J., "Wage Determinance: A Survey and Reinterpretation of Human Capital Earnings Functions," Handbook of Labor Economics, Vol.1, Chapter 10, 1986.
- Yellen, Janet L., "Efficiency Wage Models of Unemployment," American Economic Review, vol.74, 1984.

IPEA
Instituto de Pesquisa
Econômica Aplicada

Banco Mundial

PNUD
Programa das Nações
Unidas para o Desenvolvimento

SEMINÁRIO SOBRE EDUCAÇÃO, CRESCIMENTO E DESIGUALDADE NO BRASIL

SEMINAR ON EDUCATION, GROWTH AND INEQUALITY IN BRAZIL

Rio de Janeiro, Brasil

24 a 27 de Março de 1991

EDUCATION, MOBILITY AND GROWTH

Irma Adelman
Samuel Morley
Christoph Schenzler
Steven Vogel

documento 13

Education, Mobility and Growth

by

Irma Adelman*

Samuel Morley**

Christoph Schenzler**

Steven Vogel*

March 1991

Preliminary

*University of California

**Vanderbilt University

One of the unexplained puzzles in Brazilian development experience is the apparent dramatic narrowing in income differentials during the 1970's. Prior to that rapid and skill-intensive growth had caused widening in education-income differential despite substantial investments in education, particularly higher education. Between 1970 and 1980 this pattern was reversed, and real income gains for the less educated were 2-3 times greater than those of high school or college graduates. What is even more puzzling is that this occurred while the distribution of income continued to become more unequal. Since the education income differential is a key determinant of the social and private desirability of investments in education, the causes of this narrowing demand attention.

One of the possible explanations for the reversal in the differential is that it was the lagged result of the rapid increase in educated labor that took place in Brazil in the 1960's and which continued into the 1970's. It could also be that growth was somewhat less skill intensive than it had been earlier. We want to investigate a third possible contributing factor, namely, the bias introduced by new entrants.

We know that reported income patterns are based on cross section household censuses. Those censuses obviously contain different numbers of people at different points in time. If educated new entrants earned less than the average wage of their peer group, the observed average wage of the educated could appear to stagnate over time even though the wages of educated survivors rose rapidly. By survivors we mean those present in both censuses. This downward bias is always present, but it would get worse in periods like the 1970's when there is a rapid rise in the numbers of new entrants with high levels of education.

To determine the importance of this possible source of downward bias, we need some method which separates new entrants from survivors and measure their income separately. Having done that we would like to know what the income trajectories over the 1970's were for different education and income classes. Even if we separate successfully survivors from new entrants, the available data cannot tell us that. They will tell us how many people with different levels of education earned different amounts of income in both 1970 and 1980, but they will not tell us how many from each income class moved up (or down) into other income classes. In other words, they do not tell us about income mobility. That question is particularly important for the educated.

Young educated workers may well start in a relatively low income jobs, but their education gives them upward mobility. For it gives them the ability to learn new skills and take advantage of new opportunities as they arise. Indeed one of the important returns to education surely is the relatively high mobility it makes possible. We think that we have found a way to estimate this mobility, and we want to use it here to examine the effect of education and growth on the income of different education classes during the 1970's. The procedure should also shed some light on the paradoxical fact that while education differentials were narrowing, the overall distributions of income continued to become more unequal.

In the following section we describe our estimation procedure in borrowing heavily from our previous paper and discuss data problems. In section three we discuss the evidence on education, income and mobility. Section four contains our estimates of the interaction between education, income, mobility and growth, and section 5 concludes the paper.

II. The Estimation Problem

One can think of the distribution of income at any point in time as the result of a first order Markov process in which the probability that any individual will be in income class j at time t+1 will depend on which income class he or she was in at time t. Formally our interest in what happens to the income of particular groups over time could then be solved by estimating the transition matrix of the Markov process. What we have from the reported censuses are the row and column sums of transition matrices, whose jth elements are the number of people in income class j at time t and income class i at time t+1. We are looking for some way of estimating these cell entries, given our observation of the row and column sums. Since there are only 2n data points and n^2 unknowns, we need some additional data or restrictions to make progress.

Telser (1963) addressed this problem in the context of market shares for cigarettes using a time series approach. If one takes a sufficient number of observations of the distribution (in his case of cigarette smokers across brands) and if one assumes these distributions are generated by a first order Markov process, Telser showed how to derive an unbiased regression estimator of the unknown elements of the transition matrix. The method gives the transition matrix which minimizes the difference between the actual distribution at time t+1 and the distribution predicted by applying the transition matrix to the distribution at time t. Lee, Judge and Zellner (1970) propose alternative Bayesian and non-Bayesian approaches to the estimation of transition probabilities from time series data on marginal totals and examine the properties of these estimates.

Unfortunately the time series approach is not practical for the income distribution problem in LDCs because we do not have a sufficient number of censuses. But we can use regional data from the censuses themselves as an alternative. If we have regional data, and can assume either that the same first order Markov mechanism operates in each region, or that it differs across regions in a predictable way, we can proceed, as Telser did, to use regression analysis to find the transition matrix which minimizes the difference between the observed and the predicted regional distribution at time t+1, given the observed distribution at time t.

A similar problem has been addressed in sociology and political science. In 1953 Goodman proposed a simple regression to estimate the interior elements in a four way table of individual characteristics when only the regional row and column sums of the two characteristics are known. His technique made the assumption that the interior conditional probabilities were constant across regions. Crewe and Payne (1976) applied the same general technique to get an estimate of the percentage of different occupational groups voting for the two British political parties. They extended Goodman's technique by assuming that the conditional probabilities were a function of exogenous factors that vary across regions. They derived a best linear unbiased estimator which simultaneously produced an estimate of the transition matrix and of the effect of the exogenous variables on that transition matrix. Their model was applied to a two by two case--two parties and two broad occupational classes. Our model is a simple extension of Crewe and Payne to the n-dimension case, where the n dimensions are income classes and where we are trying to find the proportion of those in income class j in time t who move to class i at time t+1.

Let P be an nxn transition matrix whose ijth element, P_{ij}, is the proportion of those in income class j at time t who move to class i at time t+1. Let Xi and Y_i be the observed fraction of the total population in income class i at time t and t+1 respectively. N is the number of mutually exclusive income classes. By definition, in matrix notation.

(1) $$Y = P*X$$

$$\text{or} \quad Y_i = \quad j \; P_{ij}X_j \qquad (i - 1,\ldots,n)$$

Equation one looks like a regression model where we observe the X's and the Y's and estimate the unknown transition parameters P_{ij}. Clearly only n-1 of these equations are independent. However, rather than dropping one of the equations, we make the equivalent restriction that the sum of each column of P_{ij}'s be equal to one. The problem with equation (1) is that we do not have enough data to estimate the P_{ij}. In our case we have 5 income classes so we are trying to estimate 25 elements of the transition matrix, but we have only five observations of the marginal totals X_i and Y_i.

We proceed by using regional observations. If the Markov process could be assumed to be the same across regions, we could increase the number of observations by taking regional observed values of the distribution. However it is probably unreasonable to assume that mobility is the same across regions. Instead, one would expect it to vary positively with many variables like income growth and labor force structure that vary across regions. Surely one's chances of moving up the distribution ladder are higher in fast growing or highly industrialized regions. Following Crewe and Payne (1976) it is straightforward to modify equation (1) to take account of regional variations in the transition matrix.

We hypothesize that transition probabilities are a function of observable characteristics Z that differ across regions. Thus in the simplest form with only one Z variable:

(2) $$P_{ij} = a_{ij} + b_{ij}Z$$

In (2) Z is a variable with region-specific values. In our case, the growth rate of income was used. We also attempted to introduce education in the same manner, but it was so highly non-linear in its effects that we were forced to separately estimate the income transition matrix for each education class.

If we now substitute equation (2) into equation (1) we get:

(3) $$Y_i = (a_{ij} + b_{ij}Z)X_j \quad (i - 1,\ldots,n)$$

This is the equation system we will estimate under the two restrictions:

(4) $$0 \leq P_{ij} \leq 1 \qquad \text{for all } i,\ j$$

(5) $$P_{ij} = 1 \text{ for } j = 1,\ldots,n$$

Unfortunately, available statistical packages cannot incorporate both restrictions. Packages which allow for estimation of systems of equations will incorporate the cross-section constraint (5) but not the within-equation inequality constraint (4). Bayesian packages, which can incorporate the inequality constraint, do not allow for estimation of systems of equations and thus prohibit incorporation of cross-equation constraints.

The alternative we used was treating the problem as a non-linear programming problem representing the ordinary least squares approach. The objective function minimized is the sum of squared errors and the constraints are given by equations (6), (7) and (8) below. This yields a non-linear programming problem with non-linear inequality constraints.

A representative equation of the constraint set is given by:

$$Y_i^r = \sum (a_{ij} + b_{ij}\, Z^r)\, X_j^r + \epsilon^r \tag{6}$$

where the r superscript indicates regional observations and ϵ^r is the statistical error term.

In our estimation, we required that the inequality constraint hold for all values of Z in the sample and that the cross-equation constraint hold for the mean value of Z in the sample. That is:

$$0 \le P_{ij}^r \le 1 \quad \text{for all } i,\ j \text{ and } r \tag{7}$$

$$\sum_i \overline{P_{ij}^r} = 1 \qquad \text{for } j = 1,\ldots,n \tag{8}$$

Where $P_{ij}^r = a_{ij} + b_{ij}\, Z^r$ and $\overline{P_{ij}^r}$ is the sample mean of P_{ij}^r.

The above procedure yields unbiased estimates of the parameters under the usual assumptions that the distribution of the error term is iid.

The Data

Central to the procedure we are using here are sets of regional observations of the distribution of income by age, sex, and education. The goal, in the data preparation about to be described, is to obtain an estimate of the income distribution in 1980 of those in the distribution who survived from 1970--a population we label "survivors". That means that we have to remove new entrants from the observed 1980 group in those cohorts where the 1980 labor force is larger than the 1970; and we have to remove those who retire from the 1970 group in those cohorts where the 1970 labor force is larger than in 1980.

The problem is that we have no way of knowing which of the 1980 workers are new entrants in expanding cohorts, or which will retire in those age groups which shrank during the 1970's. We can, however, determine many of the characteristics of new entrants and retirees. For each age cohort we first disaggregate by sex, education and region with each combination defining a cell. We then subtract the 1970 from the 1980 totals in each cell. If the difference is positive we know there were new entrants between 1970 and 1980, with the particular characteristics of the cell. We assumed that the new entrants had the same distribution of income as the total observed for that cell in 1980. That permits us to estimate the income distribution of survivors in each cell by subtracting, element by element, the vector of new entrants from the 1980 cell totals.

If we then aggregate across the 120 cells (2 sex, 5 for education and 13 regions) in each age cohort, we obtain an estimate of the 1980 income distribution of survivors. It is a vector giving the observed value of the X_i's which we will compare with the Y_i's obtained from the observed 1970 distribution.

For older age cohorts which had retirements instead of new entrants over the 1970's, we use a procedure similar to that described above to estimate the 1970 income distribution of those who would retire during the next decade. In any cell if the 1980 number is smaller than the 1970 number we know there were net retirements. Here we assumed that the retirees had the same income profile in 1970 as the rest of the members of the cell, and we subtract element by element the vector of retirees from the total distribution of the population in the cell in 1970. We then aggregate across the ten cells as before to get a survivors' distribution for 1970. This distribution gives the vector of Y_i's which we will use in our regressions along with the X_i's obtained from the observed 1980 distribution.

The other complication encountered in adapting census data for income-mobility estimates is regional migration. Clearly, in a country like Brazil, there is a substantial amount of interregional migration. Since we assume that mobility differs across regionswe have the choice of either excluding migrants, placing them in their destination populations or in their originating regions. We chose the last of the three options because it allows us to use the observed regional distributions without making a correction for migrants similar to the one made for retirements. The disadvantage with our procedure is that for regions with substantial outmigration the income growth that we use is not equal to that of the originating region since some part comes from migrants to faster growing areas. An interesting question which we can fairly easily explore in an extension to this work is the effect of migration on mobility. How much do those who migrate contribute to the observed mobility patterns? Did migrants do better than those they left behind? We can get a good answer to both of these questions by either comparing the transition matrices of migrants and non-migrants or by putting migrants into the destination population.

The Brazilian public use census tapes upon which this work is based, are a 1% sample of the demographic censuses of 1970 and 1980. They contain data on earnings, age, sex, occupation, education, current and previous residence, time in present residence, and many other variables. We aggregated the data into the thirteen regions shown in the appendix. We treated as migrants all those who had resided in their current region for less than ten years. We then reassigned migrants to the region they reported as their previous residence.

We divided the 1970 population into ten age groups: 9 for five year intervals between 15 and 60 and one over 60. We created five education groups, classifying individuals according to the last grade passed. They are: no education: elementary (5 years or less): middle school (6-8 years): high school (9-12 years); and university.

Income is reported in current cruzeiros. We converted the 1970 data to 1980 cruzeiros using the Rio de Janeiro cost of living index. We then created the following five real income classes:

No income, 0-3599, 3600-4999, 5000-11999, > 12000

We set the upper limit of class one at 3599 CR$ because the 1970 minimum wage was 3600 measured in 1980 cruzeiros. Over the subsequent decade, the minimum wage rose in real terms reaching 4149 CR$ or $79 in 1980. That means, of course, that any worker holding a less than minimum wage job in 1970 would move from class one to class two by finding a minimum wage paying job in 1980. Note that the income variable that we used ostensibly includes earned income from all sources, but there is a substantial degree of underreporting, particularly of income from capital.

III

Education and Mobility

Before the oil shocks Brazil was often held up as the quintessential example of inequitable growth. Between 1960 and 1980 it enjoyed one of the world's highest growth rates with per capita income rising by 3.9% per year. But the benefits of this prodigious boom do not appear to have been distributed at all equally across the population. The Gini Coefficient rose from .50 to .57 between 1960 and 1970 and a further 2 percent during the 1970's. During the 1970's the income share of the top 20% rose from 61.7% to 63.3% while the share of the bottom 60% shrank by a similar amount.

When we look at income gains across education classes, however, the picture appears to be different. In the 1960's education differentials widened significantly exacerbating other sources of rising inequality. In the 1970's they narrowed. (See Table one) Given the continued increase in inequality during the seventies this is surprising. It could be that the rapid increase in the supply of the educated occuring over the decade, pushed down the skill differential. That is what table one seems to suggest. The other possibility is that large numbers of new entrants with high education levels, came into the labor force earning relatively low wages. That would pull down the reported average wage of the educated. Meanwhile relatively few illiterate new entrants entered the 1970's labor market so that the reported gains for this educational group may really reflect the upward mobility of those already in the labor force.

This discussion highlights the importance of separating new entrants from survivors to see what happened to the latter, and, in particular how education affected individual prospects over the decade.

We now look more closely at the detailed evidence on mobility using our model to estimate transition matrices for each sex, education and age combination. We estimated separately for male and female, for the four age groups 15-24, 25-39, 40-59 and 60 and older, and for three education classes; no education, primary school, and at least high school. Unfortunately there were not enough observations to permit a finer educational breakdown. Since in each of these age categories we had separate observations for each 5 year interval we obtained a fairly high number of degrees of freedom. (For the 25-39 regression we had 52 d.f = 13 state x 4 age groups).

In order to conserve both space and reader patience, in Table 2 we present here only the estimated transition matrices for each of our age, education, gender combinations along with an estimate of goodness of fit and the standard errors of each coefficient. The t_{ij} in the table are the probabilities of moving from class i to j. Note here that we have transposed the matrices so that rows correspond to 1970 and columns to 1980. The entries in each of these matrices are taken from the constrained estimates of equation (3). To derive the coefficients in the table we set the growth rate at its observed level for the relevant age cohort for Brazil as a whole.

As we pointed out above, these estimates were not dervied using a standard regression routine. To estimate standard errors for the estimated coefficients we used the "delete-one jackknife technique" (i.e., subsample size - sample size - 1) by subsampling with replacement from the data and estimating the coefficients for each subsample. Such a technique was required because standard calculations do not incorporate the information contained in the restrictions and so give misleading estimates of the standard errors. Following Efron (1982) the jackknife estimate of the standard error of a paramenter estimate $_{ij}$ is calculated according to:

$$s.e.\hat{\theta}_{ij} = \sqrt{\frac{n-1}{n}\sum_{k=1}^{n}(\hat{\theta}^{k}_{ij} - \bar{\theta}_{ij})^2}$$

Where $\hat{\theta}^{k}_{ij}$ is the kth subsample estimate of the parameter, and $\bar{\theta}_{ij}$ is the sample mean of the subsample estimates of the parameter.

The R^2 above each matrix measures the percentage of the variation in the observed 1980 distributions across income classes and regions explained by the model.

The first question to ask is how well the model does and are there differences in its predictive ability across age, sex or income class. To answer that see Table 3 where we have collected the R^2 for each category and calculated their simple averages. Clearly the model performs better for males than it does for females, and is better for the 25-39 year olds who are in their prime working years. For males, as one might expect, it is difficult to predict where 15-24 year olds will end up after 10 years, knowing only where they start and their education level.

There are also clear differences in fit across education classes. Generally speaking the model does significantly worse for high school graduates than it does for the other two classes. That seems to be because the range of alternatives for such graduates is far greater than it is for the less educated partly because they start further up the income pyramid.

Turning now to the implications of our estimates, we find as we noted in our earlier paper, that there clearly is a great deal of upward mobility in Brazil. If you had zero income as a male teenager in 1970 you had a 93% chance of moving up at least one class and a 49% chance of moving up at least two classes. Males in the 20-39 group had a 44% chance of moving up at least one income class. For females the picture was not quite as rosy but there is still evidence of strong upward mobility.

To see how these mobility patterns differ across education classes we took sample averages of the t_{ij} in the relevant parts of the transition matrices and report the result in table 4. The table says that for females under 60 the overall probability of suffering downward mobility was just 8.6%, while for males 3.7%. The chance of rising at least one income class was 50% for women and 63.2% for men. When we look across education classes it does appear that upward mobility is somewhat higher for the better educated, at least for females and young(<25) males. But unexpectedly for prime age males, if anything mobility is slightly lower for the educated.

The problem with the table (4) is that it biases downward upward mobility for people who start high in the distribution. We have data only on numbers in each income class, not reported income. Thus for those who started in the top class, the best we can report is zero downward mobility. To attempt to correct for this problem and get a better measure of relative mobility across income classes we constructed Table 5.

To get the table we assigned midpoint values to each education class and then calculated the expected income for those in each age, sex and education group in 1980 based on the starting point in the income pyramid in 1970. The numbers from that calculation are part A of the table. The table tells us that for males in the 25-39 age group, with no education and earning between 3600-5000 cruzeiros in 1970, expected 1980 income was 8872 CR$ roughly double the 1970 average.

To estimate mobility one cannot simply calculate rates of growth of expected income because of the problem of the top and bottom classes. For the former one would be dividing by zero and for the latter one would appear to have no upward mobility for the reason discussed earlier. We made two alternative calculations. In Table 5B we compare the expected income of the top two education classes relative to those with no education.

In it we have divided the observed 1980 entry for each age income and gender group shown in the top portion of the table by the value for education class one. This part of the table answers the question of whether those with more education had higher expected 1980 income than those with less, when they both started in the same income class in 1970. In other words, did education confer mobility regardless of where one started in the income pyramid.

The table leaves little doubt on this question. In 60 out of the 80 possible cases the top two education classes did better than the bottom (had a coefficient >1). The advantage is even clearer for the two middle age groups where only 5/40 of the observed coefficients were less than one. The mobility advantage is even clearer for workers with at least a high school education. If one compares the expected incomes of high school and grade school graduates there are only 7 cases out of 80 where the gain of the latter exceed that of the former.

One may ask a further question of these data; did education confer more of an advantage to those who started at the top than at the bottom. That is, how did those with education and a good start do relative to those with education and a bad start? Table 5C addresses that question. In it we have divided each observation in Table 5B by the value in the zero income class. We learned from Table 5B that education raised ones prospects of upward mobility. Here we learn that the relative advantages of education tend to differ across gender. For females the advantage of education for those starting in the lowest income class is so large that it dwarfs the increases in expected income further up the income pyramid. 16/24 of the female coefficients in Table 5C are less than one, implying that education confers more of an advantage to those who start at the bottom. For males the picture is somewhat different. Education apparently did not help many of those who started at the bottom, possibly reflecting a self selection bias. From Table 5A we see that 5/8 of the expected incomes of those in the two top education classes who started with zero income, were lower than the income of those with no education which is a surprising result. For those starting further up the pyramid education apparently was more helpful than it was at the bottom. 21/32 of the coefficients in Table 5C for males are greater than one. Where education really helped males was at the top of the 1970 income pyramid. For those lucky enough to start in the top class in 1970, education raised substantially the probability of staying in the top income class in 1980. One can see that by comparing the numbers in the bottom row of Table 5C with those further up the same table. 6/8 are greater than one and 5/8 are bigger than any of the other entries in this respective column. That suggests that the skill requirements associated with modern economic growth made at least a high school level of education increasingly necessary for those who occupy Brazil's best paying jobs.

Altogether the bottom portions of the table suggest quite strongly that for survivors the education differential probably widened, which contradicts the pattern we observed for the entire labor force, and helps explain why the distribution got less equal even though the skill differential narrowed.

IV

Education, Mobility and Growth

A natural further question to ask of our regression is the interaction of growth and education. We found in our previous paper that our growth variable generally had a positive effect on mobility (Adelman et al 1990).

Here we can ask did the better educated enjoy higher mobility in the faster growing states relative to those with less education? Does our evidence suggest a positive interaction between growth and education?

There are far too many individual regressions to attempt to display all the growth coefficients directly. Instead we have counted all the cases in which there is a positive growth effect and in Table 6 we indicate both how many of those there are and how many are significantly positive and negative in parentheses. A positive growth effect means that growth raises the probability of upward mobility or lowers the probability of downard mobility. We includeds the moved the bottom three terms along the main diagonal (0,0), (1,1), (2.2) with the downward mobility cases.)

Somewhat surprisingly, the table does not suggest a very strong positive growth effect. The growth coefficient is significant and has the right sign in about one-third of the regressions, but in many of these its quantitative effect is small. As the reader can see from the table, for females fast growth rates appear to lower chances of upward mobility as often as it helps. For males growth was far more helpful, particularly for those older than 25 in 1970. To some extent the failure of the growth effect reflects the poorer fit of our regressions for women and people in the highest income class. If one looks just at those cases where the $R^2 > .50$, which we have indicated with a star in the table, a more favorable picture of the effects of growth emerges. 55% of these growth coefficients are positive and 30% are significant.

Turning now to the education-growth interaction, the evidence does not suggest a strong positive relationship between education and upward mobility. Those instances where the growth coefficient is positive are not higher for the better educated except in the youngest age group. As one goes across each row of the table one does not find an increase in the number of significant positive cases. Indeed they fall for the three oldest male age groups, and the two oldest female groups. What this says is that upward mobility is not heavily dependent on growth which is an unexpected and surprising result that requires further investigation.

CONCLUSION

We draw several conclusions from this work. First, the estimation procedure works quite well and provides good estimates of mobility, but it works less well in some cases. Our fits were good for males and for prime age females, but they tended to be much worse for those with high education levels or who were over 60 in 1970. This may reflect the smaller sample size in these cases.

Second, the separation of new entrants from survivors appears to be important for understanding what happened to education differentials. Whereas the reported differential narrowed quite sharply between 1970 and 1980, that was almost surely not true for survivors alone. Education had a strong positive effect on upward mobility, at all points in the income pyramid, and was particularly strong at the top. Since the educated survivors were bunched in the upper income classes in 1970, that pattern suggests a widening of education differentials and income inequality.

Finally, we could find little evidence of a positive interaction between growth, education and upward mobility. That says the better educated tended to have high mobility regardless of the region in which they were located.

TABLE ONE

AVERAGE INCOME GAINS BY EDUCATION CLASS

	1960-1970	1970-1980
Illiterates	0%	86.0%
Primary	13.7	70.6
Secondary (lower)	9.5	4.1
High School	28.4	29.4
College	51.9	27.3
Overall		71.6

Source: 1960-70: Langoni (1973), p. 86.
1970-80: Benevides (1985), p. 120.

TABLE 2

FEMALE AGE GROUP: 15-24.

Eqations: (1) $TP0(I,J) = P0(I) + ST0(I) * GR(J),$

(2) $TP1(I,J) = P1(I) + ST1(I) * GR(J),$

(3) $TP2(I,J) = P2(I) + ST2(I) * GR(J),$

(4) $TP3(I,J) = P3(I) + ST3(I) * GR(J),$

(5) $TP0(I,J) = P4(I) + ST4(I) * GR(J),$

(6) $X(I,J) = TPO(I,J) * Y0(J) + TP1(I,J) * Y1(J) + TP2(I,J)*Y2(J) + TP3(I,J)* Y3(J) + TP4(I,J) * Y4(J),$

for $I, = 0, 1, 2, 3, 4.,$ and $J = 1 \ldots 26.$

Or, (6') $X = TY$, where T = transition probability matrix.

EDUCATION LEVEL = 1. $R^2 = 0.300$. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc} & X = 1980 \\ \begin{array}{c} \\ 1970 \\ \\ \\ Y \\ \\ \\ \\ \\ \\ \end{array} & \left[\begin{array}{ccccc} 0.310 & 0.321 & 0.142 & 0.152 & 0.075 \\ (0.006) & (0.033) & (0.013) & (0.014) & (0.010) \\ 0.041 & 0.483 & 0.135 & 0.218 & 0.124 \\ (0.003) & (0.014) & (0.005) & (0.007) & (0.005) \\ 0.0 & 0.0 & 0.104 & 0.692 & 0.204 \\ & & (0.018) & (0.030) & (0.013) \\ 0.0 & 0.0 & 0.0 & 0.642 & 0.358 \\ & & & (0.063) & (0.063) \\ 0.0 & 0.0 & 0.190 & 0.810 & 0.0 \\ & & (0.114) & (0.114) & \end{array}\right] \end{array}$$

EDUCATION LEVEL = 2. $R^2 = 0.541$. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
\substack{0.475 \\ (0.018)} & \substack{0.117 \\ (0.023)} & \substack{0.231 \\ (0.020)} & \substack{0.152 \\ (0.020)} & \substack{0.024 \\ (0.006)} \\
\substack{0.026 \\ (0.002)} & \substack{0.553 \\ (0.008)} & \substack{0.150 \\ (0.004)} & \substack{0.231 \\ (0.006)} & \substack{0.040 \\ (0.001)} \\
0.0 & 0.0 & \substack{0.304 \\ (0.055)} & \substack{0.695 \\ (0.055)} & \substack{0.001 \\ (0.003)} \\
0.0 & 0.0 & \substack{0.002 \\ (0.012)} & \substack{0.854 \\ (0.039)} & \substack{0.144 \\ (0.032)} \\
0.0 & \substack{0.290 \\ (0.164)} & \substack{0.122 \\ (0.154)} & \substack{0550 \\ (0.166)} & \substack{0.039 \\ (0.117)}
\end{bmatrix}
\end{array}$$

EDUCATION LEVEL = 3. $R^2 = 0.409$. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
\substack{0.011 \\ (0.046)} & \substack{0.005 \\ (0.023)} & \substack{0.011 \\ (0.044)} & \substack{0.817 \\ (0.118)} & \substack{0.157 \\ (0.098)} \\
\substack{0.026 \\ (0.004)} & \substack{0.228 \\ (0.011)} & \substack{0.147 \\ (0.008)} & \substack{0.151 \\ (0.015)} & \substack{0.447 \\ (0.021)} \\
\substack{0.023 \\ (0.006)} & \substack{0.001 \\ (0.005)} & \substack{0.089 \\ (0.011)} & \substack{0.549 \\ (0.026)} & \substack{0.338 \\ (0.028)} \\
\substack{0.039 \\ (0.012)} & \substack{0.167 \\ (0.024)} & \substack{0.095 \\ (0.017)} & \substack{0.484 \\ (0.034)} & \substack{0.216 \\ (0.021)} \\
0.0 & 0.0 & 0.0 & \substack{0.019 \\ (0.033)} & \substack{0.981 \\ (0.033)}
\end{bmatrix}
\end{array}$$

FEMALE AGE GROUP: 25-39.

Eqations: (1) $TP0(i,j) = P0(i) + ST0(i) * GR(j)$,
(2) $TP1(i,j) = P1(i) + ST1(i) * GR(j)$,
(3) $TP2(i,j) = P2(i) + ST2(i) * GR(j)$,
(4) $TP3(i,j) = P3(i) + ST3(i) * GR(j)$,
(5) $TP0(i,j) = P4(i) + ST4(i) * GR(j)$,
(6) $X(i,j) = TPO(i,j) * Y0(j) + TP1(i,j) * Y1(j) + TP2(i,j)*Y2(j) + TP3(i,j)* Y3(j) + TP4(i,j) * Y4(j)$,
for $i, = 0, 1, 2, 3, 4.$, and $j = 1. \dots 39.$
Or, (6') $X = TY$, where T = transition probability matrix.

EDUCATION LEVEL = 1. $R^2 = 0.544$. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc} & X \\ Y & \begin{bmatrix} \substack{0.483 \\ (0.014)} & \substack{0.097 \\ (0.019)} & \substack{0.135 \\ (0.010)} & \substack{0.166 \\ (0.012)} & \substack{0.119 \\ (0.012)} \\ \substack{0.070 \\ (0.002)} & \substack{0.671 \\ (0.005)} & \substack{0.115 \\ (0.003)} & \substack{0.113 \\ (0.003)} & \substack{0.031 \\ (0.002)} \\ 0.0 & 0.0 & \substack{0.374 \\ (0.020)} & \substack{0.441 \\ (0.014)} & \substack{0.185 \\ (0.013)} \\ 0.0 & 0.0 & \substack{0.313 \\ (0.040)} & \substack{0.687 \\ (0.040)} & 0.0 \\ 0.0 & 0.0 & \substack{0.271 \\ (0.112)} & \substack{0.729 \\ (0.112)} & 0.0 \end{bmatrix} \end{array}$$

EDUCATION LEVEL = 2. R^2 = 0.736. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
\substack{0.521\\(0.014)} & 0.0 & \substack{0.020\\(0.010)} & \substack{0.381\\(0.016)} & \substack{0.078\\(0.007)} \\
\substack{0.048\\(0.002)} & \substack{0.644\\(0.006)} & \substack{0.108\\(0.002)} & \substack{0.152\\(0.004)} & \substack{0.049\\(0.003)} \\
0.0 & 0.0 & \substack{0.502\\(0.008)} & \substack{0.498\\(0.008)} & 0.0 \\
\substack{0.005\\(0.006)} & 0.0 & 0.0 & \substack{0.698\\(0.014)} & \substack{0.297\\(0.010)} \\
\substack{0.199\\(0.051)} & \substack{0.393\\(0.070)} & 0.0 & \substack{0.093\\(0.053)} & \substack{0.315\\(0.049)}
\end{bmatrix}
\end{array}$$

EDUCATION LEVEL = 3. R^2 = 0.677. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
\substack{0.0\\(0.001)} & \substack{0.003\\(0.012)} & \substack{0.536\\(0.058)} & \substack{0.357\\(0.081)} & \substack{0.104\\(0.080)} \\
\substack{0.014\\(0.004)} & \substack{0.440\\(0.014)} & \substack{0.172\\(0.007)} & \substack{0.126\\(0.010)} & \substack{0.249\\(0.017)} \\
\substack{0.048\\(0.009)} & \substack{0.089\\(0.025)} & \substack{0.172\\(0.012)} & \substack{0.380\\(0.011)} & \substack{0.312\\(0.024)} \\
\substack{0.032\\(0.004)} & \substack{0.049\\(0.008)} & \substack{0.038\\(0.005)} & \substack{0.499\\(0.011)} & \substack{0.382\\(0.011)} \\
\substack{0.001\\(0.003)} & \substack{0.0\\(0.001)} & 0.0 & \substack{0.096\\(0.006)} & \substack{0.903\\(0.007)}
\end{bmatrix}
\end{array}$$

FEMALE AGE GROUP: 40-59.

Eqations: (1) TP0(I,J) = P0(I) + ST0(I) * GR(J),
(2) TP1(I,J) = P1(I) + ST1(I) * GR(J),
(3) TP2(I,J) = P2(I) + ST2(I) * GR(J),
(4) TP3(I,J) = P3(I) + ST3(I) * GR(J),
(5) TP0(I,J) = P4(I) + ST4(I) * GR(J),
(6) X(I,J) = TPO(I,J) * Y0(J) + TP1(I,J) * Y1(J) + TP2(I,J)*Y2(J) + TP3(I,J)* Y3(J) + TP4(I,J) * Y4(J),
for I, = 0, 1, 2, 3 , 4., and J = 1. . . . 52.

Or, (6') X = TY, where T = transition probability matrix.

EDUCATION LEVEL = 1. R^2 = 0.442. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \left[\begin{array}{ccccc}
0.550. & 0.149 & 0.194 & 0.088 & 0.019 \\
(0.080) & (0.030) & (0.032) & (0.021 & (0.137) \\
0.091 & 0.760 & 0.054 & 0.074 & 0.021 \\
(0.003) & (0.004) & (0.002) & (0.003) & (0.003) \\
0.0 & 0.0 & 0.384 & 0.514 & 0.101 \\
 & & (0.016) & (0.019) & (0.017) \\
0.0 & 0.0 & 0.531 & 0.305 & 0.164 \\
 & & (0.080) & (0.084) & (0.121) \\
0.0 & 0.010 & 0.222 & 0.533 & 0.236 \\
 & (0.062) & (0.085) & (0.096) & (0.066)
\end{array}\right]
\end{array}
$$

EDUCATION LEVEL = 2. $R^2 = 0.146$. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
\begin{matrix}0.362\\(0.052)\end{matrix} & \begin{matrix}0.192\\(0.052)\end{matrix} & 0.0 & \begin{matrix}0.425\\(0.085)\end{matrix} & \begin{matrix}0.021\\(0.138)\end{matrix} \\
\begin{matrix}0.031\\(0.006)\end{matrix} & \begin{matrix}0.641\\(0.090)\end{matrix} & \begin{matrix}0.144\\(0.021)\end{matrix} & \begin{matrix}0.105\\(0.015)\end{matrix} & \begin{matrix}0.079\\(0.129)\end{matrix} \\
\begin{matrix}0.048\\(0.013)\end{matrix} & \begin{matrix}0.114\\(0.030)\end{matrix} & \begin{matrix}0.340\\(0.049)\end{matrix} & \begin{matrix}0.162\\(0.041)\end{matrix} & \begin{matrix}0.336\\(0.094)\end{matrix} \\
\begin{matrix}0.090\\(0.020)\end{matrix} & 0.0 & \begin{matrix}0.006\\(0.017)\end{matrix} & \begin{matrix}0.723\\(0.105)\end{matrix} & \begin{matrix}0.182\\(0.116)\end{matrix} \\
0.0 & \begin{matrix}0.080\\(0.080)\end{matrix} & 0.0 & \begin{matrix}0.679\\(0.122)\end{matrix} & \begin{matrix}0.241\\(0.113)\end{matrix}
\end{bmatrix}
\end{array}$$

DUCATION LEVEL = 3. $R^2 = 0.267$. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
0.0 & \begin{matrix}0.030\\(0.034)\end{matrix} & \begin{matrix}0.599\\(0.089)\end{matrix} & \begin{matrix}0.006\\(0.032)\end{matrix} & \begin{matrix}0.364\\(0.108)\end{matrix} \\
\begin{matrix}0.008\\(0.008)\end{matrix} & \begin{matrix}0.390\\(0.020)\end{matrix} & \begin{matrix}0.187\\(0.035)\end{matrix} & \begin{matrix}0.413\\(0.044)\end{matrix} & \begin{matrix}0.003\\(0.014)\end{matrix} \\
\begin{matrix}0.039\\(0.009)\end{matrix} & \begin{matrix}0.145\\(0.036)\end{matrix} & \begin{matrix}0.126\\(0.029)\end{matrix} & \begin{matrix}0.277\\(0.058)\end{matrix} & \begin{matrix}0.413\\(0.086)\end{matrix} \\
0.0 & \begin{matrix}0.043\\(0.013)\end{matrix} & \begin{matrix}0.003\\(0.013)\end{matrix} & \begin{matrix}0.466\\(0.023)\end{matrix} & \begin{matrix}0.487\\(0.020)\end{matrix} \\
\begin{matrix}0.083\\(0.007)\end{matrix} & \begin{matrix}0.042\\(0.009)\end{matrix} & \begin{matrix}0.051\\(0.008)\end{matrix} & \begin{matrix}0.189\\(0.016)\end{matrix} & \begin{matrix}0.636\\(0.031)\end{matrix}
\end{bmatrix}
\end{array}$$

FEMALE AGE GROUP: 60+.

Eqations: (1) $TP0(i,j) = P0(i) + ST0(i) * GR(j)$,

(2) $TP1(i,j) = P1(i) + ST1(i) * GR(j)$,

(3) $TP2(i,j) = P2(i) + ST2(i) * GR(j)$,

(4) $TP3(i,j) = P3(i) + ST3(i) * GR(j)$,

(5) $TP0(i,j) = P4(i) + ST4(i) * GR(j)$,

(6) $X(i,j) = TPO(i,j) * Y0(j) + TP1(i,j) * Y1(j) + TP2(i,j)*Y2(j) + TP3(i,j)* Y3(j) + TP4(i,j) * Y4(j)$,

for $i = 0, 1, 2, 3, 4.$, and $j = 1.... 13$.

Or, (6') $X = TY$, where T = transition probablity matrix.

EDUCATION LEVEL = 1. $R^2 = 0.468$. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \left[\begin{array}{ccccc}
0.0 & \begin{matrix}0.075\\(0.155)\end{matrix} & \begin{matrix}0.009\\(0.020)\end{matrix} & \begin{matrix}0.318\\(0.091)\end{matrix} & \begin{matrix}0.598\\(0.181)\end{matrix} \\
\begin{matrix}0.078\\(0.006)\end{matrix} & \begin{matrix}0.747\\(0.022)\end{matrix} & \begin{matrix}0.132\\(0.021)\end{matrix} & \begin{matrix}0.042\\(0.009)\end{matrix} & \begin{matrix}0.001\\(0.002)\end{matrix} \\
0.0 & \begin{matrix}0.231\\(0.119)\end{matrix} & \begin{matrix}0.073\\(0.108)\end{matrix} & \begin{matrix}0.696\\(0.093)\end{matrix} & 0.0 \\
0.0 & 0.0 & \begin{matrix}0.361\\(0.067)\end{matrix} & \begin{matrix}0.639\\(0.067)\end{matrix} & 0.0 \\
0.0 & 0.0 & \begin{matrix}0.632\\(0.085)\end{matrix} & \begin{matrix}0.340\\(0.130)\end{matrix} & \begin{matrix}0.028\\(0.096)\end{matrix}
\end{array}\right]
\end{array}
$$

EDUCATION LEVEL = 2. $R^2 = 0.387$. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
\substack{0.923 \\ (0.266)} & 0.0 & \substack{0.023 \\ (0.080)} & \substack{0.054 \\ (0.186)} & 0.0 \\[1em]
\substack{0.162 \\ (0.039)} & \substack{0.564 \\ (0.042)} & \substack{0.049 \\ (0.021)} & \substack{0.058 \\ (0.026)} & \substack{0.167 \\ (0.030)} \\[1em]
\substack{0.030 \\ (0.069)} & \substack{0.011 \\ (0.025)} & \substack{0.433 \\ (0.115)} & \substack{0.507 \\ (0.114)} & \substack{0.020 \\ (0.042)} \\[1em]
\substack{0.091 \\ (0.085)} & \substack{0.438 \\ (0.114)} & \substack{0.216 \\ (0.109)} & \substack{0.253 \\ (0.109)} & \substack{0.003 \\ (0.009)} \\[1em]
0.0 & \substack{0.262 \\ (0.127)} & 0.0 & \substack{0.668 \\ (0.051)} & \substack{0.070 \\ (0.112)}
\end{bmatrix}
\end{array}
$$

EDUCATION LEVEL = 3. $R^2 = 0.536$. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
\substack{0.007 \\ (0.012)} & 0.0 & \substack{0.168 \\ (0.125)} & \substack{0.091 \\ (0.198)} & \substack{0.735 \\ (0.274)} \\[1em]
0.0 & \substack{0.031 \\ (0.029)} & \substack{0.604 \\ (0.076)} & \substack{0.359 \\ (0.053)} & \substack{0.006 \\ (0.019)} \\[1em]
\substack{0.240 \\ (0.089)} & \substack{0.001 \\ (0.004)} & \substack{0.258 \\ (0.116)} & \substack{0.105 \\ (0.088)} & \substack{0.396 \\ (0.255)} \\[1em]
\substack{0.005 \\ (0.008)} & \substack{0.014 \\ (0.015)} & \substack{0.100 \\ (0.065)} & \substack{0.310 \\ (0.114)} & \substack{0.571 \\ (0.143)} \\[1em]
\substack{0.230 \\ (0.080)} & \substack{0.024 \\ (0.021)} & \substack{0.362 \\ (0.047)} & \substack{0.111 \\ (0.083)} & \substack{0.273 \\ (0.118)}
\end{bmatrix}
\end{array}
$$

MALE AGE GROUP: 15-24.

Eqations: (1) TP0(i,j) = P0(i) + ST0(i) * GR(j),
(2) TP1(i,j) = P1(i) + ST1(i) * GR(j),
(3) TP2(i,j) = P2(i) + ST2(i) * GR(j),
(4) TP3(i,j) = P3(i) + ST3(i) * GR(j),
(5) TP0(i,j) = P4(i) + ST4(i) * GR(j),
(6) X(i,j) = TPO(i,j) * Y0(j) + TP1(i,j) * Y1(j) + TP2(i,j)*Y2(j) + TP3(i,j)* Y3(j) + TP4(i,j) * Y4(j),
for i, = 0, 1, 2, 3 , 4., and j = 1. . . . 26.

Or, (6') **X = TY**, where **T** = transition probability matrix.

EDUCATION LEVEL = 1. R^2 = 0.508. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc} & X \\ Y & \begin{bmatrix} \begin{matrix}0.118\\(0.018)\end{matrix} & \begin{matrix}0.114\\(0.044)\end{matrix} & \begin{matrix}0.129\\(0.015)\end{matrix} & \begin{matrix}0.509\\(0.057\end{matrix} & \begin{matrix}0.130\\(0.024)\end{matrix} \\ \begin{matrix}0.010\\(0.004)\end{matrix} & \begin{matrix}0.464\\(0.016)\end{matrix} & \begin{matrix}0.223\\(0.006)\end{matrix} & \begin{matrix}0.215\\(0.015)\end{matrix} & \begin{matrix}0.088\\(0.008)\end{matrix} \\ 0.0 & 0.0 & 0.0 & \begin{matrix}0.473\\(0.035)\end{matrix} & \begin{matrix}0.527\\(0.035)\end{matrix} \\ 0.0 & 0.0 & \begin{matrix}0.019\\(0.022)\end{matrix} & \begin{matrix}0.535\\(0.038)\end{matrix} & \begin{matrix}0.447\\(0.049)\end{matrix} \\ 0.0 & 0.0 & \begin{matrix}0.004\\(0.018)\end{matrix} & \begin{matrix}0.427\\(0.113)\end{matrix} & \begin{matrix}0.570\\(0.115)\end{matrix} \end{bmatrix} \end{array}$$

EDUCATION LEVEL = 2. $R^2 = 0.644$. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
0.085 & 0.168 & 0.146 & 0.494 & 0.107 \\
(0.007) & (0.034) & (0.015) & (0.033) & (0.020) \\
0.014 & 0.290 & 0.178 & 0.386 & 0.131 \\
(0.002) & (0.011) & (0.004) & (0.007) & (0.008) \\
0.0 & 0.0 & 0.020 & 0.685 & 0.295 \\
 & & (0.018) & (0.021) & (0.028) \\
0.0 & 0.0 & 0.079 & 0.309 & 0.612 \\
 & & (0.031) & (0.043) & (0.061) \\
0.0 & 0.0 & 0.006 & 0.0 & 0.994 \\
 & & (0.031) & & (0.031)
\end{bmatrix}
\end{array}
$$

EDUCATION LEVEL = 3. $R^2 = 0.383$. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
0.001 & 0.023 & 0.003 & 0.046 & 0.928 \\
(0.003) & (0.009) & (0.007) & (0.052) & (0.064) \\
0.015 & 0.065 & 0.063 & 0.323 & 0.535 \\
(0.002) & (0.006) & (0.005) & (0.027) & (0.032) \\
0.023 & 0.004 & 0.067 & 0.451 & 0.455 \\
(0.003 & (0.004) & (0.009) & (0.041) & (0.049) \\
0.008 & 0.025 & 0.014 & 0.185 & 0.768 \\
(0.001) & (0.002) & (0.004) & (0.015) & (0.018) \\
0.0 & 0.0 & 0.0 & 0.003 & 0.997 \\
 & & & (0.016) & (0.016)
\end{bmatrix}
\end{array}
$$

MALE AGE GROUP: 25-39.

Eqations: (1) TP0(I,j) = P0(I) + ST0(I) * GR(j),
(2) TP1(I,j) = P1(I) + ST1(I) * GR(j),
(3) TP2(I,j) = P2(I) + ST2(I) * GR(j),
(4) TP3(I,j) = P3(I) + ST3(I) * GR(j),
(5) TP0(I,j) = P4(I) + ST4(I) * GR(j),
(6) X(I,j) = TPO(I,j) * Y0(j) + TP1(I,j) * Y1(j) + TP2(I,j)*Y2(j) + TP3(I,j)* Y3(j) + TP4(I,j) * Y4(j),
for I, = 0, 1, 2, 3, 4., and j = 1. . . . 39.

Or, (6') **X = TY, where T = transition probablity matrix.**

EDUCATION LEVEL = 1. R^2 = 0.802. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
0.0 & 0.0 & \substack{0.175 \\ (0.040)} & \substack{0.502 \\ (0.053)} & \substack{0.322 \\ (0.047)} \\
\substack{0.041 \\ (0.001)} & \substack{0.557 \\ (0.003)} & \substack{0.188 \\ (0.002)} & \substack{0.186 \\ (0.003)} & \substack{0.028 \\ (0.002)} \\
0.0 & 0.0 & \substack{0.144 \\ (0.011)} & \substack{0.741 \\ (0.010)} & \substack{0.115 \\ (0.011)} \\
0.0 & 0.0 & \substack{0.035 \\ (0.034)} & \substack{0.277 \\ (0.018)} & \substack{0.688 \\ (0.035)} \\
0.0 & 0.0 & \substack{0.329 \\ (0.087)} & \substack{0.655 \\ (0.083)} & \substack{0.016 \\ (0.100)}
\end{bmatrix}
\end{array}$$

EDUCATION LEVEL = 2. R^2 = 0.836. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \left[\begin{array}{ccccc}
\begin{matrix}0.043\\(0.012)\end{matrix} & \begin{matrix}0.018\\(0.042)\end{matrix} & \begin{matrix}0.099\\(0.042)\end{matrix} & \begin{matrix}0.829\\(0.072)\end{matrix} & \begin{matrix}0.011\\(0.059)\end{matrix} \\
\begin{matrix}0.035\\(0.001)\end{matrix} & \begin{matrix}0.427\\(0.005)\end{matrix} & \begin{matrix}0.206\\(0.005)\end{matrix} & \begin{matrix}0.309\\(0.005)\end{matrix} & \begin{matrix}0.023\\(0.006)\end{matrix} \\
0.0 & 0.0 & \begin{matrix}0.074\\(0.014)\end{matrix} & \begin{matrix}0.592\\(0.025)\end{matrix} & \begin{matrix}0.334\\(0.021)\end{matrix} \\
0.0 & 0.0 & \begin{matrix}0.063\\(0.012)\end{matrix} & \begin{matrix}0.537\\(0.025)\end{matrix} & \begin{matrix}0.400\\(0.025)\end{matrix} \\
0.0 & 0.0 & 0.0 & \begin{matrix}0.002\\(0.007)\end{matrix} & \begin{matrix}0.998\\(0.007)\end{matrix}
\end{array}\right]
\end{array}
$$

EDUCATION LEVEL = 3. R^2 = 0.452. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \left[\begin{array}{ccccc}
\begin{matrix}0.006\\(0.016)\end{matrix} & 0.0 & 0.0 & \begin{matrix}0.018\\(0.091)\end{matrix} & \begin{matrix}0.976\\(0.099)\end{matrix} \\
\begin{matrix}0.005\\(0.008)\end{matrix} & \begin{matrix}0.336\\(0.006)\end{matrix} & \begin{matrix}0.105\\(0.016)\end{matrix} & \begin{matrix}0.554\\(0.020)\end{matrix} & 0.0 \\
\begin{matrix}0.012\\(0.006)\end{matrix} & \begin{matrix}0.002\\(0.007)\end{matrix} & \begin{matrix}0.109\\(0.033)\end{matrix} & \begin{matrix}0.148\\(0.021)\end{matrix} & \begin{matrix}0.729\\(0.049)\end{matrix} \\
\begin{matrix}0.013\\(0.003)\end{matrix} & \begin{matrix}0.001\\(0.004)\end{matrix} & \begin{matrix}0.031\\(0.008)\end{matrix} & \begin{matrix}0.239\\(0.021)\end{matrix} & \begin{matrix}0.716\\(0.031)\end{matrix} \\
\begin{matrix}0.005\\(0.001)\end{matrix} & \begin{matrix}0.016\\(0.002)\end{matrix} & \begin{matrix}0.018\\(0.003)\end{matrix} & \begin{matrix}0.162\\(0.009)\end{matrix} & \begin{matrix}0.799\\(0.013)\end{matrix}
\end{array}\right]
\end{array}
$$

MALE AGE GROUP: 40-59.

Eqations: (1) $TP0(I,J) = P0(I) + ST0(I) * GR(J),$

(2) $TP1(I,J) = P1(I) + ST1(I) * GR(J),$

(3) $TP2(I,J) = P2(I) + ST2(I) * GR(J),$

(4) $TP3(I,J) = P3(I) + ST3(I) * GR(J),$

(5) $TP0(I,J) = P4(I) + ST4(I) * GR(J),$

(6) $X(I,J) = TPO(I,J) * Y0(J) + TP1(I,J) * Y1(J) + TP2(I,J)*Y2(J) + TP3(I,J)* Y3(J) + TP4(I,J) * Y4(J),$

for $I, = 0, 1, 2, 3, 4.,$ and $J = 1. \ldots 52.$

Or, (6') $\mathbf{X} = \mathbf{TY}$, where $\mathbf{T}$ = transition probability matrix.

EDUCATION LEVEL = 1. $R^2 = 0.510$. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
0.0 & 0.485 & 0.0 & 0.495 & 0.020 \\
 & (0.097) & & (0.098 & (0.137) \\
0.037 & 0.616 & 0.200 & 0.134 & 0.012 \\
(0.005) & (0.086) & (0.112) & (0.019) & (0.002) \\
0.0 & 0.0 & 0.190 & 0.720 & 0.090 \\
 & & (0.030) & (0.102) & (0.128) \\
0.0 & 0.0 & 0.143 & 0.401 & 0.455 \\
 & & (0.033) & (0.060) & (0.078) \\
0.0 & 0.0 & 0.501 & 0.004 & 0.495 \\
 & & (0.084) & (0.032) & (0.077)
\end{bmatrix}
\end{array}
$$

EDUCATION LEVEL = 2. R^2 = 0.696. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc} & X \\ Y & \begin{bmatrix} 0.0 & \begin{matrix}0.003\\(0.022)\end{matrix} & 0.0 & \begin{matrix}0.978\\(0.139)\end{matrix} & \begin{matrix}0.019\\(0.137)\end{matrix} \\ \begin{matrix}0.021\\(0.003)\end{matrix} & \begin{matrix}0.521\\(0.073)\end{matrix} & \begin{matrix}0.162\\(0.023)\end{matrix} & \begin{matrix}0.245\\(0.035)\end{matrix} & \begin{matrix}0.051\\(0.133)\end{matrix} \\ \begin{matrix}0.034\\(0.005)\end{matrix} & \begin{matrix}0.054\\(0.009)\end{matrix} & \begin{matrix}0.078\\(0.018)\end{matrix} & \begin{matrix}0.670\\(0.096)\end{matrix} & \begin{matrix}0.164\\(0.118)\end{matrix} \\ 0.0 & 0.0 & \begin{matrix}0.148\\(0.022)\end{matrix} & \begin{matrix}0.391\\(0.028)\end{matrix} & \begin{matrix}0.461\\(0.021)\end{matrix} \\ 0.0 & 0.0 & \begin{matrix}0.016\\(0.010)\end{matrix} & \begin{matrix}0.206\\(0.047)\end{matrix} & \begin{matrix}0.779\\(0.044)\end{matrix} \end{bmatrix} \end{array}$$

EDUCATION LEVEL = 3. R^2 = 0.386. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc} & X \\ Y & \begin{bmatrix} \begin{matrix}0.470\\(0.072)\end{matrix} & \begin{matrix}0.036\\(0.093)\end{matrix} & \begin{matrix}0.0\\(0.002)\end{matrix} & \begin{matrix}0.481\\(0.097)\end{matrix} & \begin{matrix}0.013\\(0.070)\end{matrix} \\ \begin{matrix}0.039\\(0.018)\end{matrix} & \begin{matrix}0.431\\(0.078)\end{matrix} & 0.0 & \begin{matrix}0.492\\(0.073)\end{matrix} & \begin{matrix}0.038\\(0.140)\end{matrix} \\ 0.0 & \begin{matrix}0.226\\(0.049)\end{matrix} & \begin{matrix}0.0\\(0.002)\end{matrix} & \begin{matrix}0.755\\(0.112)\end{matrix} & \begin{matrix}0.019\\(0.137)\end{matrix} \\ \begin{matrix}0.001\\(0.003)\end{matrix} & \begin{matrix}0.106\\(0.022)\end{matrix} & \begin{matrix}0.096\\(0.015)\end{matrix} & \begin{matrix}0.244\\(0.037)\end{matrix} & \begin{matrix}0.552\\(0.028)\end{matrix} \\ \begin{matrix}0.007\\(0.001)\end{matrix} & \begin{matrix}0.002\\(0.007)\end{matrix} & \begin{matrix}0.027\\(0.004)\end{matrix} & \begin{matrix}0.116\\(0.010)\end{matrix} & \begin{matrix}0.849\\(0.020)\end{matrix} \end{bmatrix} \end{array}$$

MALE AGE GROUP: 60+.

Eqations: (1) TP0(i,j) = P0(i) + ST0(i) * GR(j),
(2) TP1(i,j) = P1(i) + ST1(i) * GR(j),
(3) TP2(i,j) = P2(i) + ST2(i) * GR(j),
(4) TP3(i,j) = P3(i) + ST3(i) * GR(j),
(5) TP0(i,j) = P4(i) + ST4(i) * GR(j),
(6) X(i,j) = TPO(i,j) * Y0(j) + TP1(i,j) * Y1(j) + TP2(i,j)*Y2(j) + TP3(i,j)* Y3(j) + TP4(i,j) * Y4(j),
for i, = 0, 1, 2, 3 , 4., and j = 1. . . . 13.

Or. (6') X = TY, where T = transition probabilty matrix.

EDUCATION LEVEL = 1. R^2 = 0.696. TRANSITION PROBABILITY MATRIX:

$$\begin{array}{cc} & X \\ Y & \left[\begin{array}{ccccc}
0.0 & \begin{matrix}0.026\\(0.050\end{matrix} & \begin{matrix}0.016\\(0.056)\end{matrix} & \begin{matrix}0.958\\(0.101)\end{matrix} & 0.0 \\
\begin{matrix}0.030\\(0.003)\end{matrix} & \begin{matrix}0.517\\(0.009)\end{matrix} & \begin{matrix}0.228\\(0.007)\end{matrix} & \begin{matrix}0.195\\(0.012)\end{matrix} & \begin{matrix}0.030\\(0.006)\end{matrix} \\
0.0 & 0.0 & \begin{matrix}0.002\\(0.003)\end{matrix} & \begin{matrix}0.917\\(0.043)\end{matrix} & \begin{matrix}0.081\\(0.042)\end{matrix} \\
0.0 & 0.0 & \begin{matrix}0.440\\(0.077)\end{matrix} & \begin{matrix}0.070\\(0.068)\end{matrix} & \begin{matrix}0.490\\(0.051)\end{matrix} \\
0.0 & 0.0 & \begin{matrix}0.545\\(0.087)\end{matrix} & \begin{matrix}0.240\\(0.158)\end{matrix} & \begin{matrix}0.216\\(0.115)\end{matrix}
\end{array}\right]\end{array}$$

EDUCATION LEVEL = 2. $R^2 = 0.624$. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
0.0 & 0.471 & 0.457 & 0.072 & 0.0 \\
 & (0.164) & (0.146) & (0.250) & \\
0.051 & 0.420 & 0.230 & 0.199 & 0.100 \\
(0.008) & (0.027) & (0.024) & (0.038) & (0.054) \\
0.0 & 0.0 & 0.004 & 0.671 & 0.325 \\
 & & (0.011) & (0.104) & (0.107) \\
0.009 & 0.016 & 0.179 & 0.412 & 0.384 \\
(0.008) & (0.022) & (0.056) & (0.103) & (0.102) \\
0.0 & 0.0 & 0.0 & 0.255 & 0.745 \\
 & (0.002) & & (0.129) & (0.128)
\end{bmatrix}
\end{array}
$$

EDUCATION LEVEL = 3. $R^2 = 0.543$. TRANSITION PROBABILITY MATRIX:

$$
\begin{array}{cc}
 & X \\
Y & \begin{bmatrix}
0.0 & 0.400 & 0.497 & 0.0 & 0.104 \\
 & (0.134) & (0.151) & & (0.205) \\
0.0 & 0.038 & 0.512 & 0.450 & 0.0 \\
 & (0.114) & (0.151) & (0.185) & \\
0.0 & 0.007 & 0.111 & 0.729 & 0.153 \\
 & (0.016) & (0.092) & (0.187) & (0.138) \\
0.006 & 0.528 & 0.108 & 0.007 & 0.351 \\
(0.015) & (0.050) & (0.050) & (0.019) & (0.084) \\
0.021 & 0.007 & 0.005 & 0.094 & 0.873 \\
(0.005) & (0.011) & (0.013) & (0.022) & (0.029)
\end{bmatrix}
\end{array}
$$

TABLE 3

REGRESSIONS R_2's

Education Class

Female:	1	2	3	Overall
15-24	.300	.541	.409	.42
25-39	.544	.736	.677	.65
40-59	.422	.146	.267	.29
60+	.468	.387	.536	.46
Overall	.433	.453	.472	.452
Males:				
15-24	.508	.383	.644	.51
25-39	.802	.836	.452	.70
40-59	.510	.696	.386	.53
60+	.696	.624	.543	.62
Overall	.629	.634	.506	.59

TABLE 4
MOBILITY PROBABILITIES
BY SEX AND EDUCATION CLASS

A) Downward Mobility (average of lower off-diagonals)

	Education Class 1	2	3	
Females				8.6%
15-24	10.4%	9.9%	3.7%	
25-39	13.2	7.4	3.7	
40-59	12.8	10.5	6.0	
Males				3.7%
15-24	4.6	.1	.1	
25-39	10.6	1.0	2.7	
40-59	6.9	4.8	6.2	

B) Chance of Staying in the Same Class

Females				41.4%
15-24	30.7	44.5	35.9	
25-39	44.3	53.6	40.3	
40-59	44.7	46.1	32.4	
Males				33.1%
15-24	33.7	33.9	26.3	
25-39	19.9	41.6	29.8	
40-59	34.0	35.4	39.9	

C) Upward Mobility (average of upper off diagonal)

Females				50.0%
15-24	58.9	45.6	60.4	
25-39	42.5	39.0	56.0	
40-59	42.5	43.4	61.6	
Males				63.2%
15-24	61.7	66.0	73.6	
25-39	69.5	57.4	67.5	
40-59	59.1	59.8	53.9	

Table 5: Expected Income in 1980 by age and
education class

male			female				female					
class	15-24-1	15-24-2	15-24-3	25-39-1	25-39-2	25-39-3	40-59-1	40-59-2	40-59-3	60+-1	60+-2	60+-3
0	4044.3	3009.2	9674.3	4276.4	4650.5	7115.4	2307.5	4487.9	8895.7	13110.2	557.9	13990.9
0-3600	5845.6	4781.6	10130.2	3793.7	4328.2	7231.6	3270.2	4585.8	5418.6	2958.5	5065.5	5834.4
3600-500	9797.2	7231.7	10797.9	8501.7	6391.6	9513.9	7737.2	8858.8	10308.8	6853.6	6541.1	8736.6
5-12	11543	9715.6	8645.4	7185.4	10982	11031.2	7663.8	9265.32	12369	6983.8	4312.9	12809.8
12plus	7702	6645.6	16838.5	7361.8	7206.6	16167	9254.1	10084.5	12751.2	6086.6	7575.4	7205.9

table 5b

0	1	0.744059	2.392082	1	1.087480	1.663876	1	1.944918	3.855124	1	0.042554	1.067176
0-3600	1	0.817982	1.732961	1	1.140891	1.906212	1	1.402299	1.656962	1	1.712185	1.972080
3600-500	1	0.738139	1.102141	1	0.751802	1.119058	1	1.144962	1.332368	1	0.954403	1.274746
5-12	1	0.841687	0.748973	1	1.528376	1.535224	1	1.208972	1.613951	1	0.617557	1.834216
12plus	1	0.862840	2.186250	1	0.978918	2.196066	1	1.089733	1.377897	1	1.244602	1.183895

table 5c

0	1	1	1	1	1	1	1	1	1	1	1	1
0-3600	1	1.099351	0.724457	1	1.049114	1.145645	1	0.721006	0.429807	1	40.23497	1.847941
3600-500	1	0.992043	0.460745	1	0.691325	0.672561	1	0.588694	0.345609	1	22.42771	1.194503
5-12	1	1.131210	0.313105	1	1.405429	0.922679	1	0.621605	0.418650	1	14.51210	1.718755
12plus	1	1.159639	0.913952	1	0.900171	1.319849	1	0.560297	0.357419	1	29.24716	1.109371

income class	male 15-24-1	15-24-2	males 15-24-3	25-39-1	25-39-2	males 25-39-3	40-59-1	40-59-2	males 40-59-3	60+-1	60+-2	60+-3
0	7399	7099.4	16242	10493.5	7707.8	16745	5857	8644.1	4406.7	8282	3848.8	4985.1
0-3600	5535.2	7056.4	12286.9	4369.3	5056.2	6067.7	3866.2	5052.8	5991.7	4543.8	5514.5	6129.2
3600-500	12979.5	10923.5	11867.4	8872.7	11028.2	14125.1	8467	8964.2	7350.7	9180.1	11245.7	9293.7
5-12	12228.2	13370.2	14756.2	14201	11635.4	14339.5	11758.4	11796.9	12157	10817	10842.9	7916.5
12plus	13336.7	16923.8	16974.5	7254	16983	15080.6	10603.3	15062.8	15540.5	8055.5	14832.5	15680.4

table 5B

0	1	0.959508	2.195161	1	0.734530	1.595749	1	1.475857	0.752381	1	0.464718	0.601919
0-3600	1	1.274822	2.219775	1	1.157210	1.388712	1	1.306916	1.549764	1	1.213631	1.348915
3600-500	1	0.841596	0.914318	1	1.242936	1.591973	1	1.058722	0.868158	1	1.225008	1.012374
5-12	1	1.093390	1.206735	1	0.819336	1.009752	1	1.003274	1.033899	1	1.002394	0.731857
12plus	1	1.268964	1.272766	1	2.341191	2.078935	1	1.420576	1.465628	1	1.841188	1.946545

table 5C

0	1	1	1	1	1	1	1	1	1	1	1	1
0-3600	1	1.328621	1.011212	1	1.575441	0.870256	1	0.885529	2.059811	1	2.611540	2.241021
3600-500	1	0.877112	0.416515	1	1.692149	0.997633	1	0.717360	1.153880	1	2.636021	1.681909
5-12	1	1.139532	0.549725	1	1.115455	0.632776	1	0.679790	1.374168	1	2.156991	1.215871
12plus	1	1.322515	0.579805	1	3.187328	1.302795	1	0.962542	1.947985	1	3.962157	3.233895

TABLE 6

NUMBER OF CASES WHERE GROWTH EFFECT IS POSITIVE

	Education Class 1			Education Class 2			Education Class 3			Total Coefficients 1	Total Coefficients 2	Total Coefficients 3
	Total	pos	Neg	Total	pos	neg	Total	pos	neg			
Females												
15-24	7	(6)	(4)	*13	(7)	(3)	14	(5)	(8)	17	20	22
25-39	*9	(4)	(4)	*11	(5)	(6)	*15	(7)	(3)	17	19	23
40-59	10	(6)	(5)	9	(6)	(9)	8	(4)	(7)	20	21	23
60+	9	(4)	(4)	11	(5)	(5)	*10	(1)	(4)	17	21	23
Males												
15-24	*12	(3)	(4)	*11	(6)	(3)	12	(8)	(4)	22	18	22
25-39	*10	(9)	(2)	*10	(5)	(2)	10	(6)	(6)	17	18	23
40-59	*14	(10)	(3)	*9	(7)	(4)	12	(6)	(8)	17	19	23
60+	*11	(7)	(1)	*11	(4)	(4)	*13	(5)	(3)	17	19	20

Note: In income classes 0-3 staying in the same place is treated as downward mobility. Numbers in parentheses are number of significant positive and negative growth coefficients.

REFERENCES

Adelman, I. and S. Morley, C. Schenzler and M. Warning, (1990), Measurin-Brazilian Income Mobility with Census Data (Vanderbilt, Mimeo).

Adelman, I. (1989), "What is the Evidence on Income Inequality and Development?" Giannini Foundation Working Paper 502.

Adelman, I. and P. Whittle (1980), "Static and Dynamic Indices of Income Inequality," Canadian Journal of Development Studies, 1, 27-46.

Adelman, I., K. Subbarao, and P.S. Vashista (1985), "Some Dynamic Aspects of Rural Poverty in India," Economic and Political Weekly, 20 (39), 575-287.

Bonelli, R. and G.L. Sedlacek (1989), "Distribuicao de renda: evolucao no ultimo quarto de seculo," In G.L. Sedlacek and R. Paes de Barros, eds., Mercado de Trabalho e Distrribucao de Renda: uma Coletanea, (Rio de Janeiro: IPEA), 7-24.

Bacha, E.L. and L. Taylor (1978), "Brazilian Income Distribution in the 60's: Facts, Model Results and the Controversy," Journal of Development Studies 14, 271-297.

Crewe, I. and C. D. Payne (1966), "Another Game with Nature: An Ecological Regression Model of the British Two Party Vote Ratio in 1970," British Journal of Political Science 7, 43-81.

Denslow, D. and W.G. Tyler (1983), Perspectives on Poverty and Income Inequality in Brazil: An Analysis of the Changes During the 1970's. (Washington: The World Bank Staff Working Papers #61.

Fields, G. (1977), "Who Benefits from Economic Development? A Re-examination of Brazilian Growth in the 1960's," American Economic Review 67, 570-582.

Fishlow, A. (1972), "Brazilian Size Distribution of Income," American Economic Review 62, 391-402.

Goodman, L. (1953), "Ecological Regression and the Behavior of Individuals," American Sociological Review 18, 663-664.

Hoffman, R. and A.A. Kageyama (1986), "Distribuicao da renda no Brasil, entre Familias e entre Pessoas, em 1970 e 1980," Estudos Economicos 16, 25-51.

Langoni, C. G. (1973), Distribucao de Renda e Crescimento Economico do Brasil. (Rio de Janeiro: Expressao e Cultura).

Lee, T.C., G.G. Judge, and A. Zellner (1970), Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data (Amsterdam: North Holland).

Morley, S.A. (1981), "The Effect of Changes in the Population on Several Measures of Income Distribution," American Economic Review 71, 285-294.

__________ (1982), Labor Markets and Inequitable Growth: The Case of Authoritarian Capitalism in Brazil. (New York: Cambridge University Press).

Morley, S.A. and J.G. Williamson (1975), "Crescimento, Politica Salarial e Desigualdade: o Brasil Durante a Decada de 1960," Estudos Economicos 5, 107-139.

Pfeffermann, G. and R. Webb (1979), The Distribution of Income in Brazil. (Washington: The World Bank Working Paper #356).

Prais, J., (1955), "Measuring Social Mobility," Journal of the Royal Statistical Society, Series A-118, 55-66.

Shorrocks, A.F., (1976), "Income Mobility and the Markov Assumption," Economic Journal 86, 566-577.

Telser, L. (1963), "Least Squares Estimates of Transition Probabilities," in C. Christ et al., eds. Measurement in Economics (Stanford: Stanford University Press.

PRELIMINARY AND INCOMPLETE
COMMENTS WELCOME

Wages, schooling and background: Investments in men and women in urban Brazil

John Strauss
Rand

Duncan Thomas
Yale University

March, 1991

Prepared for World Bank Conference on Education, Growth and Inequality in Brazil
Financial support from the National Institutes of Health grant P50-HD12639-11 is gratefully acknowledged. We were very ably assisted by Jan Hartman.

1. ISSUES, FRAMEWORK AND MODELS

Our aim is to examine human capital investment decisions in urban Brazil from the perspective of individuals and their families. We begin with a study of the returns to education in labor market outcomes, measured by the logarithm of wages. Why estimate yet another set of wage functions? We are particularly interested in tracing out the heterogeneity in returns to education across regions, across sectors (market and self-employed), across cohort and race. While there have been several of these sorts of descriptions for men in Brazil, female labor force outcomes are less well understood. In this study, we pay special attention to differences in returns between men and women. In addition, we estimate semi-parametric wage functions which place no parametric restrictions on the returns to own education: this turns out to be important. Nonlinearities in estimated wage functions permit discrimination among competing models of labor markets. Furthermore, differences in returns among socioeconomic sub-groups vary dramatically across the education distribution, which has implications for inequality in Brazil. The first part of this paper describes these results; the rest of the paper attempts to uncover some of the underlying reasons which might explain them. We focus, for the moment, on the role of the household.

Part of the differences in returns to education can presumably be attributed to heterogeneity in the demand for labor (Barros and Sedlacek, 1989; Heckman and Hotz, 1986), the quality of education (Behrman and Birdsall, 1983, 1984; Behrman, Birdsall and Kaplan, 1991; Birdsall, 1985; Card and Krueger, 1990), individual ability and motivation (Griliches and Mason, 1972; Griliches, 1977). An issue which has received a good deal of attention in this literature, at least in the United States, is the extent to which heterogeneity in estimated returns can be explained by variation in family background (see, for example, the review by Willis, 1986).[1]

Over and above purging estimated returns to own education of bias due to omitted variables, the impact of background characteristics, such as parental education, is of interest in and of itself. This is particularly important in developing countries where there is much concern with inter-generational mobility. Yet there have been rather few studies of labor market outcomes in developing economies which take account of family background characteristics.[2]

Indeed, there has been a good deal of interest in precisely this topic in Brazil. In two earlier studies, Pastore (1982) and Medeiros (1982) use the 1973 *Pesquisa Nacional por Amostra de Domicilios* (PNAD) and include indices of father's occupation in male earnings functions. This work has recently been updated (Pastore and Zylberstajn, 1990) with the 1982 PNAD, focussing on mobility across social classes. Haller and Saraiva (1986) and Saraiva, Pahari and Haller (1986) also use the 1982 PNAD and include father's occupational status in earnings functions for males and females, stratified by region. All these

---

[1]For example, researchers have estimated earnings functions and examined how robust estimated returns to education are to information on siblings (Griliches, 1979; Behrman, Hrubec Taubman and Wales, 1980; Bound, Griliches and Hall, 1986 and Hauser and Sewall, 1986) and to information on parents (Leibowitz, 1974; Featherman and Hauser, 1978; Pananicolau and Psacharopoulos, 1979; Solon, 1989a, 1989b; Solon *et al.*, 1989; Willis and Rosen, 1979; Altonji, 1988; and Corcoran *et al.*, 1989, 1990).

[2]In one of the first of these studies, Carnoy (1967) finds that, in Mexico, father's occupation is strongly related to his child's wages as an adult. Behrman and Wolfe (1984) estimate household income functions for adult women in Nicaragua, using parental education as controls. Father's schooling is significantly positive. Heckman and Hotz (1986) estimate male earnings functions for Panama and find positive effects for both mother's and father's schooling, with the mother's effect being larger. Using the World Bank's Living Standards Measurement Survey from Peru, Stelcner, Arriagada and Moock report similar results for males in the market sector. Sahn and Alderman (1988) find an effect of father's predicted wages on wage outcomes in Sri Lanka. Armitage and Sabot (1987) interact own education with parents' education, finding sharply rising marginal rates of return in Kenya and to a lesser extent in Tanzania.

studies find positive effects of father's occupation, used as an indicator of the family's socioeconomic status. In a recent study, Lam and Schoeni (1990) use the 1982 PNAD to analyze earnings of married men and instead of father's occupation, include parental schooling. They also include the schooling of the spouse and parents-in-law. They find positive impacts of parents' and parents'-in-law education and important interactive effects with own education (the marginal return rising with better family background). Although not a focus of their study, Lam and Schoeni find that the relative effects of father's *vs.* mother's education depends on the adult child's education level, with the effects of father's and father's-in-law education being somewhat higher than that for mother's and mother's-in-law among more highly educated males.

Our second goal in this paper is to better understand household resource allocation decisions. We take a closer look at how background affects labor market outcomes and link its role to patterns in returns to education. Again we focus on the differences in impacts on men and women, across regions and sectors of the economy and again we find the effects vary over the education distribution.

It is very difficult to test hypotheses about household decision making in the absence of individual-specific consumption data. We can, however, infer a good deal from an examination of the impact of parental education on child human capital and labor market outcomes. This is the focus of the remainder of the paper. Does the education of fathers have the same effect on the wages of sons and daughters? Is there evidence for differential effects of a mother's educations on her sons and daughters?

There are several reasons why we might observe differential effects of a parent's education on sons and daughters. They may simply reflect child rearing technology: women tend to spend more time with their daughters and fathers with their sons (Lamb, 1976, 1987; Morgan, Lye and Condran, 1988). The differences may be efficient: for example, expected returns to these investments may differ, and the returns may be parent-specific. Women, for example, may allocate more resources to a daughter if they expect to get more in return later in life than from a son.[3] Finally, differences in the effects of background might reflect differences in preferences: it may be that fathers just prefer to allocate more resources to their daughters.

We take the analysis one step further and examine the role of the education of a spouse's parents: it is rather difficult to interpret differences of the effects on a son-in-law and a daughter-in-law as reflecting the role of technology in child rearing. We do find differences which suggest that parents allocate resources differentially to sons and daughters after marriage, either because they yield different returns or because of preferences.

Documenting these differences does not explain them. The fact that the impact of parental education on men and women is stronger among younger cohorts, suggests that we can learn a great deal about household resource allocation decisions from studying investments in the human capital of children. We turn next, therefore, to the determinants of child schooling outcomes.

There are an enormous number of studies which demonstrate that parental education is a very good predictor of child education and performance in school. For Brazil, Barros and Lam (1990) examine schooling levels of 14 year-olds in Sao Paulo and the Northeast region of Brazil, using the 1982 PNAD. They focus

---

[3]Spitze and Logan (1990) report that daughters are more likely to be in contact with and assist their parents in old age and that a single mother is significantly more likely to be in contact with her children if at least one is a daughter. According to Hess and Waring (1978), after the father's death, daughters tend to give more attention to their widowed mothers and sons give less. In contrast, Hoddinott (1990) argues that in Kenya, absent daughters tend to remit more money to fathers than mothers.

on the impacts of parental education (which they aggregate into one measure) and father's income, finding rather strong effects of parental education, but very weak effects of father's income.

There have been rather few studies of investments in human capital which explicitly examine the interaction between household resources and gender. King and Lillard (1987) report that among Malays in Malaysia, mother's and father's education have a significant effect on daughter's schooling attainment, but not on sons. Among the Chinese in Malaysia, mother's education has a positive effect on both but father's education affects only son's schooling attainment. King and Bellew (1989) report that, in Peru, both parents' education significantly raises the probability a child (aged 8 through 19) attends school. Maternal education has a bigger effect on the probability a daughter is at school, relative to a son; paternal education has a bigger impact on the son. Using data on another indicator of child human capital investments, weight for age, Bhuiya *et al.* (1986) find that in Matlab, Bangladesh, mother's education has a significant positive effect on son's weight but not on daughter's. Using height as an indicator of child health, Thomas (1990a) reports that mother's education tends to have a bigger effect on the height of a daughter and father's education has a bigger effect on the height of a son: this is true in the United States, Brazil and Ghana. In urban Brazil, Thomas (1990b) finds that non-labor income attributed to mothers has a (significantly) bigger effect on the weight for height of daughters, relative to sons, and that unearned income in the hands of fathers has a bigger impact on his sons' weight for height. Desai *et al.* (1989) find that a mother's education significantly affects her daughter's intellectual ability (measured by the Peabody Picture Vocabulary Test) but not her son's ability.

In this paper we focus on the impacts of mother's and father's education on a variety of indicators of child educational attainment. Stratifying by age, we model enrollment decisions for children; whether they have attained the level they should with continuous attendance and advancement; and the gap in attainment, for those who are behind (or have dropped out). Among children who are enrolled, we examine the probability they attend public or private school, whether they attend for only half or full day and whether they report regularly doing homework. The impacts of parents on child schooling decisions can arise for several reasons other than taste. Parents spend time with their children and this may enhance the child's productivity in school. Whether homework is regulary done may be an indication of this effect. Parents also invest financial resources, which should show up in the decision (and financial ability) to attend private or public school and possibly in whether school attendance is for all day or half. The children of better educated parents are certainly more likely to be at school. But are daughters of better educated fathers more or less likely to be enrolled than sons? Are sons of better educated mothers more or less likely to do homework than sisters?

2. RESULTS

All the data for this study are drawn from the 1982 *Pesquisa Nacional por Amostra de Domilcilios* (PNAD). An annual labor force survey, much like the Current Population Survey in the United States, it contains a special supplement in each year: in 1982 the focus was on education. Each household head and spouse reported the educational status of both parents as well as occupational status of the father. Furthermore very detailed information was gathered on schooling attendance and attainment for all household members over 7 years.

The PNADs are from a stratified probability sample that is nearly national in scope.[4] Information is available on some 112,000 households, comprised of over 500,000 individuals. Approximately xx households are located in urban areas and are the subject of this study. Two sets of results are presented.

(1) WAGE FUNCTIONS

We begin with the determinants of (the logarithms of) wages among adult men and women are reported: we focus on the returns to own education and how these vary by gender, region and sector of employment. For the Northeast, we also disaggregate by cohort and by race. We show that returns to education differ not only among these sub-groups but also differ across the education distribution. Noting that the effect of including background controls, such as parents' education, has a significant effect on estimated returns to own education, we show these effects differ not only by gender and sector of employment but, they also differ *across* the education distribution. We therefore attempt to untangle some of these differences by examining investment choices within the family. We demonstrate that the influence of the education of parents (and parents-in-law) is not linear, that spouse's parents have a large and significant impact on wages and we find, in several instances, that these influences are larger on children of the same gender.

(a) Wage functions: inter-sectoral and inter-regional comparisons

The richness of the 1982 PNAD is exploited in two dimensions which are key for this study. First, the very large sample size permits semi-parametric estimates of the returns to own education: a dummy variable for each of seventeen

[4]Rural areas in the north region of the country are not sampled for cost reasons. The urban sample is national.

years of education is included in each regression.[5] We also stratify the data in multiple dimensions to permit flexibility in the effect of education on wages. Second, the survey provides unusually detailed information on the education of each individual's mother and father, as well as the education of the spouse and the spouse's parents. These characteristics are included as indicators of family background to determine the extent to which they can explain the heterogeneity in returns to education.

In addition to estimating the determinants of total wages (measured by total income for the previous month divided by total reported hours for that month), we also distinguish wages earned in the market sector from wages in the self-employment sector treating selection into these sectors as endogenous.[6]

We estimate wage functions rather than earnings functions, in order to abstract from labor force participation and labor supply decisions. This is important from the point of view of identification and also when the characteristics of the spouse and spouse's parents are included, if husbands and wives make joint labor supply decisions.

Using data only from urban Brazil, three sets of log wage functions are estimated separately for men and women, stratified into three macro-regions which we refer to as the South,[7] Northeast[8] and Center-North.[9] We restrict ourselves

[5]The regressions are semi-parametric in the sense that while there are no parametric restrictions on the form of the log wage-education function, a series of additional characteristics are included in a particular parametric form.

[6]The PNAD collects information on primary and secondary jobs, but the job classification is only known for the primary job. For this reason, while total wage is defined as total (primary plus secondary) monthly earnings divided by total monthly hours, market wage and self-employment wages are defined using primary job information only. Less than four percent of the sample reports any secondary employment.

[7]Rio de Janeiro, Sao Paulo, Brasilia, Parana, Santa Catarina, Rio Grande do Sul, Minas Gerais and Espirito Santo.

to household heads and spouses, since those are the individuals for whom we have mother and father education information. We further restrict the age group to 25-60 year olds, to avoid sample selection issues that would arise if we were to use young and old household heads. There exist 62,087 female and 58,687 male urban dwellers, aged 25-60, who are household heads or spouses. These comprise our sample for the wage regressions. Sample characteristics appear in Appendix Table 1.

The regressions include age (and its square), race and state dummies as well as terms to control for self-selection into the labor market (for women) and into the market or self-employment sectors (for both men and women).[10] These selection terms are identified with household unearned income (which should affect the decision to participate but not directly affect an individual's wage) and knowledge of father's occupation (at the time the individual started working).[11]

---

[8]Maranhao, Piaui, Ceara, Rio Grande do Norte, Paraiba, Pernambuco, Alagoas, Sergipe and Bahia.

[9]Actually this includes both the center-west and north: Amazonas, Amapa, Acre, Para, Rondonia, Roraima, Matto Grosso, Matto Grosso Sul and Goias.

[10]For this version of the paper, separate linear probability models were run for market and self-employment for males and for any participation, market or self-employment participation for females. In principal, the market and self-employment decisions for women should be estimated jointly as a polytomous probit or logit model. This strategy will be followed in a future version of the paper. In addition, time did not allow correction of standard errors for the log wage equations. This will be done in the next version.

[11]Roughly 20 percent of males and 75-80 percent of females do not report their father's occupation. This turns out to be highly correlated with labor force participation, particularly for females. The participation equations are identified from the non-labor income and knowledge of father's occupation variables. All other variables in the wage equation are also covariates in the participation equations. Linear probability models for the participation equations are estimated reasonably precisely. R^2s fall in the range of .67 for female labor force participation, .45 for female market wage participation and .20 for self-employment. The R^2s for male sectoral choice are somewhat lower. In all first stage equations the unearned income and its square are highly significant, with individual t-statistics above 7, except for self-employment.

Estimates of returns to own education are presented in Figures 1 through 8 as well as in Appendix Table 2 and Appendix Figure 1. The data are very rich and there is an enormous amount of information in the figures and tables: our task in this section is to summarize important features for the purpose of understanding how returns to investments in human capital vary in urban Brazil.

The vast majority of studies of returns to education estimate simple models which are linear (or possibly quadratic) in education (see, for example, the reviews by Psacharopoulos, 1985, and Psacharopoulos and Woodhall, 1985). But is there *a single* return to education? Our semi-parametric estimates of the conditional expectation of log total wages are presented in Figure 1: the slope between any two years is the marginal return to that year of education. A linear regression would seem adequate if the aim was simply to estimate an average return to education: it would clearly not be satisfactory, however, if interest focusses on heterogeneity in returns across socio-economic groups. The figure is repeated for males in Appendix Figure 1.

In the South, there are a series of steps in the log wage function. The average annual return to four years of education is about 10% and much of this is due to a significantly higher return to completing the fourth year (first part of elementary schooling) relative to the previous three years. Those who leave school the following year or one year before the end of elementary school (the seventh year) gain nothing (in terms of wages) from the last year at school and the average return to the second half of elementary school is lower (8%). The return to secondary schooling (9-12 years) is significantly higher (over 13%)

Income has a strong negative impact on labor force participation, as one would expect. Knowledge of own father's occupation is significant for males and has t-statistics in the order of 100 for females. Women who report their father's occupation are up to 50 percent more likely to participate in the labor market. Other variables such as own education have expected (positive) effects on market participation and negative effect, at higher levels, on self-employment.

and is especially high for the last two years. One year of post secondary schooling yields a small return and in fact returns to post-secondary schooling never reach those for secondary education.[12]

Relative to the South, men are paid less in the Northeast across the entire education distribution. The returns to education are smaller for elementary schooling (about 8% *per annum*) but then dramatically increase and are more than double (nearly 18%) for men who attended secondary school and so much of the gap between wages in the South and Northeast disappears. Completing only the first two years of post-secondary schooling yields a dramatically lower return to men in the Northeast (negative for the second year) although returns rise substantially thereafter.

At the bottom of the education distribution, men in the Center-North are paid about the same as in the South but, since returns to elementary schooling are lower than in either of the other two macro-regions, men with some secondary schooling are paid about the same as men in the Northeast. Thereafter, the returns are very similar for men in the North and Northeast. There are several *negative* returns to an additional year of schooling, including the first year of each schooling level (the fifth, ninth and twelfth years).

Among women, the returns to education are higher in the Northeast, relative to the South, throughout the education distribution except for the last two years of elementary education and the first two years of secondary school. On average,

[12]Since samples are often relatively small at the top of the education distribution (over 15 years), we will pay less attention to those returns. Note, however, that in polynomial models these observations often have a very large influence on estimated returns. In our semi-parametric specification, of course, these observations do not contaminate estimated returns elsewhere in the education distribution.

returns are about 20% higher[13] and thus much of the gap in wages at the bottom at the education distribution has disappeared at the top. Returns to elementary school are lowest in the North but by secondary school these returns track closely the Northeast. There is less evidence for steps in the female log wage functions, except in the North (which is estimated imprecisely relative to the other two).

What might explain the steps in the wage functions for men? It may be that particular years of schooling (often the first year of each level) is just not very productive. This seems an unlikely explanation as one would expect it to apply across all regions and certainly should apply equally to men and women. Noting that returns tend to peak in the last year of each education level, and then decline, Saraiva *et al.* (1986) argue that the shape of the wage functions reflects credentialism (although they admit that other factors might also matter). It is far from clear why the effects of credentials should differ by gender and region as they do (in their Table 2 and our Figure 1). It seems plausible that the steps also reflect differences in the types of individuals who quit school at the beginning of a new education level. We will return to this issue below.

The lower panel of Figure 1 compares returns across gender. In the South, returns are about the same for men and women until the last part of elementary education (6 years) whereupon returns for women are consistently higher: these differences are about 10% on average and are significant for secondary schooling and above.[14] The relative patterns are similar in the Northeast, although

[13]13.9% in the Northeast and 11.8% in the South. For 8 years post elementary education, the average returns are 17.4% and 13.7% *per annum* respectively.

[14]The average return for 8-16 years of schooling is 13.8% for women and 12.5% for men.

females earn higher returns earlier (half way through elementary school), which are significant slightly later (towards the end of secondary school). The difference in the average rate is higher in the Northeast (25%) and so while women's wages in both the South and Northeast catch up with those of men, catch up is greater in the poorer Northeast. It is clear that wage functions that are linear in education (or a low order polynomial) will fail to capture these important differences in returns across the education distribution.

Figure 2 presents the conditional expectations for log market wages. Generally the patterns for market wages and total wages are broadly similar. Returns are higher for both men and women in the Northeast, relative to the South, throughout the education distribution apart from *negative* returns to the first year of secondary school (and a small return to three years of elementary school for women). Towards the top of the education distribution, therefore, wages in the Northeast and South are about the same; the gaps between men and women, however, persist.

Returns to education in the self-employment sector, in Figure 3, are quite different. For men, relative to the Northeast, returns are higher in the South for elementary and secondary school; returns (and wages) in the Center-North track those in the South very closely. Men working in the self-employment sector in the Northeast seem not to do very well. In contrast, apart from a negative return to the second year of secondary school, women in the Northeast enjoy a higher return to every year of schooling (less than fourteen) relative to women in the South. Thus, in the Northeast, among those with no education who choose to work in the self-employment sector, women earn only about one third a man's wage; at the top of the education distribution, however, the wages of men and women are about the same.

Clearly, if education is simply a credential (or signal) then there should be no steps in the wage functions for the self-employed (who do not need the noisy signal provided by the education system). Among self-employed men in the South, returns decline in the fifth, seventh and fourteenth year at school. In the Northeast, they decline in the fourth, sixth, eleventh and thirteenth year. It is very difficult to see how these can be explained by credentialism. Notice also that among males who are self-employed in the Northeast, those who complete the first part of elementary school and those who complete secondary school earn *less* than those who had dropped out *before* graduating. We suspect, therefore, that the steps in the wage functions reflects, at least in part, heterogeneity in the types of people who leave school prior to completing a level. This suggests that the family background variables in the data, while capturing an important part of the story, do not fully control for these effects.

Returns in the market and self-employment sectors are compared in Figure 4. In the South, returns to education and wages are about the same for men in both sectors. For women, returns in the market sector are higher for elementary school and so wages there are higher among women with at least some secondary schooling. Women in the Northeast earn about the same wage, and return, in both sectors although there is more noise in the self-employment function. Among men in the Northeast, there are important differences. At the bottom of the education distribution, wages in self-employment are higher but returns to education are lower than in the market sector. The market sector returns take off at the end of elementary school and so wages are higher in the market sector at higher levels of education.

There are, therefore, large differences in the returns to education across regions: this is, by now, very well understood.[15] There are also significant differences across gender. Little attention, however, has been paid to the substantial differences in returns across the market and self-employment sectors and, perhaps most importantly, to how all these differences vary over the distribution of education. Smoothing over the differences (with parametric functional forms) seems to miss an important dimension of the wage-education function. A piece-wise linear spline function would fit these data well: but that becomes clear only after examining these semi-parametric estimates.

(b) Wage functions: cohort and race comparisons

In this section, we focus only on the Northeast and examine differences in returns after stratifying the sample on age cohorts, in the first case, and on race in the second. Estimated conditional wage functions are presented in Figure 5 for three age groups: younger (25-34 years), middle (35-44) and older (45-60): generally, returns to education tend to decline with age.

In the market sector, for the first half of elementary school, the middle and older groups earn the same returns which are about double those of the younger cohort. In the second half of elementary school, however, the returns for the older cohort take off and so the average return to the first years of schooling is about 13% for the older cohort, about three-quarters of that for the middle cohort (9.3%) and only half that for the younger cohort (6.2%).[16] After elementary school, returns for the younger cohort take off and they more than double (to 13.9% for secondary school and 17% for four years post-secondary

---

[15]See, for instance, Behrman and Birdsall (1983), Birdsall and Behrman (1984) and Birdsall (1985).

[16]For the older cohort, the return to the eighth year of school is negative: the average return to the first seven year is a massive 17.5%.

school). Since returns for the middle cohort also increase and those for the older are slightly lower, inter-cohort differences in wages are narrower at the top of the education distribution than at the bottom. The patterns for women are similar with the older cohort enjoying much higher returns to the second half of elementary school but returns for the younger and middle cohort take off during secondary school.

In the self-employment sector, the differences are less obvious. The middle cohort seems to do worst and earns lower returns than both the younger and older cohorts for the first half of elementary school although they catch up to the younger (but not older) cohorts by the beginning of secondary school.

Returns to schooling are apparently changing dramatically over time in Brazil. As education opportunities have expanded and more children have completed elementary school, economic returns to the first years of schooling have declined. Going beyond primary school, however, reaps large returns and these returns are greatest for the most recent cohorts: this has very important implications for the path of income inequality in Brazil. (Lam and Levison, 1987; Barros, 1989).

Conditional expectations of log wages are presented in Figure 6 based on regressions estimated separately for blacks, mulattos and whites (including Asians). In the market sector, returns to education (and wages) are about the same for men in the three groups who have four years or less of schooling but the returns to the second half of elementary school are higher for whites and by secondary school, returns for mulattos have also taken off, leaving the blacks behind. In the self-employment sector, there are no significant differences in returns to education for men. This is also true for women in the market sector but not in the self-employment sector where whites are consistently better rewarded for education than either mulattos or blacks (although these estimates are quite imprecise). Differential returns by race, it appears, is greatest for men in the market sector and is significant in the upper half of the education distribution; for women, it is apparent only in the self-employment sector and persists throughout the education distribution.

(c) Wage functions: effects of background on returns to education

All the estimated returns to education discussed so far have included controls for the background of the worker including parents's education.[17] We have also included controls for spouse's education (permitted to affect wages in a non-parametric form) and spouse's parents' education. While not usually included in wage or earnings functions (see Lam and Schoeni, 1990, for an exception), spouse and spouse parents characteristics may affect wages in several ways. Self-employment wages include the return to self-employment capital, which is likely to be enhanced in households with better access to resources. Market wage returns may also be affected, if, for example, there is job queueing or other forms of market segmentation. Given that we are not able to measure job-specific experience, returns to women are likely to depend on the degree of labor force attachment, which may be related to both husband's and possibly parents-in-law characteristics. Finally, these variables will in part pick up effects of marital sorting, together with the fact that own parents' characteristics only imperfectly measure family background.

Many studies have compared returns to education with and without these sorts of controls and demonstrated substantial declines in estimated returns. How do estimated returns in Brazil change when background is added to the regressions? The results are presented in Figures 7 and 8 (with solid lines

[17]And a dummy for those mothers and fathers whose education is not reported.

representing those regressions which include background and are discussed above). The comparison for males (ignoring regional heterogeneity) is discussed in Lam and Schoeni, (1990).

Estimated returns with background controls are substantially smaller in all cases and decline by over 20% in many cases. For women in the South, the differences in returns are about 15% and are approximately constant across the education distribution in the market sector but the differences increase with education in the self-employment sector (from about a 10% difference for one year of education to over 30% at high levels of education). In the Northeast, the differences rise with education in the market sector and possibly also in the self-employment sector.

Similarly, among men who are self-employed, the returns to background are biggest among the best educated in both the South and Northeast. For men working in the market sector, however, the differences are a nonlinear function of education; they are smallest at the bottom of the education distribution, increase until the beginning of secondary education and then the differences tend to decline.

There are, it appears, quite different patterns to the role of background on labor market outcomes depending on the gender, region and sector of employment of workers as well as their levels of education. Again, these results have obvious implications for the path of income inequality in Brazil.

(d) Wage functions: direct effects of background

Whereas in those regressions which include controls for background, there are significantly higher returns to education for women, at least in the upper part of the education distribution, these differences disappear (in significance) when background controls are not included. This suggests that there are interesting differences in the allocation of resources by households to men and women. We turn, therefore, to discuss the direct effects of family background on wages, paying particular attention to differences in the effects of the education of fathers and mothers (or fathers-in-law and mothers-in law) on the n wages of sons and daughters. Table 1 reports estimates for the three macro-regions, separately for market and self-employment wages.

In the Northeast we see that mother's education has large, significant positive effects on the n market wages of daughters. Daughters with mothers having greater than elementary school education enjoy a 30 percent market wage advantage over daughters of illiterate mothers. Although this effect is small compared to the own education effect, it is still sizable. While mother's education does affect the wages of sons, these effects are significantly smaller than those on daughters. This suggests that, in the Northeast, mothers have a bigger influence on the labor market outcomes of their daughters relative to their sons. Of course, we cannot tell whether this reflects differences in time or financial resource allocations, the effects of role models, or of preferences: we can tell, however, that educating women today is likely to have a bigger influence on wages of women, relative to men, in the next generation.

By contrast, in the South, the influence of mother's education is somewhat larger for sons than daughters, although the differences are jointly significant at only the .10 level (and individually significant for the best educated). In the Center-North, mothers with at least four years of education have significant, positive effects on the wages of both their sons and daughters.

Father's education has a significant impact on both sons' and daughters' log market wages in all three regions, but the differences are not statistically significant, except at the top of the education distribution. It has been noted that relative to fathers', mothers' education has a bigger effect on labor market outcomes of children;[18] this is also true in the Center-North, the Northeast and among sons in the South. It is not, however, a general fact. Among daughters in the South, the reverse is true: the influence of fathers' education is greater than that of the mother.

The effects of parental education on self-employment wages are different. In the Center-North, fathers have a significantly positive impact on the wages of sons, but not on daughters' and this difference in significant. Once again, there is evidence that parents tend to allocate more resources towards children of the same gender, in this case fathers to sons. The magnitudes of the self-employment wage differentials over having an uneducated father are larger in the Center-North than for market wages, ranging from 6 percent at low levels of education to 20 percent at high levels. Better educated fathers in the Northeast also have a positive influence on wages of sons, but not daughters. Mothers in the South and Center-North have strong, positive effects on self-employment wages of both sons and daughters, differentials ranging from 15 to almost 40 percent at higher levels.

Lam and Schoeni (1990) point out that the education of parents also has a dramatic effect on the earnings of sons-in-law. It turns out that the effects on (the wages of) daughters-in-law are also dramatic. In addition, these influences vary across gender, region and sector of employment.

Father-in-law's education has positive, significant effects on market wages of males in the Northeast, but not on the market wages of females; this difference is significant for education over 4 years. In the South, fathers-in-law also have significantly positive effects on market wages of males, not

[18]Heckman and Hotz (1986) present results for earnings of men in Panama; Behrman and Wolfe (1984) present results for schooling achievement and household income for women in Nicaragua.

on females and, again, these differences are significant. We know from the first panel of Table 1, that the education of a father significantly affects the wages of a daughter (typically more than a son). Fathers, it appears, also allocate resources towards their daughters' husbands but not to their sons' wives. Furthermore, the magnitude of the effects of the education of fathers-in-law on n wages of men is of the same order of magnitude as the effect of the man's *own* father (ranging from 10 percent at 1-3 years of schooling to 20 percent at greater than 8 years).

Mothers-in-law have positive impacts on market wages of both men and women in the Northeast, though the differences across gender are not significant. Effects of mother-in-laws are significant at lower levels of education for women and at higher levels for men, magnitudes ranging from 10 to 20 percent for the significant effects. In the Center-North, mother-in-laws have their largest effects on market wages of men, not women, particularly at lower levels of education.

For the case of self-employment wages, there are strong effects of mother-in-law's education, at higher levels, on men's wages in all three regions. The magnitudes are largest in the Northeast (20 to 50 percent at higher levels of father-in-law education) and smallest in the South. In the Center-North the differences between the effects on male and female wages are significant at greater than three years of education; they are significant at greater than eight years in the Northeast. There exist little in the way of influence of father-in-law's education on self-employment wages in any of the regions.

Family background clearly plays an important role in determining labor market outcomes. This role varies significantly by region, by sector of employment and by gender of the child. It also varies, in a non-linear fashion, by the education of the parent. We turn next to study the Northeast in more detail, reporting family background effects, first, by race and, second, by age cohort. Table 2 reports results for market wages, stratified by race: whites (including Asians), blacks and mulattos. Dramatic differences are immediately apparent: gender differentials appear for whites, but not for blacks.[19] Among whites, there is a strong positive impact of mother's education on the market wages of daughters, but none on the wages of sons. These differences are significant. The magnitudes of the wage differentials on daughters is 22 percent for those whose mothers are literate, rising to almost 50 percent for daughters of mothers with greater than elementary schooling. Among mulattos there exists some evidence for effects of mother's education above grade three on both male and female market wages, with magnitudes from 11 to 19 percent.

Father's education has an effect on the market wages of white males and mulatto females only at levels above eight years. No other effects are discernable. Among blacks, we find essentially no effects of either parent's education on market wages of either gender.

For whites the influence of a father's-in-law education is significantly positive (from 15 to 20 percent differentials) for males, but not for females; these differences are also significant. Among mulatto men, there is a positive, significant impact of father-in-laws' education, whereas there exists no discernable effect on wages of women.

The education of a mother-in-law has a positive influence on the market wages of white females, but the differences between these and the effects on male wages are not significant. For mother-in-laws of mulattos, the results are mixed, with some influence detectable on male wages. For blacks, there exist

---

[19]See Thomas (1990) for contrasting results on the influence of parental education on child health in the United States.

no discernable effects of parents-in-law, just as there are none detectable for parents.

In Table 3 we report family background results for three age groups, 25-34, 35-44 and 45-60. The main result is that the differences by gender in family background effects on market wages occur for the younger two groups. For the oldest group, no important differences arise.

Father's education affects male, but not female wages for the youngest group, 25-34 year olds. Market wage differentials range from 6.5 percent for sons of fathers with 1-3 years of schooling to 20 percent for those whose fathers have greater than an elementary level. There exist positive effects on both male and female wages of the 45-60 year old group, but no significant differences emerge. For mother's education, there exist strong, positive effects at greater than three years on female wages of the youngest age group. For 35-44 year olds we find strong effects of mother's education on daughters' market wages at all maternal education levels, with significant differences between female and male wage effects. Wage differentials over daughters of uneducated mothers are in the range of 15 to 30 percent for both of these age groups.

Among the effects of parents-in-law, we find positive effects of education of fathers-in-law on male market wages for both 25-34 and 35-44 year olds, with no effects on female wages for either group. The differences are significant for the latter age group. In contrast, education of mothers-in-law has positive influence on wages of both men and women aged 25-44, the effects being a little larger for women. For the older age group; there are, again, no important impacts of either parent-in-law on market wages.

The education of both own parents and the parents of a spouse affects labor market outcomes of workers in urban Brazil. The effect of background, however, is not the same on sons and daughters and differs for mothers and fathers (as well as parents-in-law). Excluding background from the wage functions, then we find there are no significant differences in returns to education for men and women. Including background, however, we find that women earn a greater return to education which is significant beyond elementary school. We also find that a mother's education has a bigger effect on the market wage of her daughter, relative to a son, in the Northeast whereas a father's education has a bigger effect on the self-employment wage of a son, relative to a daughter, in the Center-North. The influence of a father in law is apparently greater on men in both the South and the Northeast. Finally, we find these sorts of differences tend to dissipate with age of the worker.

2. CHILD EDUCATION OUTCOMES

While it is not possible to determine the underlying reason for these differences, they do suggest that there is a good deal to learn about resource allocation decisions within households. We turn next, therefore, to examining exactly this issue within the context of child schooling decisions.

References

Altonji, Joseph. 1988. The Effects of Family Background and Social Characteristics on Education and Labor Market Outcomes. mimeo.

Altonji, J. and T. Dunn. 1990. Effects of parental characteristics on the returns to education and labor market experience. Mimeo.

Armitage, Jane and Richard Sabot. 1987. Socioeconomic Background and the Returns to Schooling in Two Low-income Economies. *Economica*, 54: 103-108.

Barros, R. and D. Lam. 1990. Income inequality, inequality in education and the demand for schooling in Brazil. Mimeo.

Behrman, J. and N. Birdsall. 1983. The quality of schooling: Quantity alone is misleading. *American Economic Review*, 73:928-46.

Behrman, J. and N. Birdsall. 1984. Does Geographic Aggregation Cause Over Estimates of the Returns to Schooling? *Oxford Bulletin of Economics and Statistics*, 46: 55-72.

Behrman, J. and N. Birdsall. 1985. The quality of schooling: Reply. *American Economic Review*, 75:1202-5.

Behrman, J., N. Birdsall and R. Kaplan. 1990. The quality of schooling in Brazil and labor market oucomes: some further explorations. Mimeo.

Behrman, Jere, Z. Hrubec, Paul Taubman and Terrence Wales. 1980). *Socioeconomic Success: A Study of the Effects of Genetic Endowments, Family Environment and Schooling*. Amsterdam: North Holland.

Behrman, J. and B. Wolfe. 1984. The socioeconomic impact of schooling in a developing country. *Review of Economics and Statistics*, 66.2:296-303.

Bhuyia, A., B. Yojtniak, S. D'Souza and S. Zimicki. 1986. Socio-economic determinants of child nutritional status: biys versus girls, *Food and Nutrition Bulletin*, 8.

Birdsall, Nancy. 1985. Public Inputs and Child Schooling in Brazil. *Journal of Development Economics*, 18:67-86.

Boissiere, M., John Knight and Richard Sabot. 1985. Earnings, Schooling, Ability and Cognitive Skills. *American Economic Review*, 75: 1016-1030.

Bound, John, Zvi Griliches and Bronwyn Hall. 1986. Wages, Schooling and IQ of Brothers and Sisters: Do Family Factors Differ? *International Economic Review*, 27: 77-105.

Carnoy, Martin. 1967. Earnings and Schooling in Mexico. *Economic Development and Cultural Change*, 15:408-419.

Card, D. and A. Krueger. 1990. School quality and B;ack/White relative earnings: A direct assessment. Mimeo, Princeton University.

Desai, S., P. L. Chase-Lansdale and R. T. Michael. 1989. Mother or market? Effects of maternal employment on the intellectual ability of 4-year-old children, *Demography*, 26.4:545-63.

Dos Reis, Jose and Ricardo Paes de Barros. 1989. Income Inequality and the Distribution of Education: A Study of the Evolution of Brazilian Regional Differences in Inequality. processed, Department of Economics, Yale University.

Featherman, D. and R. Hauser. 1978. *Opportunity and Change*. New York: Academic Press.

Fields, G. 1977. Who Benefits from Economic Development? A Reexamination of Brazilian Growth in the 1960's. *American Economic Review*, 67: 570-582.

Fishlow, A. 1972. Brazilian Size Distribution of Income. *American Economic Review*, 62: 391-402.

Griliches, Z. 1977. Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica*, 45: 1-22.

Griliches, Z.. 1979. Sibling Models and Data in Economics: Beginnings of a Survey. *Journal of Political Economy*, 87: S37-S64.

Griliches, Zvi and William Mason. 1972. Education, Income and Ability. *Journal of Political Economy*, 80: S74-S103.

Haller, A and H. Saraiva. 1986. Income and education: Brazil, 1982. Mimeo, University of Wisconsin, Madison.

Hauser, R. and W. Sewell. 1986. Family Effects in Simple Models of Education, Occupational Status and Earnings: Findings from the Wisconsin and Kalamazoo Studies. *Journal of Labor Economics*, 4: S83-S115.

Heckman, James J.. 1974. Shadow Prices, Market Wages and Labor Supply. *Econometrica*, 42:679-694.

Heckman, J.. 1979. Sample Selection Bias as Specification Error. *Econometrica*, 47:153-162.

Heckman, J. and J. Hotz. 1986. The sources of inequality for males in Panama's labor market. *Journal of Human Resources*, 21.4:507-42.

Heckman, J. and R. Robb. 1985. Alternative Methods for Evaluating the Impact of Interventions. In J. Heckman and B. Singer eds., *Longitudinal Analysis of Labor Market Data*. Cambridge: Cambridge University Press.

Heckman, J. and G. Sedlacek. 1985. Heterogeneity Aggregation and Wage Functions: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, 93:1077-1125.

Hess, B. and J. Waring. 1978. Parent and child in later life: Rethinking the relationships. In R. Lerner and G. Spanier (eds.) *Child Influences on*

Marital and Family Interactions: A Life-Span Perspective, 241-73. New York:Academic Press.

Hoddinott, J. 1990. Rotten kids or manipulative parents: Are children old age security in Western Kenya? Mimeo, Oxford University.

King, E. and R. Bellwew. 1989. Gains in the education of Peruvian women in 1940's-1980's -- Patterns and explanations. Mimeo. PHR, World Bank.

King, E. and L. Lillard. 1987. Education policy and schooling attainment in Malaysia and the Philippines, *Economics of Education Review*, 6.2:167-81.

Lam, D. and D. Levison. 1990. Age, experience and schooling: Decomposing earnings inequality in the U.S. and Brazil. University of Michigan Population Studies Center Research Report.

Lam, D. and D. Levison. 1989. Declining inequality in schooling in Brazil and its effects on inequality in earnings. University of Michigan Population Studies Center Research Report 89-170.

Lam, D. and R. Schoeni. 1990. Effects of family background on earnings and returns to schooling: Evidence from Brazil. Mimeo, University of Michigan.

Lamb, M. E. 1976. The role of the father: an overview. In M. Lamb, (ed.), *The Role of the Father in Child Development*, New York:Wiley.

Lamb, M. E. 1987. *The Father's Role: Cross Cultural Perspectives*, London:Lawrence Erlbaum Associates.

Langoni, Carlos. 1973). *Distribuicao de Renda e Desenvolvimento Economico do Brasil*. Rio de Janeiro: Editora Expressao e Cultura.

Langoni, C.. 1977. Income Distribution and Economic Development: The Brazilian Case. in Michael Intriligator (ed.), *Frontiers of Quantitative Economics*, Volume B. Amsterdam: North Holland Press.

Leibowitz, Arleen. 1974. Home Investments in Children. in T. W. Schultz (ed.), *Economics of the Family*. Chicago: University of Chicago Press.

Medeiros, Jose. 1982. Alcance e Limitacoes de Teoria do Capital Humano: Diferencas de Ganhos no Brasil em 1973. Serie Ensaios Economicos, Volume 17, Instituto de Pesquisas Economicas, Sao Paulo.

Morgan, S. P., D. Lye and G. Condran. 1988. Sons, daughters and the risk of marital disruption, *American Journal of Sociology*, 94.1:110-29.

Papanicolau, J. and George Psacharapoulos. 1979. Socioeconomic Background, Schooling and Monetary Rewards in the United Kingdom. *Economica*, 46: 435-439.

Pastore, J. 1982. *Inequality and social mobility in Brazil*. Madison:University of Wisconsin Press.

Pastore, J. and H. Zylberstajn. 1990. Social mobility in Brazil 1973-82: The role of education in status determination. Mimeo, University of Sao Paulo.

Psacharapoulos, George. 1985. Returns to Education: A Further International Update and implications. *Journal of Human Resources*, 20: 583-604.

Psacharopolous, G. and Maureen Woodhall. 1985. *Education for Development: An Analysis of Investment Choices*. Oxford: Oxford University Press for the World Bank.

Sahn, David and Harold Alderman. 1988. The Effects of Human Capital on Wages, and the Determinants of Labor Supply in a Developing Country. *Journal of Development Economics*, 29: 157-184.

Saraiva, H, A. Pahari and A. Haller. 1986. Beyond credentialism: The productivity of ideology and the ideology of productivity. Mimeo, University of Wisconsin, Madison.

Schultz, T. P.. 1988. Education Investments and Returns to Economic Development. in Hollis Chenery and T. N. Srinivasan (eds.), *Handbook of Development Economics*, Volume 1. Amsterdam: North Holland Press.

Sedlacek, Guilherme and Ricardo Paes de Barros (eds.). 1989). *Mercado de Trabalho e Distribuicao de Renda: Uma Coletanea*. Instituto de Planejamento Economico e Social, Instituto de Pesquisas, Monografia No. 35, Rio de Janeiro.

Stelcner, Morton, Ana-Maria Arriagada and Peter Moock. 1987. Wage Determinants and School Attainment Among Men in Peru. Living Standards Measurement Study Working Paper No. 38. Washington D.C.: World Bank.

Spanier, G. B. and P. C. Glick. 1981. Marital instability in the United States: Some correlates and recent changes, *Family Relations*, 31:319-38.

Spitze, G. and J. Logan. 1990. Sons, daughters and intergenerational social suport. *Journal of Marriage and the Family*, 52:420-30.

Thomas, D. 1990a. Like father, like son: Gender differences in household resource allocations. Mimeo, Yale University.

Thomas, D. 1990b. Intra-household resource allocation: an inferential approach, *Journal of Human Resources*, 25.4:635-664.

Tinbergen, Jan. 1951. Remarks on the Distribution of Labor Incomes. *International Economic Papers*, 1:195-207.

Willis, R.. 1986. Wage Determinants: a Survey and Reinterpretation of Human Capital Earnings Functions. In Orley Ashenfelter and Richard Layard eds. *Handbook of Labor Economics*, Volume 1. Amsterdam; North Holland Press.

Willis, R. and Sherwin Rosen. 1979. Education and Self-Selection. *Journal of Political Economy*, 87: S7-S36.

[redacted]J. and H. Zylberstajn. 1990. Social mobility in Brazil 1973-82: The role [redacted]education in status determination. Mimeo, University of Sao Paulo.

ca[redacted]oulos, George. 1985. Returns to Education: A Further International [redacted]ate and implications. *Journal of Human Resources*, 20: 583-604.

opolous, G. and Maureen Woodhall. 1985. *Education for Development: An [redacted]lysis of Investment Choices*. Oxford: Oxford University Press for the [redacted]ld Bank.

David and Harold Alderman. 1988. The Effects of Human Capital on Wages, [redacted] the Determinants of Labor Supply in a Developing Country. *Journal of [redacted]elopment Economics*, 29: 157-184.

a, H, A. Pahari and A. Haller. 1986. Beyond credentialism: The productivity [redacted] ideology and the ideology of productivity. Mimeo, University of [redacted]consin, Madison.

z, T. P.. 1988. Education Investments and Returns to Economic Development. [redacted]Hollis Chenery and T. N. Srinivasan (eds.), *Handbook of Development [redacted]nomics*, Volume 1. Amsterdam: North Holland Press.

ek, Guilherme and Ricardo Paes de Barros (eds.). 1989). *Mercado de Trabalho [redacted]istribuicao de Renda: Uma Coletanea*. Instituto de Planejamento [redacted]nomico e Social, Instituto de Pesquisas, Monografia No. 35, Rio de Janeiro.

[redacted] Morton, Ana-Maria Arriagada and Peter Moock. 1987. Wage Determinants [redacted] School Attainment Among Men in Peru. Living Standards Measurement Study Working Paper No. 38. Washington D.C.: World Bank.

[redacted]G. B. and P. C. Glick. 1981. Marital instability in the United States: [redacted]e correlates and recent changes, *Family Relations*, 31:319-38.

[redacted]. and J. Logan. 1990. Sons, daughters and intergenerational social [redacted]ort. *Journal of Marriage and the Family*, 52:420-30.

. D. 1990a. Like father, like son: Gender differences in household resource [redacted]ocations. Mimeo, Yale University.

[redacted]D. 1990b. Intra-household resource allocation: an inferential approach, *Journal of Human Resources*, 25.4:635-664.

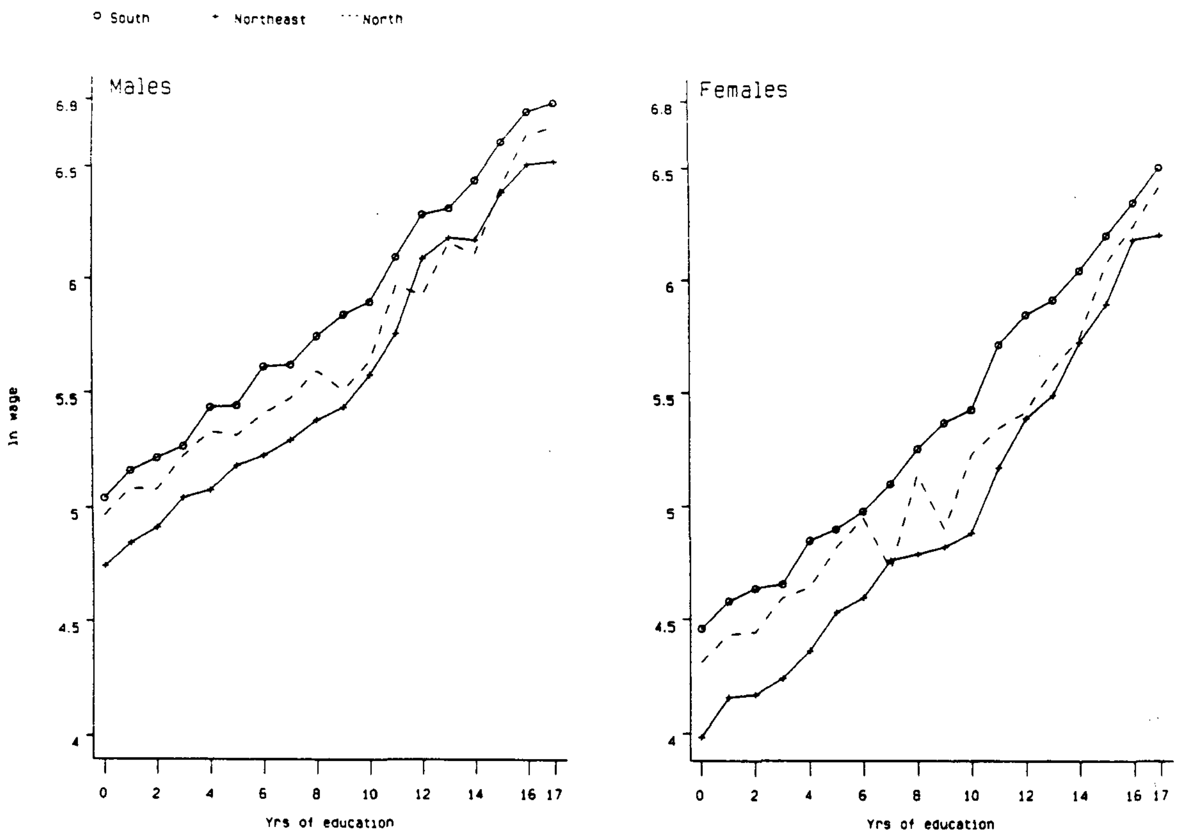
[redacted], Jan. 1951. Remarks on the Distribution of Labor Incomes. *International Economic Papers*, 1:195-207.

[redacted].. 1986. Wage Determinants: a Survey and Reinterpretation of Human [redacted]ital Earnings Functions. In Orley Ashenfelter and Richard Layard eds. *Handbook of Labor Economics*, Volume 1. Amsterdam; North Holland Press.

[redacted]. and Sherwin Rosen. 1979. Education and Self-Selection. *Journal of [redacted]itical Economy*, 87: S7-S36.

Figure 1
Returns to education : ℓn(total wage)

(a) by region



(b) by gender

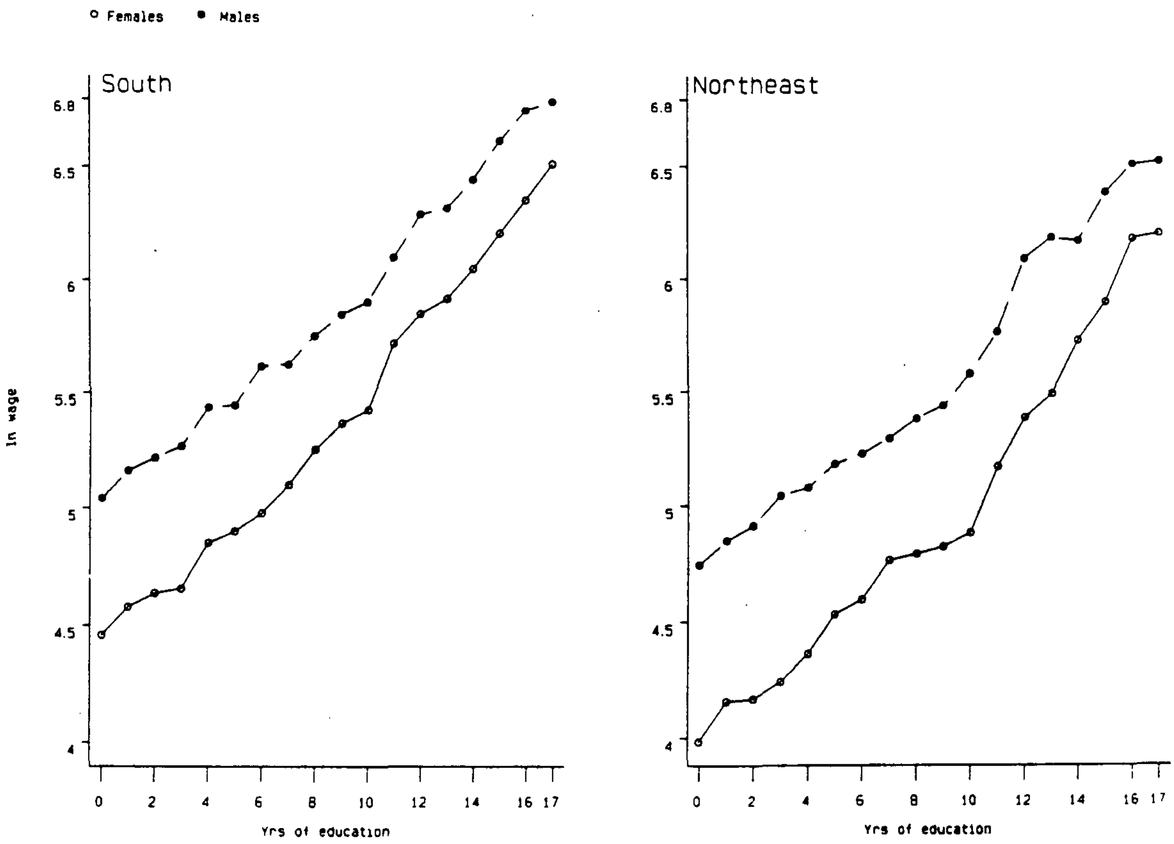
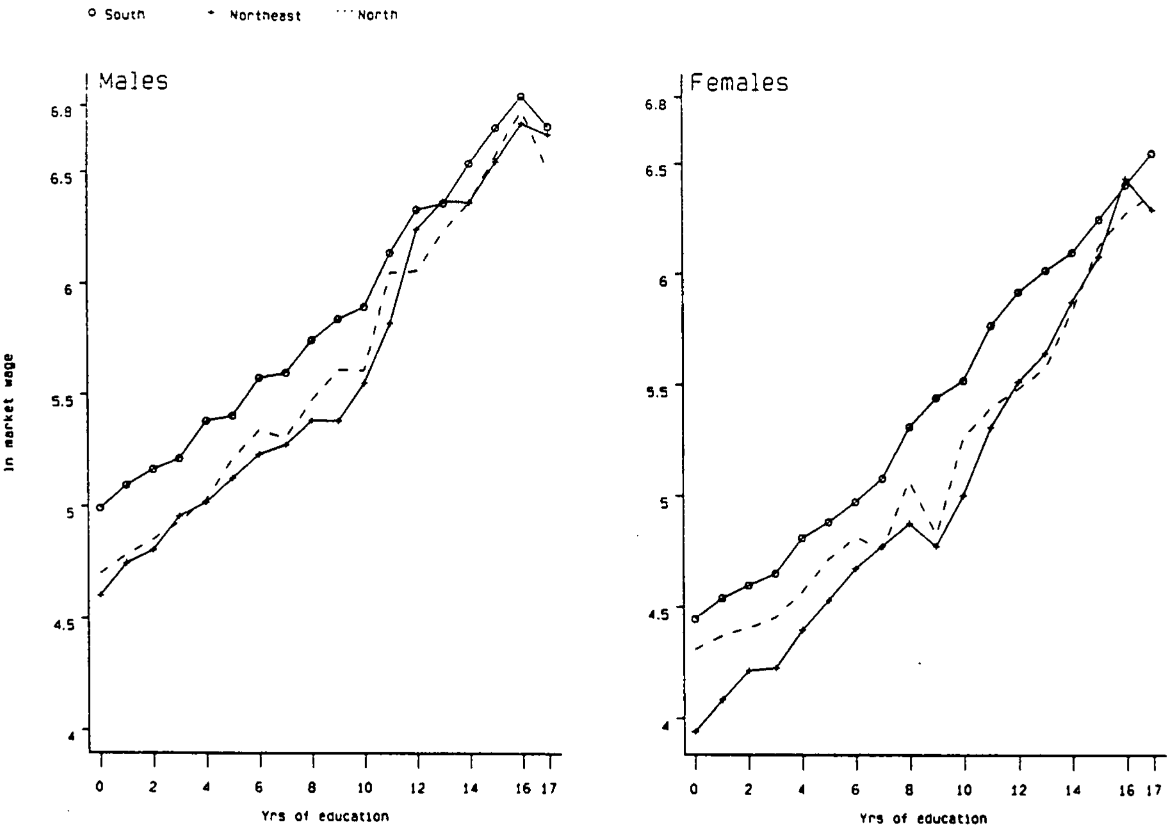


Figure 2
Returns to education : ℓn(market wage)

(a) by region



(b) by gender

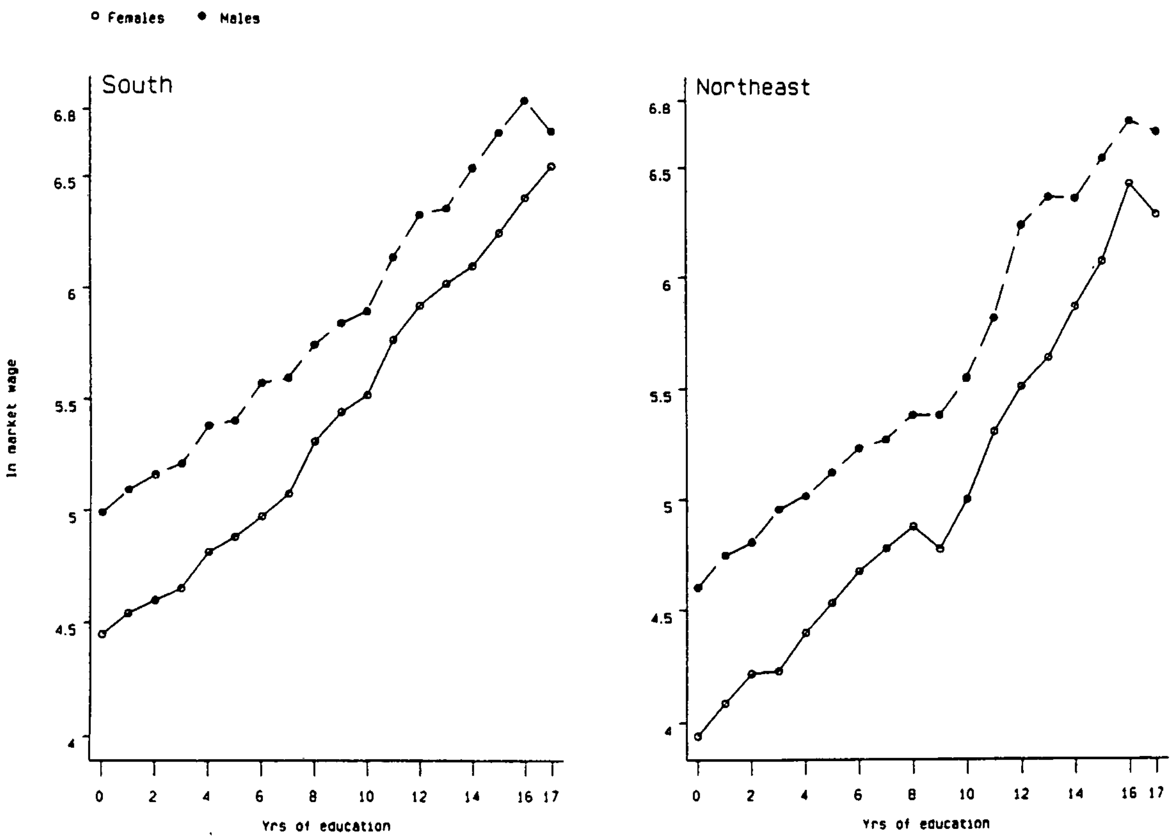
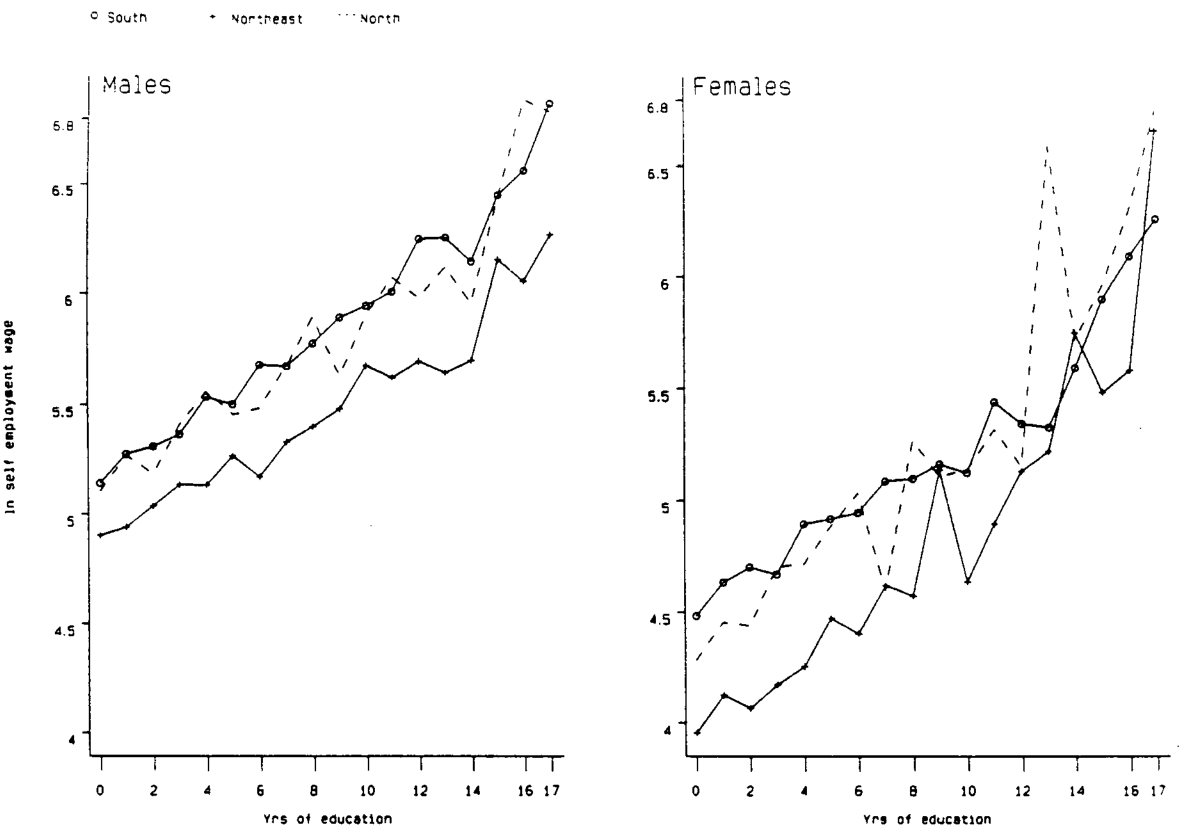


Figure 3
Returns to education : ℓn(self employment wage)

(a) by region



(b) by gender

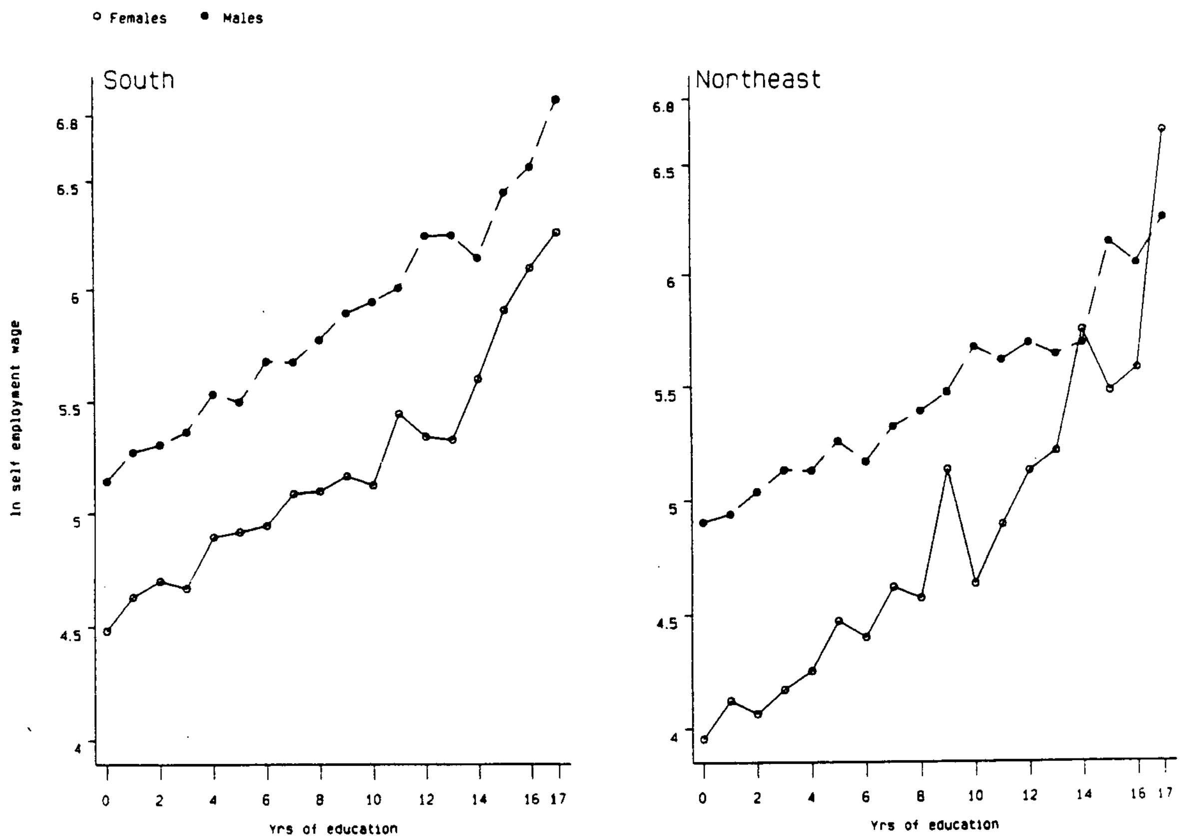


Figure 4
Comparison of returns to education in market and self-employment sectors

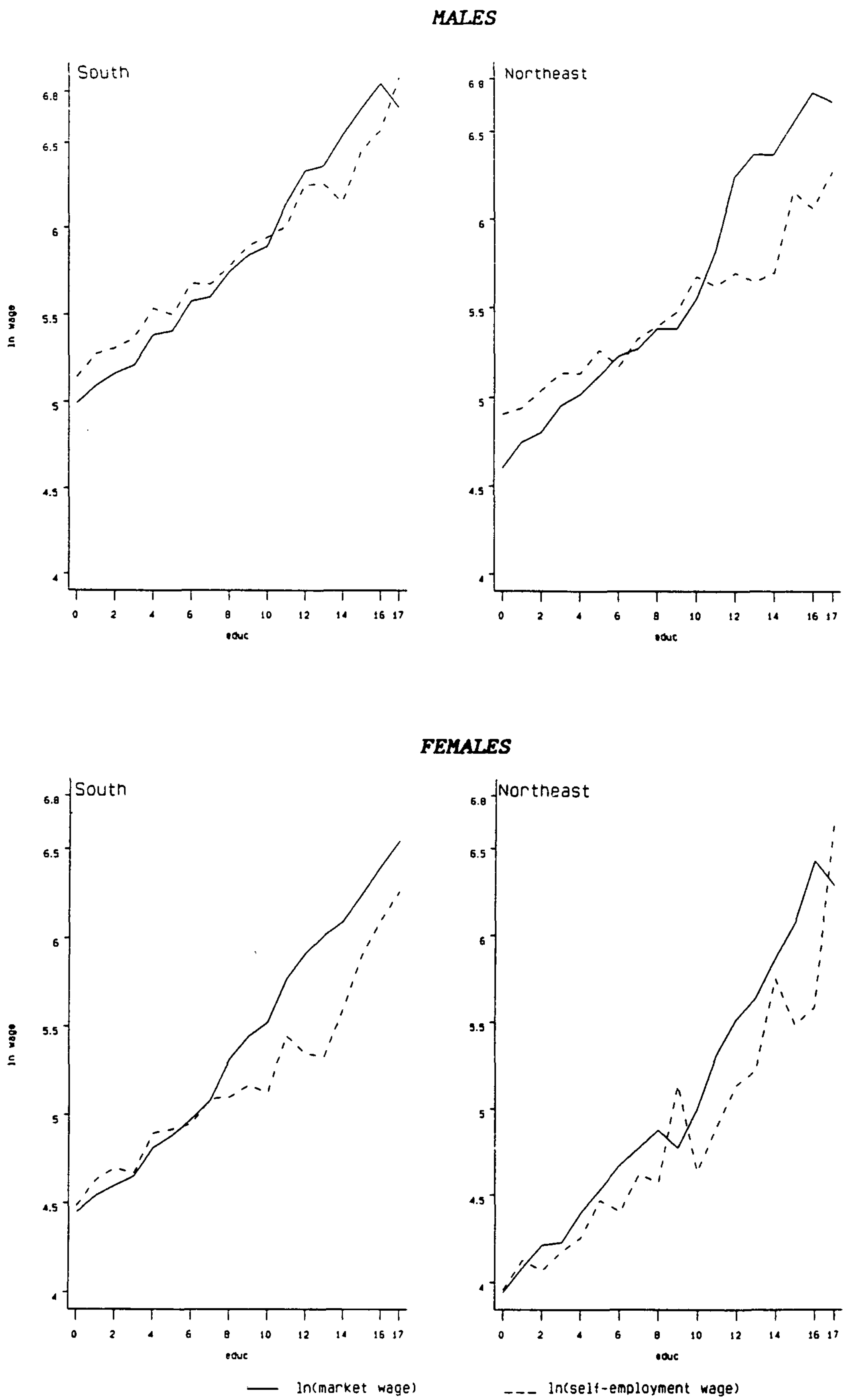
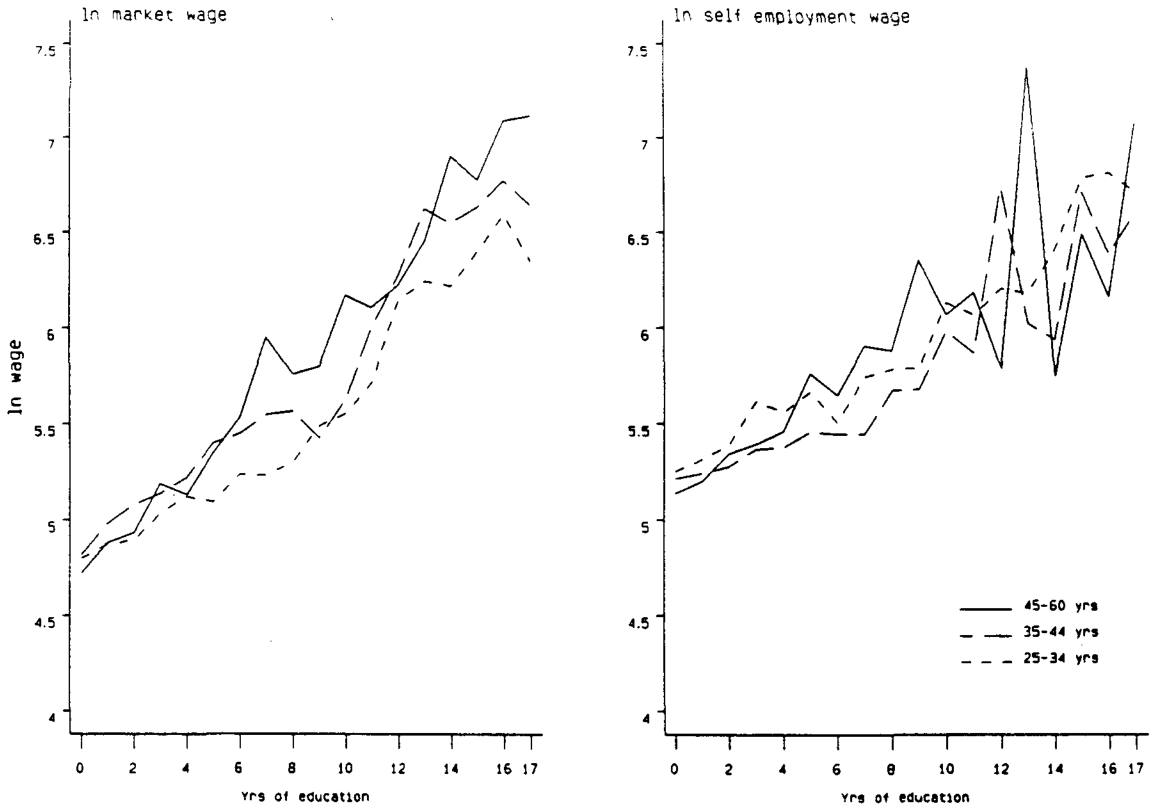


Figure 5
Returns to education : by age cohort

Northeast Males



Northeast Females

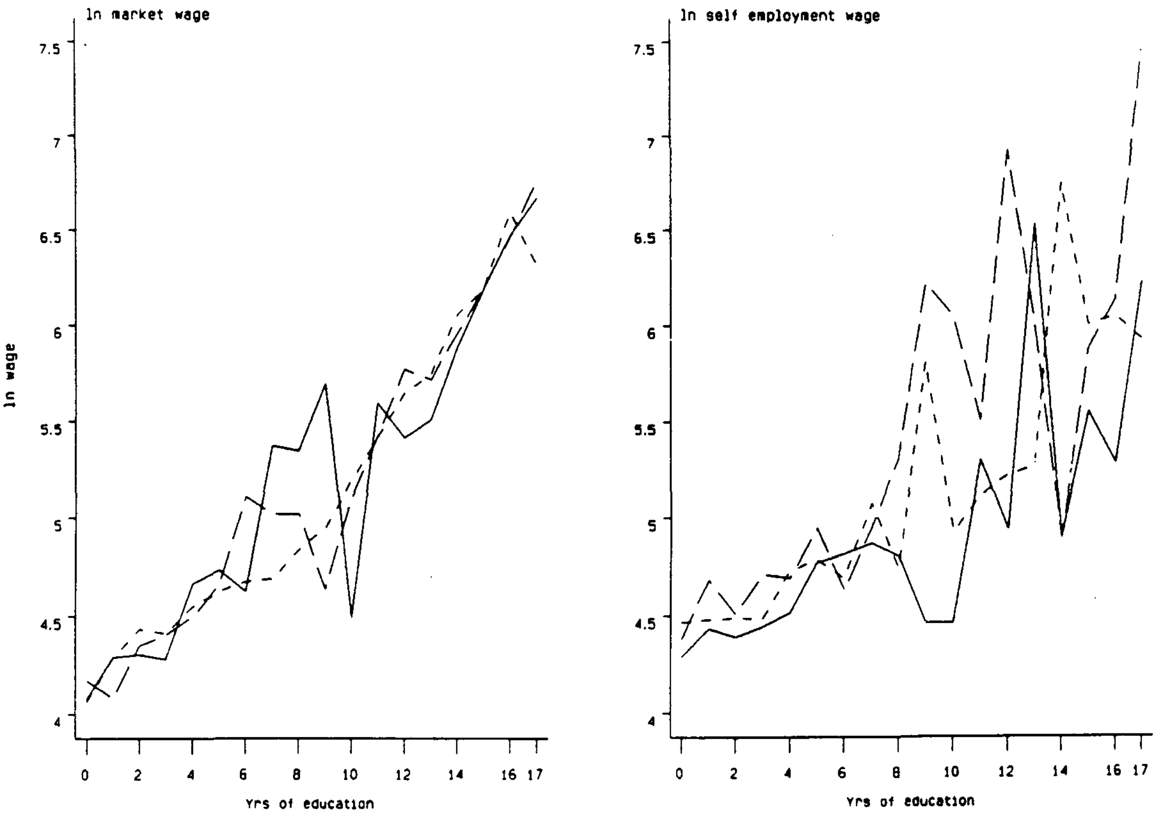
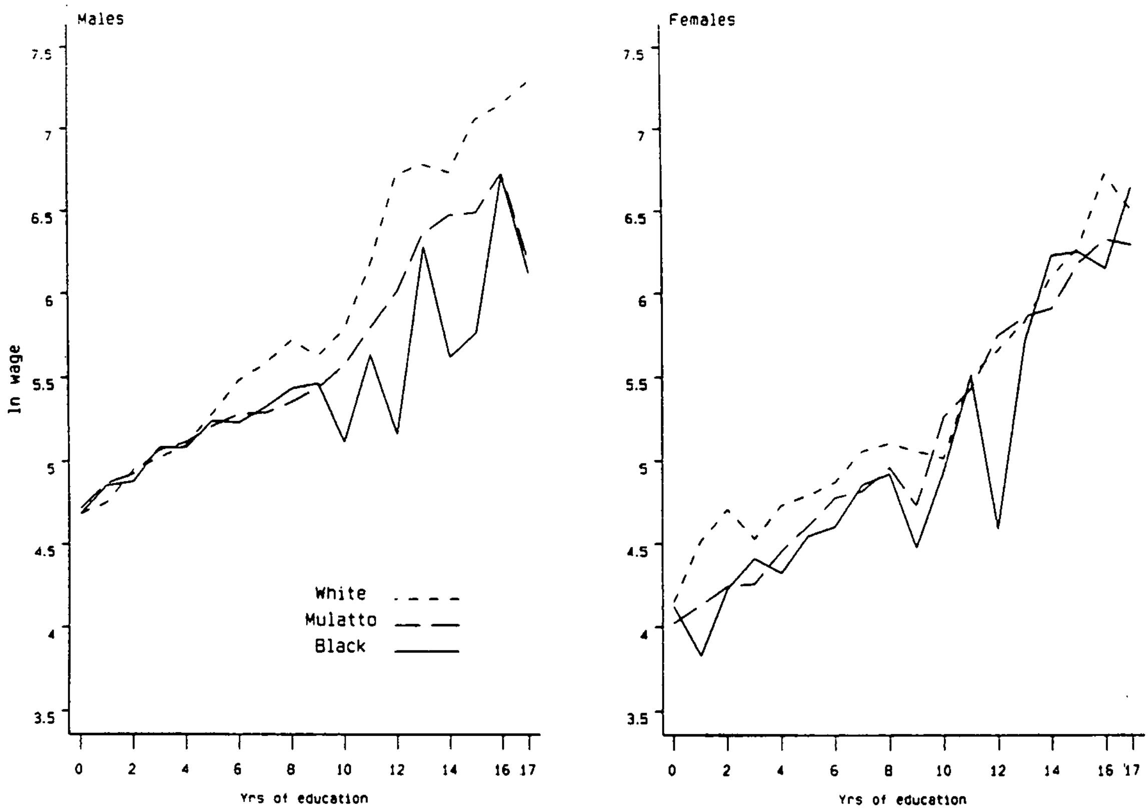


Figure 6
Returns to education: by race
Northeast

(a) ℓn market wages



(b) ℓn self-employment wages

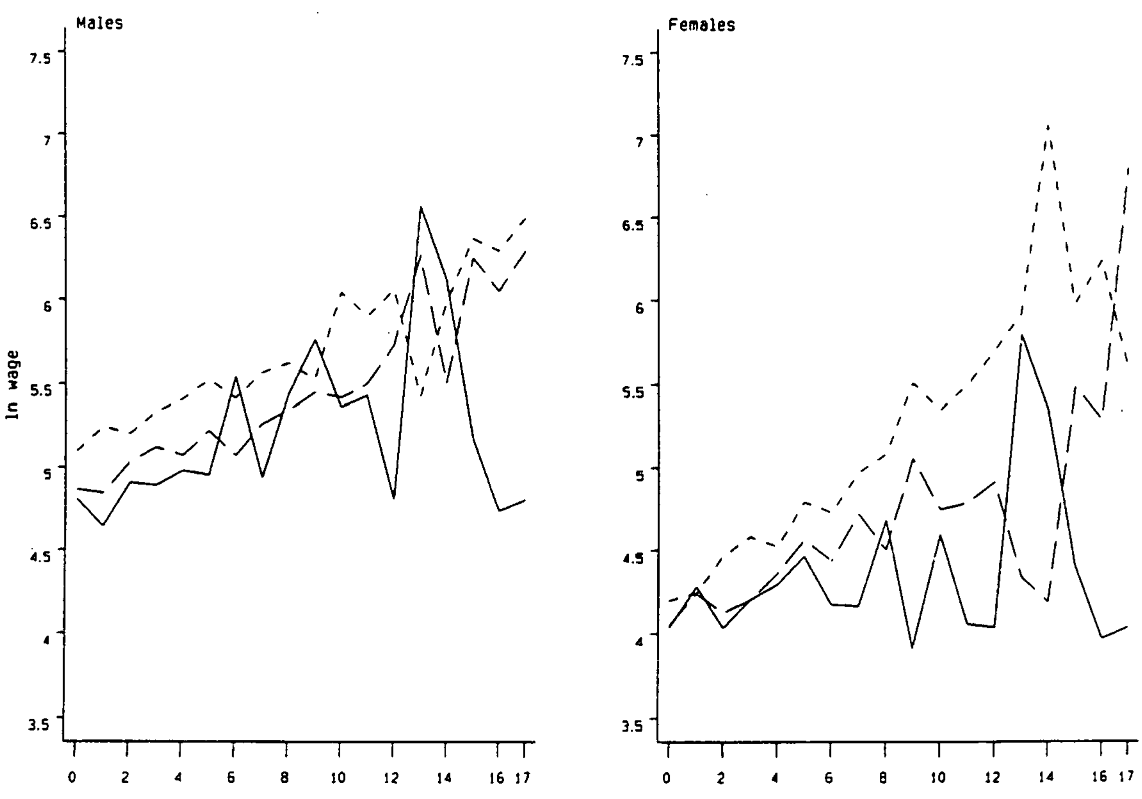
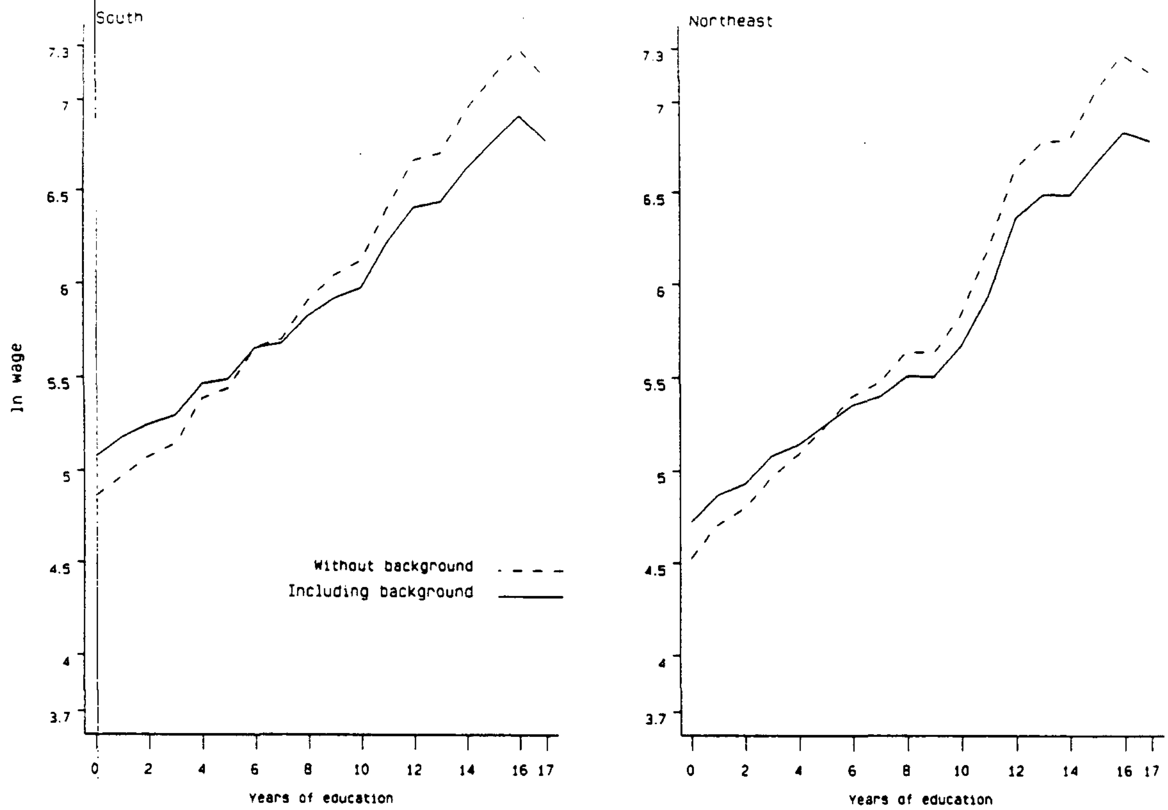


Figure 7

Returns to education: effect of controlling for spouse's and parents' education

ℓn market wages

(a) Males



(b) Females

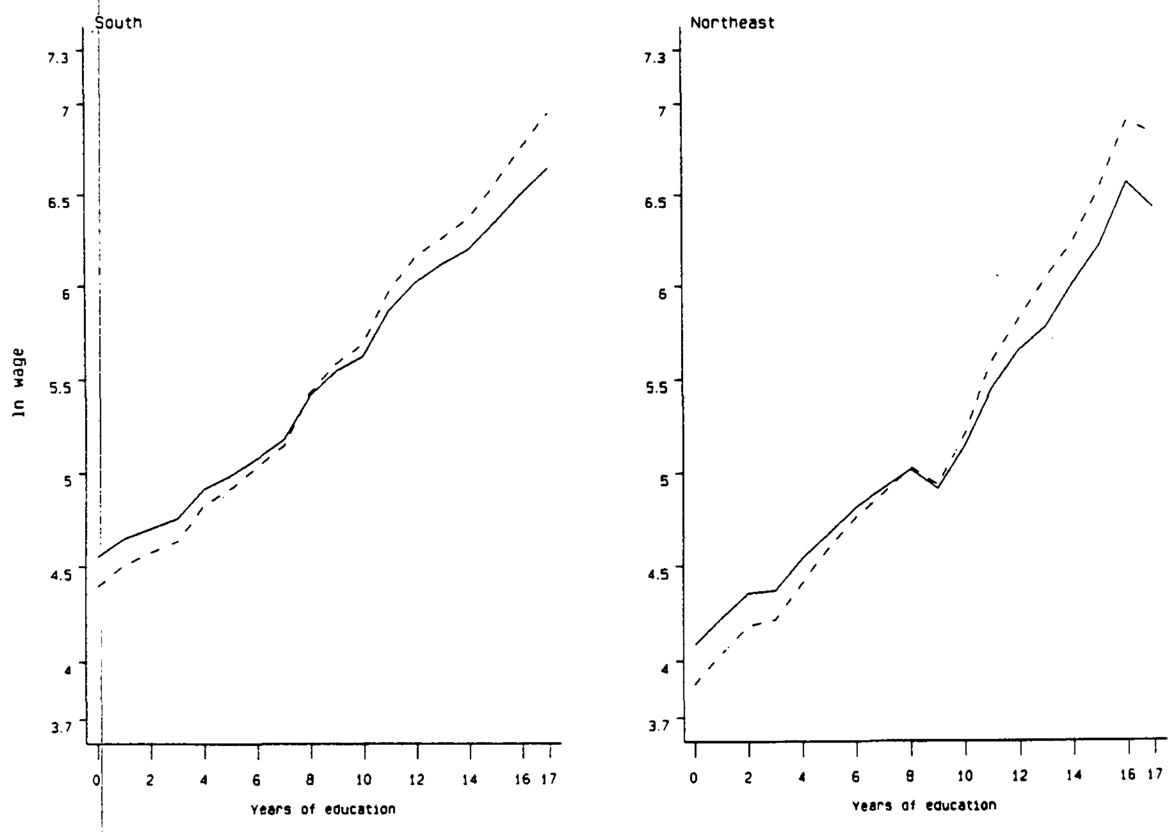
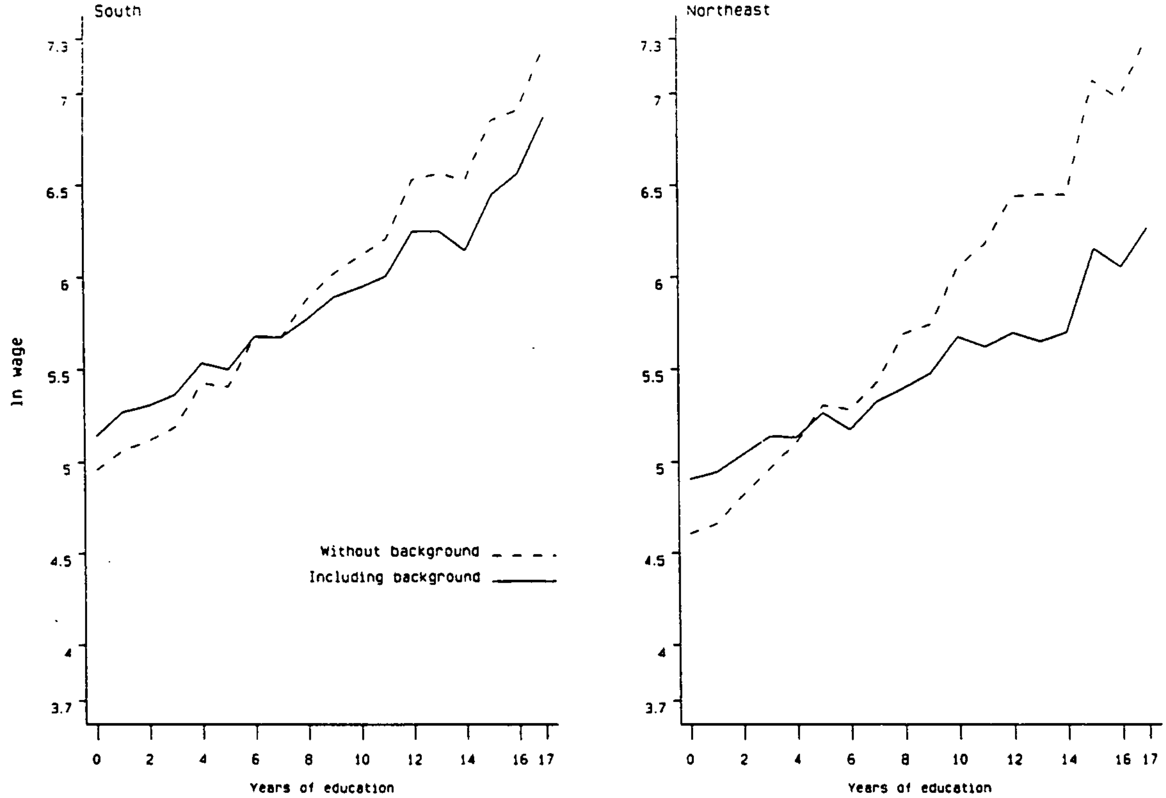


Figure 8
Returns to education: effect of controlling for spouse's and parents' education ln self employment wages

(a) Males



(b) Females

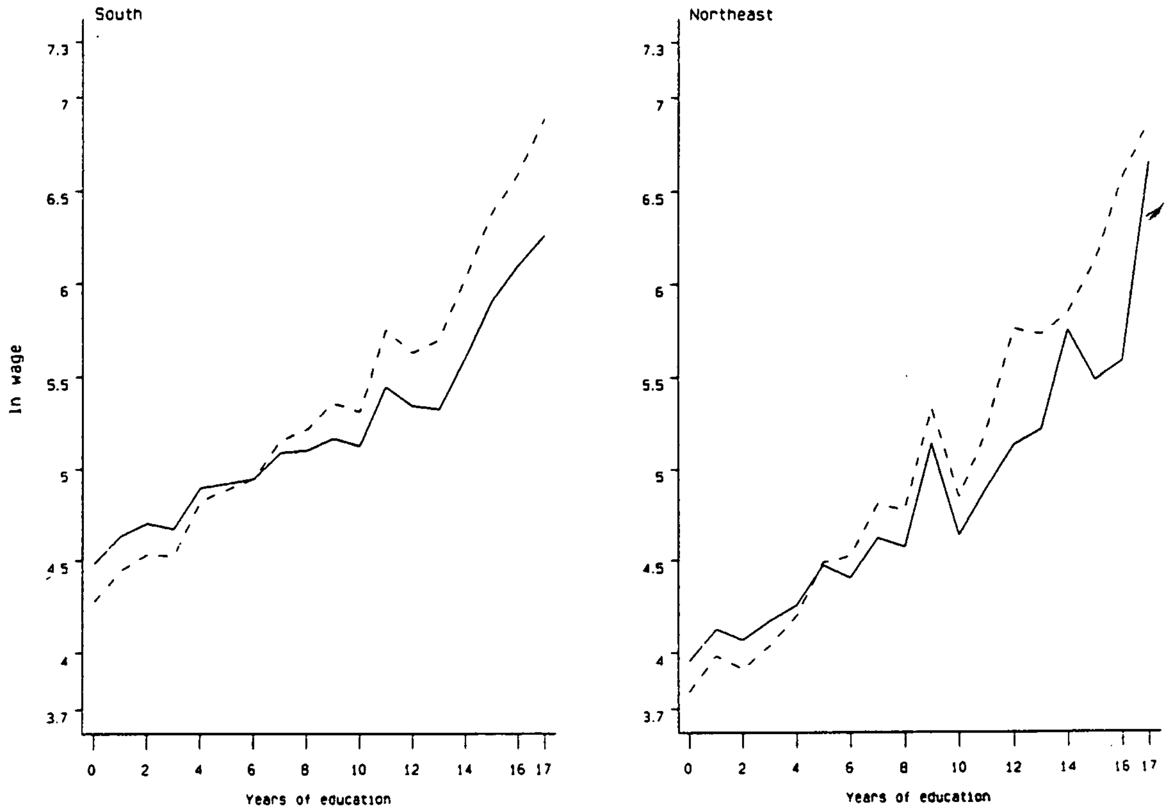


Table 1
Effects of parents' and spouse's parents' education on log(wages) of men and women

	SOUTH			CENTER-NORTH			NORTHEAST		
	Men	Women	Diff	**Men**	**Women**	Diff	Men	**Women**	Diff
ln market wage									
(1) if father									
literate	0.033	0.026	0.007	0.052	-0.022	0.074	-0.000	0.019	-0.019
	[2.34]	[1.15]	[0.27]	[1.90]	[0.51]	[1.46]	[0.00]	[0.46]	[0.40]
1-3 yrs education	0.036	0.047	-0.011	0.038	0.009	0.029	0.043	-0.022	0.065
	[2.64]	[2.22]	[0.45]	[1.18]	[0.18]	[0.51]	[1.57]	[0.49]	[1.25]
4-8 yrs education	0.079	0.127	-0.047	0.088	0.066	0.023	0.095	0.109	-0.014
	[5.15]	[5.53]	[1.72]	[2.39]	[1.24]	[0.35]	[2.77]	[2.16]	[0.24]
≥ 9 yrs education	0.126	0.198	-0.072	0.022	0.222	-0.201	0.229	0.154	0.075
	[5.30]	[6.13]	[1.81]	[0.36]	[2.88]	[2.07]	[4.57]	[2.14]	[0.85]
(1) if mother									
literate	0.041	0.025	0.017	0.041	0.071	-0.030	0.014	0.105	-0.091
	[2.94]	[1.09]	[0.63]	[1.57]	[1.76]	[0.62]	[0.59]	[2.58]	[1.90]
1-3 yrs education	0.053	0.019	0.034	0.015	0.053	-0.038	0.060	0.152	-0.092
	[3.95]	[0.93]	[1.38]	[0.48]	[1.18]	[0.70]	[2.16]	[3.47]	[1.76]
4-8 yrs education	0.113	0.070	0.043	0.153	0.150	0.003	0.104	0.222	-0.118
	[7.39]	[3.15]	[1.58]	[4.31]	[2.84]	[0.04]	[3.01]	[4.34]	[1.90]
≥ 9 yrs education	0.217	0.125	0.092	0.409	0.286	0.123	0.046	0.307	-0.261
	[8.35]	[3.69]	[2.17]	[6.46]	[3.52]	[1.20]	[0.83]	[3.89]	[2.67]
(1) if spouse's father									
literate	0.013	-0.038	0.051	-0.002	0.066	-0.067	0.034	-0.036	0.071
	[0.91]	[1.38]	[1.65]	[0.05]	[1.24]	[1.13]	[1.46]	[0.70]	[1.24]
1-3 yrs education	0.035	-0.026	0.061	0.004	0.108	-0.105	0.100	-0.013	0.113
	[2.66]	[0.98]	[2.08]	[0.12]	[1.89]	[1.62]	[3.92]	[0.22]	[1.74]
4-8 yrs education	0.070	0.015	0.055	0.054	0.051	0.003	0.138	-0.072	0.211
	[4.85]	[0.53]	[1.74]	[1.52]	[0.78]	[0.04]	[4.35]	[1.09]	[2.86]
≥ 9 yrs education	0.159	0.032	0.127	0.164	-0.005	0.169	0.209	-0.082	0.291
	[6.56]	[0.84]	[2.80]	[2.83]	[0.06]	[1.54]	[4.18]	[0.93]	[2.86]
(1) if spouse's mother									
literate	0.019	0.066	-0.047	0.076	-0.032	0.108	0.034	0.090	-0.057
	[1.32]	[2.45]	[1.54]	[2.82]	[0.63]	[1.89]	[1.44]	[1.80]	[1.02]
1-3 yrs education	0.031	0.070	-0.040	0.064	-0.049	0.113	0.078	0.138	-0.059
	[2.37]	[2.79]	[1.41]	[2.18]	[0.89]	[1.80]	[2.98]	[2.35]	[0.92]
4-8 yrs education	0.074	0.037	0.038	0.099	0.027	0.072	0.113	0.224	-0.11
	[5.17]	[1.35]	[1.22]	[2.85]	[0.42]	[0.99]	[3.52]	[3.35]	[1.49]
≥ 9 yrs education	0.097	0.007	0.090	-0.051	0.013	-0.064	0.129	0.181	-0.051
	[3.67]	[0.18]	[1.84]	[0.76]	[0.13]	[0.53]	[2.40]	[1.91]	[0.47]
ln(self-employment wage)									
(1) if father									
literate	-0.003	0.038	-0.041	0.068	-0.129	0.197	-0.003	0.047	-0.050
	[0.10]	[1.01]	[0.86]	[1.76]	[2.02]	[2.62]	[0.08]	[1.00]	[0.82]
1-3 yrs education	0.018	-0.029	0.047	0.101	-0.113	0.214	0.035	-0.009	0.044
	[0.66]	[0.79]	[1.01]	[2.26]	[1.52]	[2.45]	[0.78]	[0.16]	[0.60]
4-8 yrs education	0.046	0.015	0.031	0.083	0.036	0.047	0.152	0.066	0.085
	[1.49]	[0.38]	[0.61]	[1.50]	[0.42]	[0.46]	[2.61]	[0.85]	[0.86]
≥ 9 yrs education	0.118	0.259	-0.140	0.219	-0.126	0.344	0.386	0.142	0.243
	[2.52]	[3.77]	[1.67]	[2.28]	[0.81]	[1.87]	[3.86]	[1.02]	[1.38]
(1) if mother									
literate	0.107	-0.016	0.123	0.143	0.100	0.043	0.098	0.027	0.071
	[3.83]	[0.40]	[2.54]	[3.79]	[1.52]	[0.57]	[2.64]	[0.55]	[1.11]
1-3 yrs education	0.073	0.085	-0.012	0.152	0.162	-0.011	0.072	0.153	-0.081
	[2.74]	[2.35]	[0.28]	[3.38]	[2.22]	[0.13]	[1.54]	[2.61]	[1.06]
4-8 yrs education	0.109	0.204	-0.095	0.151	0.180	-0.029	0.079	0.167	-0.087
	[3.57]	[5.02]	[1.85]	[2.64]	[2.03]	[0.28]	[1.33]	[2.09]	[0.86]
≥ 9 yrs education	0.262	0.216	0.046	0.202	0.383	-0.181	0.179	0.059	0.120
	[4.98]	[2.76]	[0.48]	[1.73]	[1.81]	[0.74]	[1.61]	[0.37]	[0.60]
(1) if spouse's father									
literate	0.029	0.068	-0.039	-0.039	-0.005	-0.034	0.020	-0.053	0.073
	[0.99]	[1.50]	[0.72]	[1.00]	[0.06]	[0.39]	[0.58]	[0.89]	[1.04]
1-3 yrs education	0.039	0.047	-0.008	-0.037	0.048	-0.085	0.027	0.105	-0.078
	[1.46]	[1.06]	[0.15]	[0.85]	[0.51]	[0.81]	[0.63]	[1.45]	[0.90]
4-8 yrs education	0.034	0.093	-0.058	0.031	-0.099	0.131	0.055	-0.119	0.174
	[1.16]	[1.78]	[0.97]	[0.58]	[0.85]	[1.01]	[0.98]	[1.16]	[1.46]
≥ 9 yrs education	0.111	0.116	-0.005	0.119	0.353	-0.234	0.189	0.305	-0.115
	[2.34]	[1.42]	[0.06]	[1.32]	[1.74]	[1.05]	[1.91]	[1.88]	[0.59]
(1) if spouse's mother									
literate	0.032	0.028	0.004	0.056	0.037	0.019	0.045	0.038	0.007
	[1.10]	[0.63]	[0.07]	[1.45]	[0.49]	[0.22]	[1.24]	[0.61]	[0.09]
1-3 yrs education	0.046	0.023	0.022	0.106	0.120	-0.014	0.061	0.037	0.025
	[1.77]	[0.53]	[0.44]	[2.43]	[1.28]	[0.14]	[1.38]	[0.50]	[0.28]
4-8 yrs education	0.124	0.127	-0.002	0.231	-0.052	0.282	0.228	0.203	0.024
	[4.32]	[2.50]	[0.04]	[4.14]	[0.45]	[2.19]	[3.95]	[1.98]	[0.20]
≥ 9 yrs education	0.195	0.077	0.118	0.262	-0.269	0.531	0.473	0.028	0.444
	[3.72]	[0.84]	[1.11]	[2.42]	[1.24]	[2.17]	[4.28]	[0.15]	[2.02]

Table 2
Effects of parents' and spouse's parents' education
on log(wages) of men and women
Urban Northeast : by race

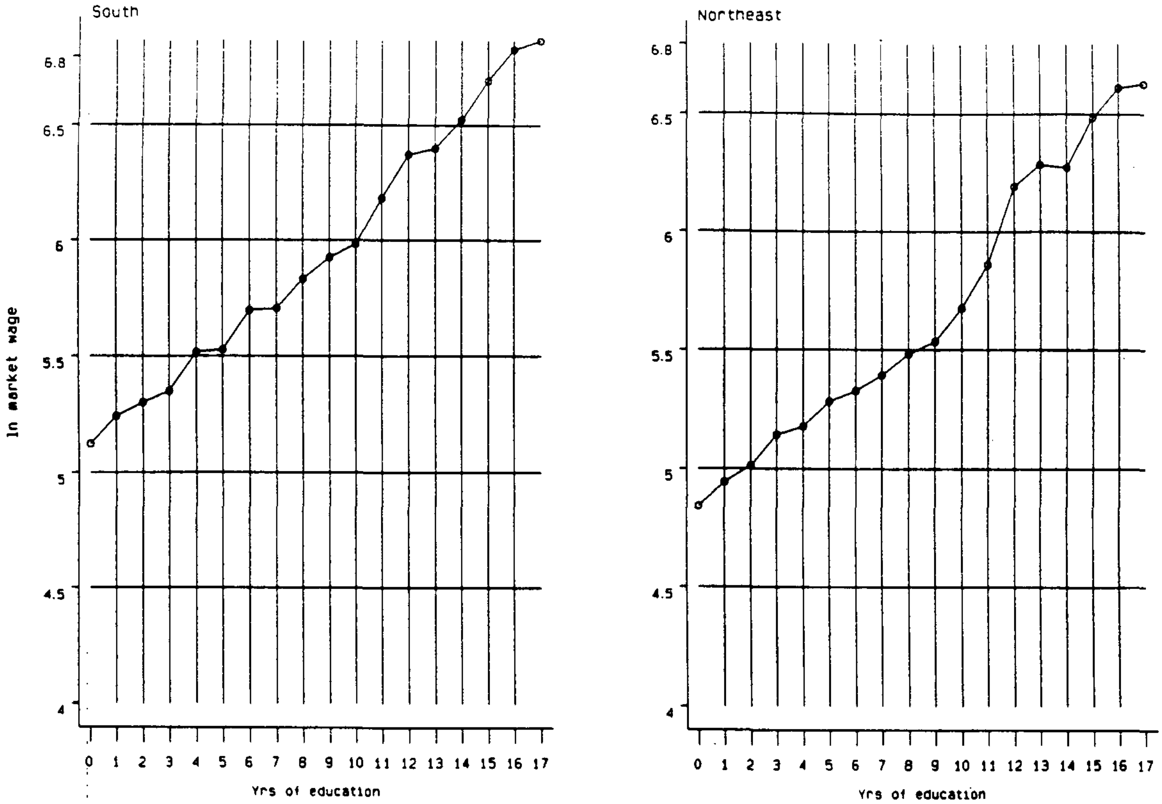
	White			Black			Mulatto		
	Men	Women	Diff	Men	Women	Diff	Men	Women	Diff
1 market wage									
) if father									
literate	-0.022	0.015	-0.037	-0.018	-0.037	0.019	-0.002	0.026	-0.028
	[0.42]	[0.20]	[0.40]	[0.14]	[0.32]	[0.14]	[0.04]	[0.50]	[0.46]
1-3 yrs education	0.002	-0.120	0.122	0.198	0.059	0.139	0.017	0.025	-0.008
	[0.02]	[1.49]	[1.27]	[2.55]	[0.43]	[0.87]	[0.52]	[0.42]	[0.11]
4-8 yrs education	0.079	0.035	0.045	0.014	0.205	-0.191	0.084	0.125	-0.041
	[1.20]	[0.40]	[0.43]	[0.12]	[1.09]	[0.85]	[1.88]	[1.78]	[0.48]
≥ 9 yrs education	0.296	0.037	0.259	0.011	-0.869	0.880	0.141	0.299	-0.158
	[3.81]	[0.35]	[1.96]	[0.00]	[1.54]	[1.43]	[1.74]	[2.51]	[1.11]
.) if mother									
literate	0.018	0.224	-0.206	0.088	0.113	-0.025	0.036	0.045	-0.009
	[0.43]	[2.82]	[2.22]	[1.24]	[0.97]	[0.18]	[0.91]	[0.84]	[0.14]
1-3 yrs education	0.007	0.246	-0.239	-0.008	0.006	-0.014	0.100	0.113	-0.013
	[0.26]	[3.02]	[2.48]	[0.09]	[0.04]	[0.08]	[2.52]	[1.93]	[0.18]
4-8 yrs education	0.057	0.326	-0.269	0.227	0.168	0.059	0.119	0.196	-0.077
	[0.92]	[3.73]	[2.54]	[1.81]	[0.78]	[0.24]	[2.46]	[2.74]	[0.89]
≥ 9 yrs education	-0.089	0.484	-0.573	0.158	1.018	-0.860	0.123	0.134	-0.011
	[1.10]	[4.16]	[3.98]	[0.35]	[1.62]	[1.12]	[1.32]	[1.05]	[0.07]
:) if spouse's father									
literate	-0.010	-0.120	0.110	0.092	-0.122	0.214	0.021	0.043	-0.022
	[0.36]	[1.31]	[1.05]	[1.42]	[0.71]	[1.16]	[0.59]	[0.63]	[0.29]
1-3 yrs education	0.176	-0.091	0.267	0.107	0.086	0.021	0.057	0.078	-0.021
	[3.32]	[0.91]	[2.38]	[1.46]	[0.31]	[0.08]	[2.02]	[0.96]	[0.24]
4-8 yrs education	0.184	-0.102	0.286	0.166	-0.108	0.274	0.089	-0.051	0.140
	[3.12]	[0.96]	[2.37]	[1.55]	[0.41]	[0.96]	[2.16]	[0.54]	[1.35]
≥ 9 yrs education	0.163	-0.103	0.266	-0.382	-	-	0.204	-0.012	0.216
	[1.79]	[0.82]	[1.78]	[1.18]	-	-	[2.78]	[0.09]	[1.32]
:) if spouse's mother									
literate	0.029	0.181	-0.152	-0.007	-0.277	0.270	0.044	0.076	-0.032
	[0.70]	[2.06]	[1.51]	[0.08]	[1.52]	[1.39]	[1.32]	[1.12]	[0.44]
1-3 yrs education	0.055	0.216	-0.161	-0.023	0.245	-0.268	0.098	0.068	0.030
	[1.16]	[2.18]	[1.44]	[0.28]	[0.86]	[0.91]	[2.72]	[0.87]	[0.35]
4-8 yrs education	0.124	0.199	-0.075	0.281	0.041	0.240	0.113	0.274	-0.161
	[2.10]	[1.91]	[0.63]	[2.39]	[0.12]	[0.66]	[2.55]	[2.80]	[1.50]
≥ 9 yrs education	0.125	0.167	-0.042	0.390	-	-	0.119	0.147	-0.028
	[1.53]	[1.29]	[0.27]	[1.13]	-	-	[1.36]	[0.88]	[0.15]

Table 3
Effects of parents' and spouse's parents' education
on log(wages) of men and women
Urban Northeast : by age group

	25-34			35-44			45-60		
	Men	**Women**	**Diff**	**Men**	**Women**	**Diff**	**Men**	**Women**	**Diff**
ln market wage									
1) if father									
literate	-0.003	0.039	-0.042	0.003	-0.078	0.080	0.043	0.127	-0.084
	[0.08]	[0.63]	[0.60]	[0.06]	[1.16]	[1.01]	[0.79]	[1.43]	[0.81]
1-3 yrs education	0.065	-0.094	0.160	-0.010	-0.044	0.034	0.087	0.080	0.007
	[1.70]	[1.51]	[2.16]	[0.21]	[0.57]	[0.37]	[1.33]	[0.76]	[0.06]
4-8 yrs education	0.108	-0.01	0.117	0.046	0.053	-0.007	0.179	0.403	-0.224
	[2.26]	[0.13]	[1.34]	[0.73]	[0.58]	[0.06]	[2.26]	[3.47]	[1.58]
≥ 9 yrs education	0.201	0.055	0.146	0.244	0.055	0.189	0.241	0.497	-0.256
	[2.97]	[0.58]	[1.24]	[2.59]	[0.39]	[1.10]	[2.07]	[2.50]	[1.10]
(1) if mother									
literate	0.047	0.012	0.035	-0.011	0.152	-0.163	0.014	0.165	-0.152
	[1.38]	[0.20]	[0.49]	[0.26]	[2.21]	[2.00]	[0.25]	[1.82]	[1.41]
1-3 yrs education	0.060	0.100	-0.039	0.002	0.203	-0.200	0.130	0.236	-0.106
	[1.60]	[1.65]	[0.54]	[0.04]	[2.65]	[2.19]	[1.91]	[2.09]	[0.80]
4-8 yrs education	0.082	0.243	-0.161	0.108	0.196	-0.089	0.139	0.175	-0.036
	[1.77]	[3.38]	[1.87]	[1.70]	[2.02]	[0.76]	[1.66]	[1.48]	[0.25]
≥ 9 yrs education	0.107	0.286	-0.179	0.103	0.487	-0.384	-0.058	0.070	-0.128
	[1.38]	[2.87]	[1.41]	[1.00]	[3.08]	[2.03]	[0.42]	[0.30]	[0.47]
1) if spouse's father									
literate	0.056	0.031	0.026	0.052	-0.069	0.121	-0.006	-0.018	0.013
	[1.70]	[0.42]	[0.32]	[1.21]	[0.80]	[1.26]	[0.12]	[0.13]	[0.09]
1-3 yrs education	0.056	0.018	0.037	0.104	-0.018	0.122	0.130	-0.063	0.193
	[1.52]	[0.23]	[0.42]	[2.30]	[0.17]	[1.06]	[2.31]	[0.37]	[1.08]
4-8 yrs education	0.085	0.039	0.046	0.162	-0.099	0.260	0.089	-0.276	0.364
	[1.90]	[0.44]	[0.46]	[2.67]	[0.88]	[2.05]	[1.17]	[1.33]	[1.64]
≥ 9 yrs education	0.247	0.040	0.207	0.158	-0.168	0.326	0.180	-0.181	0.362
	[3.50]	[0.36]	[1.54]	[1.73]	[1.05]	[1.76]	[1.56]	[0.65]	[1.20]
(1) if spouse's mother									
literate	0.002	0.144	-0.142	-0.002	0.163	-0.165	0.097	-0.219	0.316
	[0.06]	[2.02]	[1.79]	[0.05]	[1.97]	[1.77]	[1.92]	[1.56]	[2.11]
1-3 yrs education	0.104	0.156	-0.052	0.086	0.203	-0.117	0.066	-0.016	0.082
	[2.80]	[1.96]	[0.58]	[1.86]	[2.04]	[1.07]	[1.07]	[0.10]	[0.45]
4-8 yrs education	0.100	0.234	-0.135	0.157	0.261	-0.105	0.075	0.167	-0.093
	[2.32]	[2.60]	[1.34]	[2.58]	[2.30]	[0.81]	[0.99]	[0.81]	[0.42]
≥ 9 yrs education	0.134	0.187	-0.052	0.150	0.278	-0.128	0.191	-0.076	0.267
	[1.90]	[1.56]	[0.37]	[1.53]	[1.48]	[0.60]	[1.33]	[0.26]	[0.80]

Appendix Figure 1
Returns to education : ℓn wages

Males



Appendix Table 1
Education and Wages of Urban Brazilians

	South		North		Northeast	
	Males	**Females**	**Males**	**Females**	**Males**	**Females**
#25-60 year old household heads and spouses	37,299	39,035	9,149	9,373	12,239	13,679
% with:						
0 years education	11.5	16.1	17.6	21.9	26.7	29.2
1-3 yrs education	19.7	20.7	25.4	25.4	21.8	21.1
4-8 yrs education	45.3	43.2	37.8	36.0	33.4	33.1
≥ 9 yrs education	11.5	11.7	11.5	11.9	10.9	11.6
12+ yrs education	12.0	8.3	7.7	4.6	7.3	5.1
% participating on job	90.3	38.0	94.1	36.7	90.9	38.5
for market wage	65.9	26.4	55.4	21.8	59.1	21.1
for self-employment	24.5	11.6	38.7	12.7	31.8	17.4
Mean log total wage of participants	5.69	5.23	5.45	4.93	5.27	4.69
(standard deviation)	(0.94)	(0.99)	(0.90)	(0.95)	(0.96)	(1.09)
Mean log market wage of participants	5.69	5.34	5.37	5.02	5.30	4.96
(standard deviation)	(0.94)	(0.98)	(0.89)	(0.92)	(0.97)	(1.08)
Log self-employment wage of participants	5.70	4.96	5.54	4.78	5.19	4.37
(standard deviation)	(0.93)	(0.97)	(0.92)	(0.97)	(0.95)	(1.00)

Appendix Table 2

Urban Brazil : Wage functions : Total, market and self employment wage functions

Own education

	ln(total wages)			ln(market wages)			ln(self-emp wages		
	South	North	Northeast	South	North	Northeast	South	North	Northeas
[redacted] 1 yr education	0.120	0.121	0.173	0.094	0.062	0.143	0.149	0.168	0.169
	[0.034]	[0.063]	[0.051]	[0.040]	[0.081]	[0.074]	[0.064]	[0.097]	[0.075
[redacted] 2 yrs education	0.176	0.132	0.185	0.150	0.098	0.274	0.216	0.154	0.110
	[0.026]	[0.049]	[0.041]	[0.031]	[0.059]	[0.061]	[0.048]	[0.082]	[0.060
(1) 3 yrs education	0.196	0.285	0.257	0.203	0.145	0.287	0.186	0.420	0.216
	[0.023]	[0.047]	[0.038]	[0.027]	[0.058]	[0.053]	[0.043]	[0.075]	[0.058
[redacted] 4 yrs education	0.390	0.331	0.376	0.364	0.260	0.457	0.412	0.428	0.300
	[0.020]	[0.044]	[0.034]	[0.023]	[0.052]	[0.047]	[0.039]	[0.077]	[0.052
(1) 5 yrs education	0.439	0.509	0.549	0.434	0.408	0.591	0.434	0.599	0.515
	[0.026]	[0.051]	[0.039]	[0.031]	[0.060]	[0.053]	[0.049]	[0.088]	[0.061
(1) 6 yrs education	0.517	0.639	0.614	0.526	0.505	0.733	0.462	0.750	0.447
	[0.039]	[0.079]	[0.077]	[0.046]	[0.091]	[0.095]	[0.075]	[0.141]	[0.132
[redacted] 7 yrs education	0.640	0.408	0.779	0.630	0.448	0.836	0.606	0.314	0.663
	[0.041]	[0.083]	[0.082]	[0.049]	[0.089]	[0.098]	[0.076]	[0.171]	[0.144
(1) 8 yrs education	0.793	0.830	0.805	0.862	0.759	0.936	0.617	0.982	0.617
	[0.027]	[0.062]	[0.049]	[0.031]	[0.069]	[0.061]	[0.056]	[0.120]	[0.086
[redacted] 9 yrs education	0.909	0.592	0.839	0.995	0.516	0.832	0.683	0.825	1.182
	[0.058]	[0.107]	[0.123]	[0.064]	[0.112]	[0.129]	[0.128]	[0.235]	[0.299
[redacted] 10 yrs education	0.966	0.917	0.901	1.073	0.951	1.061	0.643	0.854	0.679
	[0.048]	[0.087]	[0.085]	[0.052]	[0.090]	[0.099]	[0.116]	[0.219]	[0.159
(1) 11 yrs education	1.256	1.035	1.187	1.320	1.089	1.367	0.963	1.034	0.938
	[0.026]	[0.052]	[0.041]	[0.029]	[0.061]	[0.054]	[0.065]	[0.147]	[0.101
[redacted] 12 yrs education	1.387	1.103	1.404	1.470	1.173	1.574	0.861	0.871	1.177
	[0.057]	[0.150]	[0.119]	[0.057]	[0.141]	[0.121]	[0.220]	[0.771]	[0.399
[redacted] 13 yrs education	1.455	1.296	1.508	1.569	1.263	1.700	0.845	2.309	1.263
	[0.050]	[0.144]	[0.117]	[0.051]	[0.138]	[0.120]	[0.156]	[0.538]	[0.396
(1) 14 yrs education	1.588	1.435	1.744	1.649	1.520	1.931	1.115	1.441	1.798
	[0.039]	[0.092]	[0.091]	[0.041]	[0.096]	[0.096]	[0.170]	[0.376]	[0.455
[redacted] 15 yrs education	1.744	1.772	1.914	1.800	1.820	2.138	1.423	1.680	1.529
	[0.030]	[0.075]	[0.058]	[0.034]	[0.082]	[0.074]	[0.097]	[0.232]	[0.184
[redacted] 16 yrs education	1.892	1.936	2.200	1.957	1.966	2.491	1.616	2.022	1.630
	[0.045]	[0.132]	[0.091]	[0.049]	[0.134]	[0.102]	[0.108]	[0.354]	[0.275
(1) 17 yrs education	2.050	2.112	2.222	2.098	2.054	2.351	1.781	2.466	2.709
	[0.079]	[0.179]	[0.175]	[0.085]	[0.183]	[0.172]	[0.188]	[0.417]	[0.777
[redacted] 1 yr education * male	0.000	-0.002	-0.071	0.007	0.022	0.002	-0.018	-0.006	-0.132
	[0.040]	[0.072]	[0.060]	[0.047]	[0.092]	[0.083]	[0.081]	[0.115]	[0.091
(1) 2 yrs education * male	0.001	-0.018	-0.014	0.021	0.055	-0.071	-0.051	-0.074	0.024
	[0.031]	[0.058]	[0.049]	[0.036]	[0.069]	[0.068]	[0.062]	[0.096]	[0.077
[redacted] 3 yrs education * male	0.031	-0.025	0.044	0.016	0.086	0.068	0.035	-0.118	0.015
	[0.028]	[0.055]	[0.046]	[0.032]	[0.068]	[0.061]	[0.056]	[0.089]	[0.075
[redacted] 4 yrs education * male	0.006	0.035	-0.041	0.025	0.072	-0.041	-0.021	0.026	-0.071
	[0.024]	[0.052]	[0.041]	[0.028]	[0.062]	[0.054]	[0.050]	[0.089]	[0.071
(1) 5 yrs education * male	-0.034	-0.159	-0.108	-0.020	0.105	-0.068	-0.079	-0.252	-0.155
	[0.032]	[0.060]	[0.048]	[0.037]	[0.073]	[0.062]	[0.063]	[0.107]	[0.083
[redacted] 6 yrs education * male	0.058	-0.193	-0.128	0.056	0.136	-0.103	0.071	-0.372	-0.180
	[0.046]	[0.092]	[0.091]	[0.053]	[0.107]	[0.109]	[0.092]	[0.167]	[0.165
[redacted] 7 yrs education * male	-0.056	0.101	-0.227	-0.025	0.158	-0.164	-0.075	0.251	-0.239
	[0.048]	[0.096]	[0.095]	[0.056]	[0.106]	[0.112]	[0.094]	[0.191]	[0.178
(1) 8 yrs education * male	-0.083	-0.198	-0.166	-0.111	0.017	-0.153	0.015	-0.193	-0.124
	[0.032]	[0.072]	[0.059]	[0.036]	[0.084]	[0.073]	[0.070]	[0.141]	[0.116
[redacted] 9 yrs education * male	-0.106	-0.047	-0.145	-0.148	0.396	-0.051	0.069	-0.302	-0.609
	[0.068]	[0.126]	[0.139]	[0.074]	[0.138]	[0.148]	[0.157]	[0.287]	[0.330
[redacted] 10 yrs education * male	-0.106	-0.243	-0.065	-0.172	-0.041	-0.112	0.159	-0.056	0.092
	[0.057]	[0.105]	[0.102]	[0.061]	[0.114]	[0.117]	[0.136]	[0.255]	[0208]
(1) 11 yrs education * male	-0.199	-0.027	-0.169	-0.175	0.261	-0.148	-0.098	-0.067	-0.222
	[0.032]	[0.064]	[0.052]	[0.036]	[0.083]	[0.069]	[0.079]	[0.169]	[0.132
[redacted] 12 yrs education * male	-0.140	-0.141	-0.055	-0.133	0.181	0.067	0.245	0.004	-0.385
	[0.070]	[0.180]	[0.157]	[0.071]	[0.179]	[0.160]	[0.251]	[0.811]	[0.515
[redacted] 13 yrs education * male	-0.181	-0.101	-0.066	-0.203	0.271	0.068	0.265	-1.295	-0.522
	[0.062]	[0.172]	[0.142]	[0.064]	[0.174]	[0.149]	[0.184]	[0.573]	[0.450
(1) 14 yrs education * male	-0.187	-0.288	-0.314	-0.102	0.144	-0.168	-0.113	-0.591	-1.006
	[0.051]	[0.119]	[0.122]	[0.055]	[0.132]	[0.132]	[0.192]	[0.411]	[0.492
[redacted] 15 yrs education * male	-0.171	-0.322	-0.269	-0.096	0.056	-0.193	-0.117	-0.349	-0.280
	[0.038]	[0.091]	[0.075]	[0.043]	[0.111]	[0.097]	[0.113]	[0.258]	[0.232
(1) 16 yrs education * male	-0.182	-0.252	-0.429	-0.109	0.110	-0.376	-0.193	-0.235	-0.478
	[0.052]	[0.149]	[0.106]	[0.057]	[0.160]	[0.121]	[0.122]	[0.381]	[0.307
(1) 17 yrs education * male	-0.302	-0.401	-0.436	-0.389	-0.242	-0.287	-0.049	-0.733	-1.346
	[0.090]	[0.204]	[0.200]	[0.098]	[0.218]	[0.202]	[0.209]	[0.448]	[0.817
i[redacted]ercept	3.340	3.317	3.135	3.304	3.543	3.417	3.562	3.193	2.930
	[0.103]	[0.232]	[0.189]	[0.117]	[0.260]	[0.232]	[0.221]	[0.447]	[0.332
(1) male	-0.334	-0.206	-0.102	-0.195	0.820	-0.056	-0.248	0.446	0.382
	[0.123]	[0.271]	[0.225]	[0.147]	[0.371]	[0.284]	[0.277]	[0.529]	[0.422

Appendix Table 2 (continued)

Urban Brazil : Total, market and self employment wage functions

Age, race and parents' education

	ln(total wages)			ln(market wages)			ln(self-emp wages		
	South	North	Northeast	South	North	Northeast	South	North	Northeas
Sample selection term									
r(participate) females	-0.038	-0.037	-0.074	-0.036	-0.088	-0.198	-0.215	-0.155	-0.07
	[0.017]	[0.036]	[0.029]	[0.029]	[0.069]	[0.074]	[0.099]	[0.159]	[0.095
r(participate) males	.	.	.	-0.169	-1.114	-0.399	0.326	0.364	-0.285
				[0.055]	[0.192]	[0.111]	[0.157]	[0.238]	[0.24
ge									
ge	0.046	0.043	0.044	0.046	0.031	0.029	0.039	0.052	0.0
	[0.005]	[0.012]	[0.009]	[0.006]	[0.013]	[0.012]	[0.011]	[0.022]	[0.01
ge**2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.001	-0.001
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000
ge * male	0.034	0.024	0.028	0.036	0.011	0.042	0.025	-0.010	0.0
	[0.006]	[0.013]	[0.011]	[0.007]	[0.016]	[0.013]	[0.013]	[0.027]	[0.0
ge**2 * male	0.000	0.000	0.000	0.000	0.000	-0.001	0.000	0.000	0.0
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000
Race									
(1) black	-0.034	-0.153	0.014	-0.004	-0.149	0.038	-0.109	-0.115	-0.0
	[0.021]	[0.054]	[0.032]	[0.024]	[0.060]	[0.042]	[0.044]	[0.101]	[0.0
(1) white	0.091	0.091	0.105	0.106	0.062	0.056	0.052	0.114	0.1
	[0.014]	[0.026]	[0.021]	[0.016]	[0.029]	[0.026]	[0.029]	[0.050]	[0.037
(1) black c' * male	-0.013	0.115	-0.014	-0.044	0.173	0.000	0.080	0.124	-0.042
	[0.026]	[0.065]	[0.039]	[0.029]	[0.072]	[0.049]	[0.060]	[0.122]	[0.07
(1) white * male	0.002	0.002	-0.027	-0.043	-0.069	-0.023	0.088	-0.046	-0.0
	[0.017]	[0.031]	[0.026]	[0.019]	[0.036]	[0.031]	[0.037]	[0.059]	[0.0
Father's education									
(1) father literate	0.029	-0.068	0.030	0.026	-0.022	0.019	0.038	-0.129	0.0
	[0.020]	[0.037]	[0.030]	[0.023]	[0.042]	[0.040]	[0.038]	[0.064]	[0.0
(1) father educ 1-3 yrs	0.019	-0.045	-0.015	0.047	0.009	-0.022	-0.029	-0.113	-0.0
	[0.019]	[0.041]	[0.034]	[0.021]	[0.047]	[0.044]	[0.037]	[0.074]	[0.055
(1) father educ 4-8 ys	0.089	0.041	0.111	0.127	0.066	0.109	0.015	0.036	0.066
	[0.020]	[0.047]	[0.042]	[0.023]	[0.053]	[0.050]	[0.040]	[0.085]	[0.078
(1) father educ 9+ yrs	0.184	0.138	0.152	0.198	0.222	0.154	0.259	-0.126	0.1
	[0.030]	[0.072]	[0.065]	[0.032]	[0.077]	[0.072]	[0.069]	[0.156]	[0.1
(1) father literate * male	0.000	0.136	-0.026	0.007	0.074	-0.019	-0.041	0.197	-0.0
	[0.023]	[0.043]	[0.037]	[0.027]	[0.051]	[0.047]	[0.048]	[0.075]	[0.061
(1) father educ 1-3 yrs * male	0.017	0.150	0.064	-0.011	0.029	0.065	0.047	0.214	0.044
	[0.022]	[0.049]	[0.042]	[0.025]	[0.057]	[0.052]	[0.046]	[0.087]	[0.07
(1) father educ 4-8 ys * male	-0.016	0.080	0.009	-0.047	0.023	-0.014	0.031	0.047	0.08
	[0.024]	[0.057]	[0.053]	[0.027]	[0.065]	[0.061]	[0.051]	[0.103]	[0.1
(1) father educ 9+ yrs * male	-0.060	0.036	0.107	-0.072	-0.201	0.075	-0.140	0.344	0.2
	[0.037]	[0.089]	[0.081]	[0.040]	[0.097]	[0.088]	[0.084]	[0.185]	[0.176
Mother's education									
(1) mother literate	0.015	0.070	0.066	0.025	0.071	0.105	-0.016	0.100	0.0
	[0.020]	[0.036]	[0.031]	[0.023]	[0.041]	[0.041]	[0.039]	[0.066]	[0.0
(1) mother educ 1-3 yrs	0.049	0.092	0.147	0.019	0.053	0.152	0.085	0.162	0.153
	[0.018]	[0.040]	[0.035]	[0.021]	[0.045]	[0.044]	[0.036]	[0.073]	[0.059
(1) mother educ 4-8 yrs	0.119	0.171	0.187	0.070	0.150	0.222	0.204	0.180	0.16
	[0.020]	[0.047]	[0.043]	[0.022]	[0.053]	[0.051]	[0.041]	[0.089]	[0.0
(1) mother educ 9+ yrs	0.173	0.310	0.245	0.125	0.286	0.307	0.216	0.383	0.0
	[0.032]	[0.081]	[0.072]	[0.034]	[0.081]	[0.079]	[0.078]	[0.212]	[0.1
(1) mother literate * male	0.048	0.020	-0.013	0.017	-0.030	-0.091	0.123	0.043	0.071
	[0.023]	[0.043]	[0.038]	[0.026]	[0.048]	[0.048]	[0.048]	[0.076]	[0.064
(1) mother educ 1-3 yrs * male	0.020	0.019	-0.080	0.034	-0.038	-0.092	-0.012	-0.011	-0.0
	[0.022]	[0.048]	[0.043]	[0.025]	[0.055]	[0.052]	[0.045]	[0.086]	[0.0
(1) mother educ 4-8 yrs * male	0.003	-0.019	-0.100	0.043	0.003	-0.118	-0.095	-0.029	-0.0
	[0.024]	[0.057]	[0.054]	[0.027]	[0.064]	[0.062]	[0.051]	[0.107]	[0.10
(1) mother educ 9+ yrs * male	0.067	-0.077	-0.145	0.092	0.123	-0.260	0.046	-0.181	0.120
	[0.039]	[0.099]	[0.089]	[0.042]	[0.103]	[0.097]	[0.095]	[0.243]	[0.199
(1) father educ missing	0.032	-0.013	-0.025	0.060	0.042	-0.101	-0.019	-0.086	0.0
	[0.026]	[0.052]	[0.046]	[0.030]	[0.060]	[0.059]	[0.054]	[0.092]	[0.0
(1) mother educ missing	0.016	0.075	0.069	-0.017	0.043	0.113	0.088	0.142	0.095
	[0.031]	[0.058]	[0.052]	[0.034]	[0.064]	[0.065]	[0.064]	[0.108]	[0.086
(1) father educ missing * male	-0.015	0.091	-0.002	-0.035	0.057	0.089	0.026	0.201	-0.073
	[0.031]	[0.061]	[0.054]	[0.034]	[0.070]	[0.067]	[0.065]	[0.108]	[0.0
(1) mother educ missing * male	0.009	-0.025	0.006	0.031	0.053	-0.024	-0.025	-0.112	-0.0
	[0.035]	[0.067]	[0.060]	[0.039]	[0.074]	[0.073]	[0.075]	[0.124]	[0.1
F [illegible]	11,18[.00]	3.03 [.02]	3,00[.02]	11,59[.00]	2,69[.03]	2.36[.05]	4.95[.00]	1.93[.10]	0,62[.65
[illegible]	10,11[.00]	5.66 [.00]	8.21[.00]	9,22[.00]	1.96[.10]	5.45[.00]	1,72[.14]	1.96[.10]	3.78
[illegible]	1,23[.30]	3.52 [.01]	1,45[.22]	1,34[.25]	2.56[.04]	1,02[.40]	2,05[.08]	3.[illegible][.01]	0.59[.47]
F [illegible]	11,47[.00]	5.00 [.00]	7.31[.00]	4.06[.00]	3.73[.00]	6.26[.00]	7,37[.00]	2,14[.07]	2.2
[illegible]	33.53[.00]	9.21[.00]	3.41[.00]	22.01[.00]	12.12[.00]	2,58[.04]	8.26[.00]	4,77[.00]	1.94
[illegible]	1,96[.10]	0.35	1,43[.22]	1,33[.26]	0.74[.55]	2,35[.05]	3.70[.01]	0.27[.90]	1,19[.31]

Appendix Table 2 (continued)
Urban Brazil : Total, market and self employment wage functions
Spouse and spouse's parents' education

	ln(total wages)			ln(market wages)			ln(self-emp wages		
	South	North	Northeast	South	North	Northeast	South	North	Northeas
Spouse's education									
(1) spouse exist	-0.218	-0.170	-0.226	-0.172	-0.079	-0.168	-0.273	-0.260	-0.292
	[0.023]	[0.045]	[0.030]	[0.027]	[0.054]	[0.042]	[0.044]	[0.076]	[0.047
years education of spouse	0.040	-0.011	0.056	0.018	-0.038	0.056	0.079	0.006	0.037
	[0.019]	[0.038]	[0.032]	[0.023]	[0.045]	[0.046]	[0.035]	[0.064]	[0.048
(1) spouse education ge 1	-0.021	0.074	-0.053	0.028	0.028	-0.051	-0.104	0.160	0.000
	[0.050]	[0.094]	[0.077]	[0.061]	[0.113]	[0.112]	[0.091]	[0.157]	[0.113
(1) spouse education ge 4	-0.029	0.119	0.040	-0.012	0.196	0.027	-0.091	0.052	0.061
	[0.037]	[0.079]	[0.072]	[0.044]	[0.093]	[0.099]	[0.067]	[0.137]	[0.110
(1) spouse education ge 8	0.062	-0.031	-0.109	0.118	-0.054	0.042	-0.034	-0.116	-0.410
	[0.045]	[0.098]	[0.097]	[0.050]	[0.107]	[0.110]	[0.096]	[0.188]	[0.184
(1) spouse education ge 11	0.020	0.348	-0.117	-0.016	0.313	-0.011	0.236	0.358	-0.253
	[0.066]	[0.135]	[0.143]	[0.068]	[0.137]	[0.156]	[0.184]	[0.316]	[0.308
(1) spouse yrs ed 1-4	-0.022	0.039	-0.018	-0.016	0.050	-0.077	-0.031	0.065	0.105
	[0.022]	[0.045]	[0.041]	[0.026]	[0.052]	[0.055]	[0.042]	[0.080]	[0.068
(1) spouse yrs ed 5-8	-0.012	-0.083	0.031	0.007	-0.035	0.046	-0.084	-0.106	-0.006
	[0.026]	[0.056]	[0.058]	[0.028]	[0.059]	[0.064]	[0.068]	[0.119]	[0.118
(1) spouse yrs ed 9-11	0.012	0.040	-0.039	0.015	0.016	-0.009	0.052	-0.028	-0.033
	[0.024]	[0.053]	[0.053]	[0.025]	[0.055]	[0.058]	[0.066]	[0.118]	[0.114
(1) spouse exist * male	0.055	0.068	0.196	0.045	0.040c'	0.081	0.106	0.178	0.415
	[0.029]	[0.058]	[0.044]	[0.034]	[0.069]	[0.056]	[0.062]	[0.100]	[0.075
years educatn of spouse * male	-0.013	0.030	-0.028	0.007	0.056	-0.026	-0.054	0.008	-0.027
	[0.022]	[0.043]	[0.037]	[0.026]	[0.050]	[0.051]	[0.042]	[0.072]	[0.059
(1) spouse educatn ge 1 * male	0.062	0.002	0.077	0.013	-0.014	0.068	0.133	-0.108	0.055
	[0.057]	[0.107]	[0.089]	[0.068]	[0.128]	[0.123]	[0.111]	[0.178]	[0.140
(1) spouse educatn ge 4 * male	0.104	-0.036	-0.014	0.089	-0.129	-0.029	0.153	0.036	0.036
	[0.042]	[0.089]	[0.082]	[0.050]	[0.105]	[0.109]	[0.082]	[0.153]	[0.132
(1) spouse educatn ge 8 * male	-0.032	0.147	0.201	-0.125	0.183	0.053	0.123	0.315	0.544
	[0.051]	[0.111]	[0.109]	[0.056]	[0.122]	[0.123]	[0.112]	[0.209]	[0.213
(1) spouse educatn ge 11* male	0.018	-0.153	0.148	0.060	-0.263	0.033	-0.224	-0.075	0.317
	[0.078]	[0.156]	[0.162]	[0.081]	[0.160]	[0.175]	[0.208]	[0.354]	[0.348
(1) spouse yrs ed 1-4 * male	0.013	-0.037	0.019	0.010	-0.065	0.068	0.026	-0.075	-0.078
	[0.025]	[0.051]	[0.047]	[0.030]	[0.059]	[0.061]	[0.051]	[0.089]	[0.081
(1) spouse yrs ed 5-8 * male	0.008	0.008	-0.052	-0.011	0.033	-0.067	0.079	0.003	-0.025
	[0.031]	[0.065]	[0.065]	[0.033]	[0.068]	[0.072]	[0.078]	[0.135]	[0.135
(1) spouse yrs ed 9-11 * male	-0.017	0.024	0.043	-0.022	0.001	0.011	-0.053	0.115	0.045
	[0.029]	[0.061]	[0.061]	[0.030]	[0.064]	[0.066]	[0.075]	[0.134]	[0.131
Spouse's parents' education									
(1) spouse father literate	0.000	0.056	-0.034	-0.038	0.066	-0.036	0.068	-0.005	-0.053
	[0.024]	[0.045]	[0.038]	[0.028]	[0.053]	[0.051]	[0.045]	[0.079]	[0.059
(1) spouse father educ 1-3 yrs	0.006	0.094	0.034	-0.026	0.108	-0.013	0.047	0.048	0.105
	[0.023]	[0.051]	[0.045]	[0.026]	[0.057]	[0.060]	[0.045]	[0.095]	[0.073
(1) spouse father educ 4-8 yrs	0.038	0.033	-0.066	0.015	0.051	-0.072	0.093	-0.099	-0.119
	[0.025]	[0.060]	[0.055]	[0.028]	[0.065]	[0.066]	[0.052]	[0.117]	[0.103
(1) spouse father educ 9+ yrs	0.046	0.093	0.044	0.032	-0.005	-0.082	0.116	0.353	0.305
	[0.035]	[0.090]	[0.078]	[0.038]	[0.094]	[0.088]	[0.082]	[0.204]	[0.162
(1) spouse father lit * male	0.023	-0.092	0.070	0.051	-0.067	0.070	-0.039	-0.034	0.073
	[0.027]	[0.051]	[0.043]	[0.031]	[0.060]	[0.056]	[0.054]	[0.089]	[0.070
(1) spouse fthr ed 1-3yr* male	0.037	-0.120	0.032	0.061	-0.105	0.113	-0.008	-0.085	-0.078
	[0.026]	[0.057]	[0.051]	[0.029]	[0.064]	[0.065]	[0.053]	[0.105]	[0.086
(1) spouse fthr ed 4-8yr* male	0.027	-0.012	0.169	0.055	0.003	0.210	-0.058	0.131	0.174
	[0.028]	[0.067]	[0.062]	[0.032]	[0.074]	[0.073]	[0.060]	[0.129]	[0.119
(1) spouse fthr ed 9+ yr* male	0.100	0.082	0.179	0.127	0.169	0.291	-0.005	-0.234	-0.115
	[0.042]	[0.104]	[0.091]	[0.045]	[0.110]	[0.102]	[0.095]	[0.224]	[0.194
(1) spouse mother literate	0.053	-0.004	0.068	0.066	-0.032	0.090	0.028	0.037	0.038
	[0.023]	[0.044]	[0.039]	[0.027]	[0.051]	[0.050]	[0.045]	[0.076]	[0.063
(1) spouse mother educ 1-3 yrs	0.056	0.017	0.094	0.070	-0.049	0.138	0.023	0.120	0.037
	[0.022]	[0.050]	[0.045]	[0.025]	[0.056]	[0.059]	[0.044]	[0.094]	[0.074
(1) spouse mother educ 4-8 yrs	0.068	0.017	0.196	0.037	0.027	0.224	0.127	-0.052	0.203
	[0.024]	[0.059]	[0.056]	[0.027]	[0.064]	[0.067]	[0.051]	[0.115]	[0.103
(1) spouse mother educ 9+ yrs	0.033	-0.029	0.132	0.007	0.013	0.181	0.077	-0.269	0.028
	[0.038]	[0.097]	[0.085]	[0.041]	[0.100]	[0.095]	[0.092]	[0.218]	[0.185
(1) spouse mthr lit * male	-0.027	0.075	-0.029	-0.047	0.108	-0.057	0.004	0.019	0.007
	[0.027]	[0.049]	[0.044]	[0.030]	[0.057]	[0.056]	[0.054]	[0.086]	[0.075
(1) spouse mthr ed 1-3yr* male	-0.020	0.077	-0.013	-0.040	0.113	-0.059	0.022	-0.014	0.025
	[0.025]	[0.056]	[0.051]	[0.028]	[0.063]	[0.064]	[0.051]	[0.104]	[0.088
(1) spouse mthr ed 4-8yr* male	0.019	0.110	-0.047	0.038	0.072	-0.111	-0.002	0.282	0.024
	[0.027]	[0.067]	[0.063]	[0.031]	[0.073]	[0.074]	[0.059]	[0.129]	[0.120
(1) spouse mthr ed 9+ yr* male	0.090	0.112	0.085	0.090	-0.064	-0.051	0.118	0.531	0.444
	[0.045]	[0.114]	[0.100]	[0.049]	[0.120]	[0.109]	[0.106]	[0.244]	[0220]
(1) spouse father educ missing	0.025	0.032	-0.102	-0.005	0.053	-0.190	0.085	-0.049	-0.020
	[0.028]	[0.057]	[0.049]	[0.032]	[0.067]	[0.070]	[0.055]	[0.099]	[0.074
(1) spouse mother educ missing	0.055	0.053	0.179	0.037	-0.011	0.299	0.078	0.116	0.070
	[0.031]	[0.060]	[0.053]	[0.036]	[0.070]	[0.073]	[0.062]	[0.104]	[0.082
(1) spouse fthr ed miss * male	-0.017	-0.024	0.082	0.017	0.050	0.208	-0.088	0.058	-0.094
	[0.034]	[0.069]	[0.061]	[0.039]	[0.080]	[0.080]	[0.072]	[0.122]	[0.102
(1) spouse mthr ed miss * male	-0.016	-0.021	-0.156	-0.005	-0.060	-0.307	-0.033	-0.029	0.014
	[0.039]	[0.074]	[0.066]	[0.044]	[0.086]	[0.085]	[0.081]	[0.131]	[0.112
F sp fthr ed female	0.82[.51]	1.05[.38]	1.44[.20]	1.37[.24]	1.15[.33]	0.43[.79]	1.07[.37]	1.51[.19]	2.83[.02]
" males	12.19[.00]	4.93[.00]	6.41[.00]	12.39[.00]	2.46[.04]	7.02[.00]	1.46[.21]	1.12[.34]	0.82[.51]
" diff	1.73[.14]	2.24[.06]	2.34[.05]	2.36[.05]	2.18[.07]	2.60[.03]	0.40[.81]	1.27[.28]	1.65[.16]
F sp mthr ed female	3.09[.01]	0.11[.98]	3.36[.01]	2.90[.02]	0.51[.73]	3.07[.02]	1.64[.16]	1.13[.34]	1.10[.35]
" male	12.66[.00]	5.80[.00]	7.90[.00]	7.50[.00]	4.31[.00]	3.79[.00]	5.38[.00]	14.36[.00]	5.43[.00]

Appendix Table 2 (continued)

Urban Brazil : Total, market and self employment wage functions

State dummies : South and Northeast

		ln(total wages)			ln(market wages)			ln(self-emp wages		
		South	North	Northeast	South	North	Northeast	South	North	Northeas
)uth										
) Sao Paulo		0.205	.	.	0.164	.	.	0.292	.	.
		[0.017]			[0.019]			[0.035]		
) Sao Paulo	* male	0.017	.	.	0.073	.	.	-0.104	.	.
		[0.020]			[0.023]			[0.044]		
) Parana		0.064	.	.	0.017	.	.	0.179	.	.
		[0.022]			[0.024]			[0.046]		
) Parana	* male	-0.031	.	.	-0.031	.	.	-0.058	.	.
		[0.026]			[0.029]			[0.056]		
) Catarina		0.085	.	.	0.087	.	.	0.105	.	.
		[0.038]			[0.042]			[0.084]		
) Catarina	* male	-0.129	.	.	-0.148	.	.	-0.146	.	.
		[0.045]			[0.049]			[0.098]		
) Rio G do Sul		0.082	.	.	0.049	.	.	0.182	.	.
		[0.018]			[0.021]			[0.040]		
) Rio G do Sul	* male	-0.049	.	.	-0.014	.	.	-0.157	.	.
		[0.022]			[0.025]			[0.049]		
) Minas Gerais		-0.087	.	.	-0.100	.	.	-0.055	.	.
		[0.018]			[0.021]			[0.034]		
) Minas Gerais	* male	0.053	.	.	0.076	.	.	-0.044	.	.
		[0.021]			[0.025]			[0.044]		
.) Espirito Santo		-0.069	.	.	-0.069	.	.	-0.056	.	.
		[0.043]			[0.050]			[0.086]		
) Espirito Santo	* male	0.125	.	.	0.079	.	.	0.191	.	.
		[0.051]			[0.058]			[0.105]		
.) Brasilia		0.272	.	.	0.296	.	.	0.170	.	.
		[0.020] c'			[0.022]			[0.043]		
.) Brasilia	* male	-0.092	.	.	-0.124	.	.	0.020	.	.
		[0.024]			[0.026]			[0.055]		
ortheast										
.) Maranhao		.	.	-0.331	.	.	-0.615	.	.	-0.123
				[0.047]			[0.064]			[0.072
.) Maranhao	* male	.	.	-0.078	.	.	0.147	.	.	-0.236
				[0.057]			[0.078]			[0.097
.) Piaui		.	.	-0.200	.	.	-0.339	.	.	-0.081
				[0.054]			[0.071]			[0.085
.) Piaui	* male	.	.	-0.182	.	.	-0.067	.	.	-0.305
				[0.066]			[0.087]			[0.111
.) Ceara		.	.	-0.137	.	.	-0.122	.	.	-0.129
				[0.026]			[0.033]			[0.044
.) Ceara	* male	.	.	-0.053	.	.	-0.079	.	.	-0.029
				[0.032]			[0.040]			[0.057
.) Rio G Norte		.	.	-0.167	.	.	-0.181	.	.	-0.183
				[0.050]			[0.059]			[0.089
.) Rio G Norte	* male	.	.	-0.028	.	.	-0.092	.	.	0.094
				[0.060]			[0.070]			[0.109
.) Paraiba		.	.	-0.252	.	.	-0.297	.	.	-0.153
				[0.045]			[0.052]			[0.084
.) Paraiba	* male	.	.	0.005	.	.	0.019	.	.	-0.043
				[0.054]			[0.063]			[0.102
.) Pernambuco		.	.	-0.094	.	.	-0.232	.	.	0.090
				[0.026]			[0.031]			[0.046
.) Pernambuco	* male	.	.	-0.008	.	.	0.114	.	.	-0.130
				[0.031]			[0.037]			[0.060
.) Alagoas		.	.	0.207	.	.	0.099	.	.	0.413
				[0.058]			[0.068]			[0.108
.) Alagoas	* male	.	.	-0.178	.	.	-0.163	.	.	-0.269
				[0.070]			[0.080]			[0.129
.) Sergipe		.	.	0.108	.	.	-0.029	.	.	0.356
				[0.080]			[0.092]			[0.151
.) Sergipe	* male	.	.	-0.053	.	.	-0.096	.	.	-0.108
				[0.095]			[0.112]			[0.176

Appendix Table 2 (continued)

Urban Brazil : Total, market and self employment wage functions
<u>State dummies : North</u>

	ln(total wages)			ln(market wages)			ln(self-emp wages		
	South	North	Northeast	South	North	Northeast	South	North	Northeas
North									
(1) Rondonia	.	0.367	.	.	0.460	.	.	0.248	.
		[0.062]			[0.064]			[0.138]	
[redacted] Rondonia * male	.	-0.096	.	.	-0.115	.	.	-0.032	.
		[0.072]			[0.077]			[0.153]	
(1) Acre	.	0.245	.	.	0.226	.	.	0.273	.
		[0.080]			[0.085]			[0.165]	
(1) Acre * male	.	-0.209	.	.	0.044	.	.	-0.469	.
		[0.096]			[0.103]			[0.193]	
[redacted] Amazonas	.	0.383	.	.	0.398	.	.	0.396	.
		[0.039]			[0.043]			[0.077]	
(1) Amazonas * male	.	-0.13	.	.	-0.104	.	.	-0.180	.
		[0.047]			[0.052]			[0.089]	
(1) Roraima	.	0.499	.	.	0.433	.	.	1.045	.
		[0.129]			[0.125]			[0.343]	
[redacted] Roraima * male	.	-0.052	.	.	-0.031	.	.	-0.640	.
		[0.159]			[0.167]			[0.378]	
(1) Para	.	0.004	.	.	-0.052	.	.	0.069	.
		[0.031]			[0.035]			[0.058]	
(1) Para * male	.	-0.011	.	.	0.061	.	.	-0.05	.
		[0.037]			[0.042]			[0.068]	
[redacted] Amapa	.	0.362	.	.	0.505	.	.	0.209	.
		[0.105]			[0.118]			[0.189]	
(1) Amapa * male	.	-0.202	.	.	-0.233	.	.	0.005	.
		[0.126]			[0.141]			[0.233]	
[redacted] Matto Grosso Sul	.	0.051	.	.	0.009	.	.	0.146	.
		[0.045]			[0.050]			[0.085]	
[redacted] Matto Grosso Sul * male	.	-0.062	.	.	0.054	.	.	-0.200	.
		[0.052]			[0.059]			[0.096]	
(1) Matto Grosso	.	0.111	.	.	0.106	.	.	0.222	.
		[0.052]			[0.055]			[0.103]	
[redacted] Matto Grosso * male	.	-0.036	.	.	0.055	.	.	-0.168	.
		[0.059]			[0.065]			[0.115]	

Dealing with poor students

Claudio de Moura Castro
Sonia Dantas Guimaraes
Joao Batista Araujo e Oliveira
Sergio Costa Ribeiro

March 1991

Introduction

Good schools for poor students are not part of the experience of the overwhelming majority of the Brazilian population. To the extent that they exist, they will certainly be an exception to the usual pattern. We usually find good private school for the rich, contrasting with poor, and bad public schools for the poor. Exceptions only confirm the rule. However, these same exceptions might also suggest that improvements in the public school system could lead to better education for poor students.

We do not assume that these schools are good, or that they make any difference for these students. We also do not attempt to defend the private school system, particularly because it would be hard to define as private a school financed by a public enterprise. Rather this paper attempts to examine these schools, how much they cost, what they are, and whether they make any difference for the children.

The schools focused in this paper belong to an unique group. They are schools financed by large corporations working in heavy construction, mining and other industrial fields in remote areas of Brazil and abroad. They cater to the children of their workers, thus mixing in the same schools the children of higher level staff (about 10-20%) and the children of the lowest paid workers. They are managed by the Sistema Pitagoras de Ensino, a private

education group, which attempts to keep quality standards comparable with those of their own elite school located in Belo Horizonte, a major Brazilian capital, Belo Horizonte. These schools are compared with the public schools to which the children of the workers would go, if they were not available. Even though they will be treated here as private schools, it is hard to really define them as such, given that most of the sponsoring firms are public enterprises.

The paper thus compares a set of schools operating under exceptional circumstances with those schools that one can expect in extremely poor and remote Brazilian towns. Hence, it compares some unusual privâte schools for the poor with the normal public schools. It addresses three major questions. First, it asks whether these private schools make any difference for the poor. Second, it asks whether the gains are worth the cost. Third, it speculates about what makes the difference for the poor in these schools.

Sampling and methodology.

Three out of the thirteen plant-compound schools presently operated by the Pitagoras group were chosen, and compared with three municipal and three state schools. Two campuses of the elite school in Belo Horizonte were also included for sake of comparison, thus totalling 11 schools in the sample. Details about the sampling process are described elsewhere (Castro, Guimaraes, Oliveira and Ribeiro, 1991). Suffice it to say that this was not meant to be a random sample. We are dealing with exceptions, having chosen the three private schools which would increase the variety of situations and permit relevant comparisons with public schools. The public schools chosen were those closer to the private schools of our sample. Overall, the poor students in these plant-compound schools come from backward, remote rural areas.

Data were collected from students in grades 1 and 3. Tests of Portuguese and Math previously developed by the Fundacao Carlos Chagas were used. The main features of these tests is that they correspond to the

curricula effectively utilized by average Brazilian public school. It is worth mentioning that the test questions were prepared by teachers of public schools and revised by the Carlos Chagas Foundation.

These tests were then validated by means of a large sampling of Brazilian schools. However, the sampling mechanism does not allow us to say that. these schools are technically representative of the population of Brazilian public schools (H. Marelin Vianna, "Avaliação do Rendimento de Alunos do Primeiro grau da Rede Publica em Vinte Cidades" , Educação e Seleção, n. 19 1989). What can be said is that most of the students in this sample leave in middle and large-size cities. These are neither the leading public schools nor the backcountry schools that are typical of our sample.

Additional questionnaires applied to the students provided data on the socio-economic background and other individual and family characteristics. In addition, several questionnaires adapted from the ECIEL studies (ref.) were used with principals, teachers of the grades studied as well as with administrators and other personnel from the sponsoring firms. School data were also collected and checked with administrative and financial services of the sponsoring firms, the educational group headquarters, as well as interviews with local authorities. A checklist was also used to compute the costs of the buildings and instructional materials in each school. The results of the cost analysis are presented elsewhere (Castro et alia, op. cit).

There were two major reasons to choose grades 1 and 3. First, standardized tests and comparative results were readily available from the Carlos Chagas Foundation. Second, most Brazilian students seldom go beyond these grades, thus maximizing the relevance of our comparisons. The tests were administered from late October 1990 on, within a space of three weeks, so that not much differences in total amount of teaching during the year would be present. One of the authors participated directly in most of the data gathering, and made extensive interviews and observations which are partly incorporated in the present paper.

Description of the schools

Public schools in our sample repeat the usual pattern of Brazilian public schools. Teachers earn very little, with salaries in the range of one to two minimum wages for a 20/hour/week contract; the schools are physically decrepit, the exceptions confirming the norm. There are no adequate facilities, some schools do not even have enough desks for teachers and students. Instructional materials very often are not available - absence is more often the case. Principals describe themselves as impotent to take any relevant decision concerning administrative, financial or pedagogic matters. The schools have no autonomy over such matters. These schools typically have a fairly large number of janitors, and yet only one had pedagogical or technical staffs. Students are mostly from low socio-economic backgrounds. Municipal schools are always works off that state schools, even in the sense that they do not have basic inputs, such as desks or textbooks. State schools,with one exception, seem to have at least the bare minimum.

By contrast, the private schools have a number of unique features. Pitagoras started the operation of such schools in 1966. It presently operates 18 plant-compound schools, with a total population of 20,000 students, scattered over 7 states in Brazil and two countries. Only three schools buildings are owned by the group, the others are managed under contractual arrangements.

The norms for each contract with sponsoring firms vary along a number of dimensions, including financial arrangements. However, the Pitagoras group explicitly attempts to keep the same basic rules and standards of the elite schools at headquarters. But other than that, the system is highly decentralized and school principals enjoy much autonomy over most matters, including pedagogic ones.

Few schools adopt the curricula, materials and tests of the central school in Belo Horizonte. Most develop their own materials, adapt some of the ones used in the capital, or buy books from the market. The group attempts to control quality through a careful selection of the school principal

- typically a seasoned teacher and administrator with experience with the group's norms, values and standard operating procedures.

Five schools from this group were selected. The first two are the elite schools in Belo Horizonte. They are typical Brazilian elite schools, fully equipped and staffed, offering courses from K-12, operating 5 hours a day, 25 hours per week. A major difference with all the other schools, including the other plant-compound schools, is that most teachers have higher education degrees. The other three schools in this group are plant-compound schools.

The School in Tucurui belongs to an hydroelectric public enterprise which started its operations in 1973. The private system was invited to run the schools in 1986. During the peak of the construction, 14 school complexes were simultaneously operating at various sites, but only 2 such units remain open. These schools have been phased out in the recent past. These schools operates in three shifts. In 1990, they went from 4672 students enroled in the beginning of year to 2398 in the end. As a result of these decaying enrollments, other things changed. Students, who were previously segregated according to location and other criteria which reflected the socio-economic background became mixed in these two schools. The school had to adapt to this new circumstance, and develop new strategies to deal with poorer students. As a result of this, noticeable changes were observed in the repetition rates, which went from 58% in the first grade for the poor students-only schools in 1988 to 38% in the heterogeneous school. At the time of the data collection, the issue of integration was was still very much alive, polarizing most of the efforts of management and staff, and in spite of the overall climate of de-mobilization. One of the consequences of these processes is the gradual departure from some practices of the elite school, and a more adaptive approach to the local realities.

The School in Niquelandia belongs to a mining corporation, which started its operations in 1980. The school was opened in 1983 and was previously operated by the firm with some assistance from a Federal University. Pitagoras took over the existing school in 1990, the year of the study, and inherited most of the locally recruited staff. The school then had

554 students, and changes in staffing and pedagogical practices were still beginning to be implemented. They reflect some of the previous experience acquired in Tucurui with mixed, heterogeneous classes. However, they were applied in the opposite direction, in the sense of grouping students by age, which, ultimately, reflects learning ability and competence. However, these groups receive special treatment, in the hope of increasing they changes of success.

The School in Teofilandia was set up by Pitagoras in 1989. It belongs to a major public corporation in the field of mining. It presently enrols 532 students. Compared with the two other schools in this group, it attempts to closely follow the standards and practices of the elite school in Belo Horizonte. For example, recruitment excluded most of the local teachers, only 5 out of 200 local candidates being accepted. This school clearly distances itself from the local community in the sense of the ethos and norms that it attempts to maintain. It relies almost exclusively on instructional materials from the headquarters. For example, the school re-classifies all students according to levels of competence and irrespective of previous school attainment. As a result, classes are slightly more homogeneous, from an achievement point of view.

All these private schools are perceived as different from the public schools by the local communities. They are well built, well maintained, and well staffed. One of the schools allows private students to attend, but the others are restricted to the children of the workers. As a result, students have to leave these schools when their parents quit their work, thus creating major problems for the students. Principals in the public schools remarked the shock experienced by students returning to public schools, due to the difference in the level of resources between these two groups.

Results

The major purpose of this investigation was to identify school effects on poor students. Thus, family and educational background of each student had to be controlled. Achievement in the Portuguese and Mathematics test were also computed. (See Appendix for sample questionnaire).

Table 1 displays the means on the Portuguese and Math tests for each school, grouped in private and public as well as the results obtained by the Carlos Chagas Foundation in the sample utilized to validate their tests. Since the differences between the Pitagoras school at headquarters and the others were quite important, they were presented separately.

The first results that called our attention and surprised us were the comparisons between the entire set of the schools in our sample (excluding the private school at headquarters) and those of the Carlos Chagas Foundation. While not a technically random sample of the Brazilian population of schools, this sample was meant to reflect a perception of the "average" Brazilian public school. And in fact, this is consistent with the differences in scores observed between this average and those of the elite school at the headquarters of the private system. Indeed, the elite school shows results which are much higher than those of the Chagas sample.

But even more surprising are the comparisons between the Carlos Chagas data and those of our sample - except for the results from the Belo Horizonte schools. The distances between averages are still higher. Comparing the distributions of the schools in our sample and those of the Chagas sample we find practically no overlap between the two. In other words, there are practically no students in our sample obtaining scores which reach the mean of the Chagas sample. Even when we examine the Chagas sample broken down by towns, we do not find any in which the means are as low as those of our sample. These are extraordinary results. These differences are enormous and will have to frame all further analysis.

If we consider public schools alone, and as our studies indicate (see Castro et alia) the public schools in our sample are not particularly worse off in equipment and overall situation compared to the average Brazilian poor school. They are in very bad shape but the average Brazilian public school is not significantly better (see Castro & Fletcher, "As escolas que os brasileiros frequentam". IPEA, 1986). Hence, if the test results are so much worse, this means that the clientele must lie further away from the Brazilian mainstream culture. In other words, these are schools catering to the most

remote and isolated groups, sharing even less of the values, norms and cognitive styles that are being promoted by the schools.

Given the strong relationship between socio-economic status and the kind of learning that is measured by such tests, the net effect of the schools could not be compared in a simple and straightforward manner, as one could expect at first. This is a population that falls way below an average definition of "poor", particularly in the sense that our sample contains students from very remote, backward areas. Our socio-economic indicators cannot be used to compare the sample with that of the Chagas sample, as the latter did not include such socio-economic controls. But our own analysis and our field observations suggest that the populations in our sample are indeed below the average Brazilian "poor". Or, at least, they come from a very empoverished background, as compared to the more cosmopolitan character of the Carlos Chagas Foundation sample.

When we compare the private schools in these towns with those operated by the state and the municipalities, the results are not immediately obvious. As Table 1 indicates, the means between the private and the public schools are quite significant in most cases, indicating better results for the private schools. Yet, these results do not control for the socio-economic status of the students and there were a priori reasons to believe that the private schools could cater to a different clientele.

Table 2 displays the simplified box-plots for each school and sampled grade with the distribution of scores. Each page combines achievement and SES scores for a given grade (1th or 3th) and discipline (Portuguese or Math). The median is shown as a vertical bar. The box shows the range of 25% of the scores above and the 25% scores below the median. The horizontal line show the full range of scores observed. To construct this scale we combined the scale of ownership of consumer durable, characteristics of the dwelling and occupational status of the father.

As we analyze these pairs of graphs, two salient points emerge. The first is that the sample becomes critically small in terms of statistical significance of results. Many of the cells are really to small to tell if the differences are caused by chance or result from the nature of the parent-

population. The second point - not unrelated to the first - is that the results are not easy to interpret.

Nevertheless, some trends emerge from additional probing. Table 3 shows the significance of these comparisons in conventional ANOVA tests. In order to compute them we have broken down the samples into "higher" and "lower" socio-economic groups (we emphatically note that "higher" does not mean "high", as these are both very poor groups).

The municipal schools yielded consistently lower results in all analyses. This is in itself an important result. It shows that one can make education even worse with a school that is still less adequate. Municipal schools are clearly and consistently worse than those operated by the state, despite the fact that they are not significantly less expensive. In other words, compared to the mediocre and lackluster state-operated school, the municipalities do even worse. This is in itself a major finding.

But since our goal was to compare the best that public schools could offer with the private system, we eliminated completely the municipal institutions from the ANOVA presented in Table 3.

The first result is that both schools on Table 3 do a very poor job of educating the poor student. This was already evident in Table 1 and becomes even more explicit here.

But some further observations seem warranted. The private Niquelandia schools have been taken over by the Pitagoras system in 1990 (note that field work was conducted later in this same year). This means that they are still in a process of structuring their activities in the school and that the third year students did not attend schools managed by Pitagoras during their past educational career. Perhaps the Pitagoras philosophy is still not fully implemented, and that might help explain why none of the differences are significant. While the differences between the means exist and favour the private schools, the differences are not sufficiently large to be significant with sample cells that often have less than ten observations.

The Teofilandia school has been taken over by Pitagoras in 1989, the students having a maximum of two years of exposure to the new

administration. In these comparisons we find already some advantages resulting from attendance to the private school. Unfortunately, the Portuguese test could not be applied to first year students due to time constraints. On the whole we find differences in the expected directions but they are not always significant. As described earlier, this school follows very closely the practices of headquarters, being physically isolated from the town and quite well equiped for a school in such remote location.

Finally, Tucurui is the town in which the results are more clear cut. The private system operates there already for five years and has had a chance to implement its ideas and philosophy. In most cases, the private schools perform significantly better.

Overall, there seems to have a phenomenon of organizational learning over time: schools take time to structure, to consolidate, and to be able to differentiate their treatment to the clienteles they cater form. Our sample does not allow us to attribute any effect to the time students spent on Pitagoras' schools

One last comment seems warranted. As we examine the data contained in Table 4, there is a suggestion that these private schools are more effective in offering chances for the less deprived students than for the even more deprived (at least in mathematics). In other words, these schools offer something that the less poor can grab but that the very deprived students of these towns are unable to benefit. Let us not forget that the poor of our sample come from a distinctly rural background, as compared to the more cosmopolitan students of the Carlos Chagas Foundation. In other words, to reach this clientele that is so extraordinarily removed from the Brazilian mainstream society, it takes time, and it might not be enough to act as a conventional school.

To sum up this section, our data seem to indicate that schools make a difference. But there is no magic to this process. The more structured and experienced the school, the better the results. Learning to behave as such, however, takes enormous managerial effort and takes time to learn. As they are now, these schools seem to be more responsive to the better-off students, and less effective with the poorest ones. Even though these schools

try to depart from the practices of the school at the headquarters, it is not clear which differentiated practices are most suited to these students in the lower end of the SES scale.

Discussion

To sum up the analysis of our data, the following comments seem warranted:

Our research starts from the basic difficulty of capturing the net effect of schools. Students coming from richer and better educated families attend good schools while poor students coming from uneducated families hardly ever have access to schools of better quality. From the high degree of multicolinearity found in research comparing rich and poor students, very little can be said about the ability of better schools to help poor students to overcome their environmental deprivation. For that reason, we tried to identify situations in which poor students have access to better schools. The case of the Pitagoras system offered an interesting possibility. In these schools, there is the clear and explicit decision to offer, in far away places, an education which is as close as possible to the high standards of the schools they operate at headquarters.

The semi-experimental sample design indeed permited a good control of socio-economic variables. **We found schools catering to the poor that were well installed, had good equipment and showed the environment, structure and ethos that we associate with good schools**. But we got more than we asked for. We do not have a sample of the typical clientele that we could find in poor neighborhoods of middle or large-sized Brazilian towns. Instead, we got a sample of the end of the line.

The overall consequences of attending better schools are clear and undisputable, even when we break down the school into higher and lower socio economic classes. **Better schools breed students who read better and have a higher command of mathematics**. This is certainly not a trivial result.

Students from the Pitagoras schools, on the whole, perform better than those of the public (state and municipal) system.

Comparisons with the better performing public schools show differences that are not always significant but the pattern of superiority is clear.

The decision to choose only those Pitagoras schools that were geographically close to a public school lead us to select units that were operating under the Pitagoras management for a relatively short time. In one case, the school had been taken over by Pitagoras less than one year before the research, in another, less than two years and in the third, less than five years. The consequences of these differences in time of implementation show very clearly in the results. The longer and deeper the presence of Pitagoras in the school, the more important are the differences in learning. In other words, **the Pitagoras management learns as time passes and this shows in an improvement in the performance of its students. It takes time for schools to learn, as much as for students**.

Another interesting result comes from comparing the results for the poorer and for the less poor groups. Pitagoras schools are less effective the more distant the students are from their Belo Horizonte elite clientele. The group of the "less poor" students benefit more from the enriched Pitagoras environment and methods. After all, this is not that surprising. Pitagoras tries to do their best and adapt their methods to the much poorer clientele of the plant compounds but they still offer an instruction that is closer to the world of those that are less stranded from the Brazilian mainstream culture. In other words, **Pitagoras is less effective in dealing with clienteles that are too distant from their own culture**.

A more surprising result was the difference between the municipal and the state schools. **We found that the municipal schools performed consistently very badly**. In fact, their scores were always much worse than those of any school from our sample. By contrast, **two state schools performed better, in some cases getting very close to the Pitagoras schools**. This is not a trivial result either, as it illustrates that some **extra attention and better management pay handsomely in terms of better learning of the students**.

The scores in reading and mathematics indicate better results for the Pitagoras schools. Field observation captures differences in environment and in less measurable dimensions that are even greater than what the scores show. That, of course has a price. Hence, when we ask how much it costs to offer a superior education, two main results emerge. Pitagoras schools are significantly more expensive than those of the public system. In rough terms, these schools cost per student up to twice the cost of public institutions. However, these **costs are not any higher than those of countries with similar levels of economic development**.

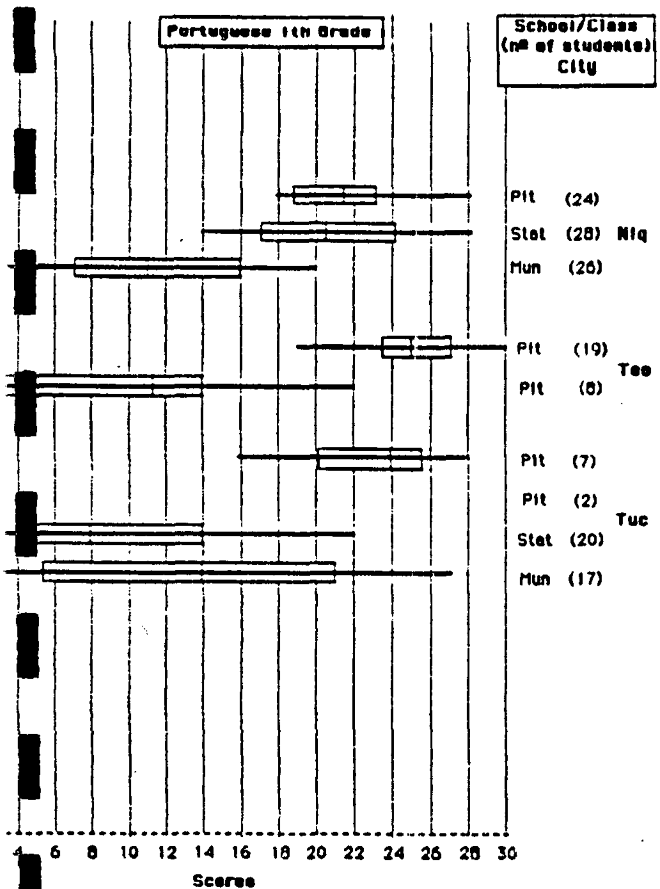
The more unexpected result come from the cost comparisons between state and municipal schools. **Both systems spend more or less the same per student but the performance differences can be very large indeed in favour of some state schools**. State schools can range from very bad to decent while **municipal schools are always deplorable** on all counts. Field observation confirms that the state school that showed higher scores operates closer to the image we have of a suitable school. By contrast, **everything seems to militate against education in the municipal schools**. This is a most important result as it demonstrates that up to a certain threshold of performance, it is not a matter of costs but of spending money and operating the schools in manners that favour learning. It is as if municipal schools adopted procedures in which the intentions of educating students were not at all present. Needless to say, **one does not have to spend one single extra cruzeiro to have a school that performs much better than those of the municipal system** (notice the similarity in the poor performance of the municipal schools across the three States included in the sample).

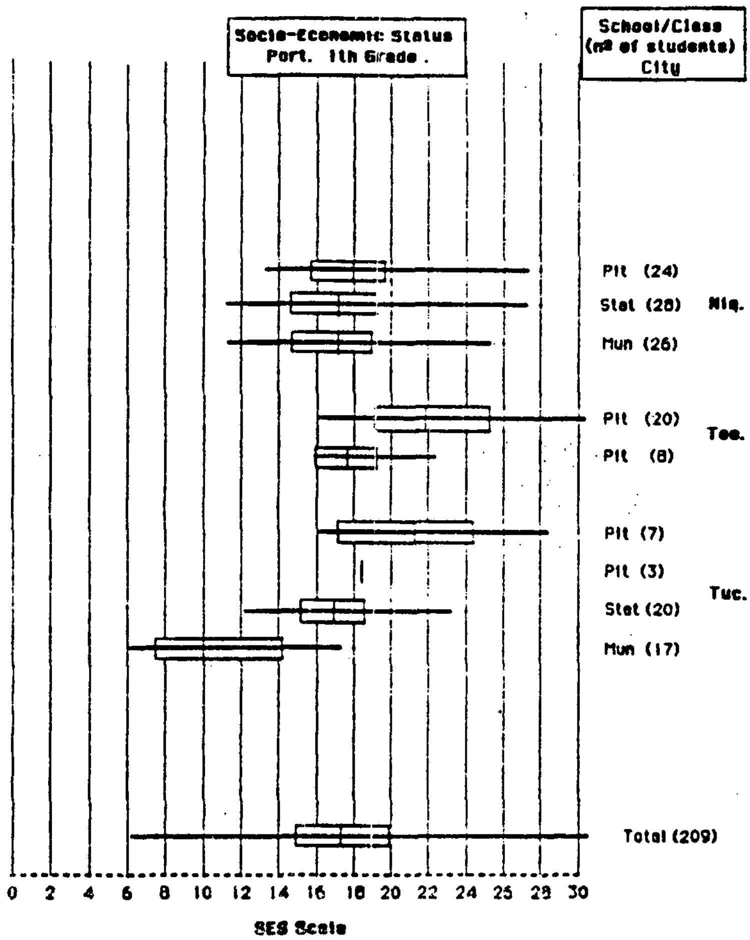
Those tragic results of the municipal schools lead us to examine more carefully their functioning. What emerged from the analysis is not a pattern but the absence in any pattern. There are no rules. For instance, the school can have no janitors or can have almost as many janitors as teachers. Class size, ratios between anything and anything vary wildly from one school to the next. **The only permanent fixture is that spending follows no rule in which a better learning could be a goal.**

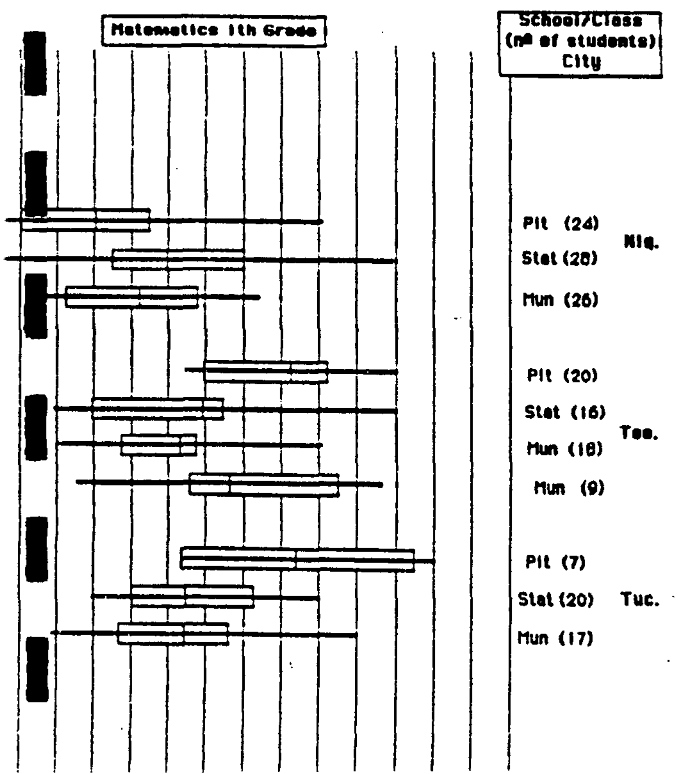
We wrote a paper about how to bring better education to the poor. The Pitagoras system is making a serious and earnest effort to offer a decent instruction in the schools they operate in the farthest frontiers of the country. This is confirmed by the overall environment of the schools and by the soul-searching discussions about the proper pedagogy to deal with poor students. There is a firm decision to bring their experience and their practices to the remote regions in which they operate - and they have all the interest in responding to the demands of the large enterprises that have given Pitagoras the contracts to operate the schools. **Yet, it is not clear how much of what they do in an upper class district in Belo Horizonte is transportable to these frontier schools**. Should they track students, should they use Emilia Ferreiro methods away from headquarters? Should they instead remain faithful to what worked best at headquarters? There is a slight evidence that by going further away from the practices of Belo Horizonte they can bring the levels of the students somewhat closer to Belo Horizonte. But they are far from having the answers. And unfortunately, it does not seem that others have the answers either. For tho se reasons, this seems to be an experiment that deserves following up closely.

Schools	T[redacted]Mean[redacted]			
	1th Grade		3th Grade	
	Portuguese	Matematics	Portuguese	Matematics
Colégio Pitágoras - Unidade CODEMIN (Niquelândia)	21.6	10.0	13.9	10.6
Colégio Estadual Paulo Francisco da Silva (Niquelândia)	20.9	12.6	13.2	11.4
Escola Municipal de 1º Grau Juscelino Kubitschek de Oliveira (Niquelândia)	11.6	10.8	10.4	8.4
Colégio Pitágoras - Cidade Jardim (Belo Horizonte)	27.0	25.7	20.8	21.6
Instituto Pitágoras de Educação - Pampulha (Belo Horizonte)	27.6	23.9	22.0	22.1
Centro Educacional Rafael Lopes de Araújo (Teofilândia)	-	13.3	10.2	6.4
Escola Pitágoras (Teofilândia)	20.6	17.5	15.4	13.1
Colégio Estadual de 1º Grau Plínio Carneiro da Silva (Teofilândia)	-	13.5	10.6	10.1
Escola Municipal de 1º Grau Gumercindo Gomes Pereira (Tucuruí)	13.6	13.3	10.5	4.9
Escola Estadual de 1º Grau Deputado Raimundo Ribeiro de Souza (Tucuruí)	10.1	13.8	14.8	10.5
Colégio Pitágoras - Escola de 1º Grau e Educação Especial S. Pedro de Alcobaça (Tucuruí)	23.1	18.5	-	-
Colégio Pitágoras - Unidade Jorge Antonello 1º e 2º Graus (Tucuruí)	-	-	18.6	14.9
Mean for All Sample	**19.54**	**15.72**	**14.58**	**12.18**
Colégio Pitágoras - Unidade CODEMIN (Niquelândia)	21.6	10.0	13.9	10.6
Escola Pitágoras (Teofilândia)	20.6	17.5	15.4	13.1
Colégio Pitágoras - Escola de 1º Grau e Educação Especial S. Pedro de Alcobaça (Tucuruí)	23.1	18.5	-	-
Colégio Pitágoras - Unidade Jorge Antonello 1º e 2º Graus (Tucuruí)	-	-	18.6	14.9
Means for Colégios Pitágoras whithout/Belo Horizonte	**21.73**	**15.35**	**15.97**	**12.89**
Colégio Pitágoras - Cidade Jardim (Belo Horizonte)	27.6	23.9	22.0	22.1
Instituto Pitágoras de Educação - Pampulha (Belo Horizonte)	27.0	25.7	20.8	21.6
Means for Colégios Pitágoras in Belo Horizonte	**27.28**	**24.80**	**21.41**	**21.85**
Colégio Estadual Paulo Francisco da Silva (Niquelândia)	20.9	12.6	13.2	11.4
Escola Municipal de 1º Grau Juscelino Kubitschek de Oliveira (Niquelândia)	11.6	10.8	10.4	8.4
Centro Educacional Rafael Lopes de Araújo (Teofilândia)		13.3	10.2	6.4
Colégio Estadual de 1º Grau Plínio Carneiro da Silva (Teofilândia)		13.5	10.6	10.1
Escola Municipal de 1º Grau Gumercindo Gomes Pereira (Tucuruí)	13.6	13.3	10.5	4.9
Escola Estadual de 1º Grau Deputado Raimundo Ribeiro de Souza (Tucuruí)	10.1	13.8	14.8	10.5
Means for Public Schools	**14.03**	**12.88**	**11.61**	**8.60**
Means for public schools in 20 Brasilian cities (Fundação Carlos Chagas) *	**20.42**	**22.74**	**18.54**	**20.45**

Table 2a


Portuguese 1th Grade
School/Class (nº of students) City
Pit (24)
Stat (28) Niq
Mun (26)
Pit (19)
Tee
Pit (8)
Pit (7)
Pit (2)
Tuc
Stat (20)
Mun (17)
4 6 8 10 12 14 16 18 20 22 24 26 28 30
Scores


Socio-Economic Status Port. 1th Grade .
School/Class (nº of students) City
Pit (24)
Stat (28) Niq.
Mun (26)
Pit (20)
Tee.
Pit (8)
Pit (7)
Pit (3)
Tuc.
Stat (20)
Mun (17)
Total (209)
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
SES Scale


Matematics 1th Grade
School/Class (nº of students) City
Pit (24)
Niq.
Stat (28)
Mun (26)
Pit (20)
Stat (16)
Tee.
Mun (18)
Mun (9)
Pit (7)
Stat (20) Tuc.
Mun (17)

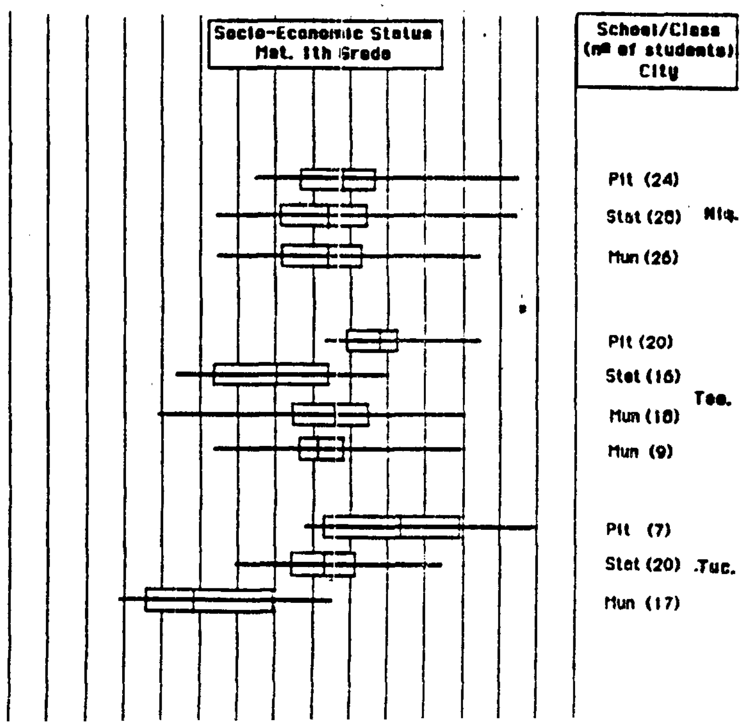
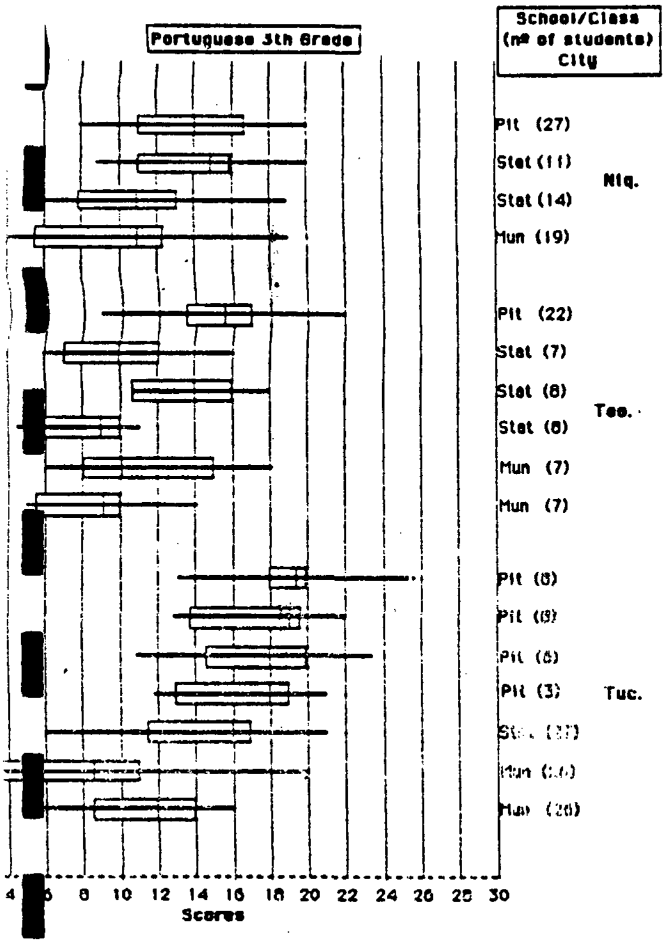
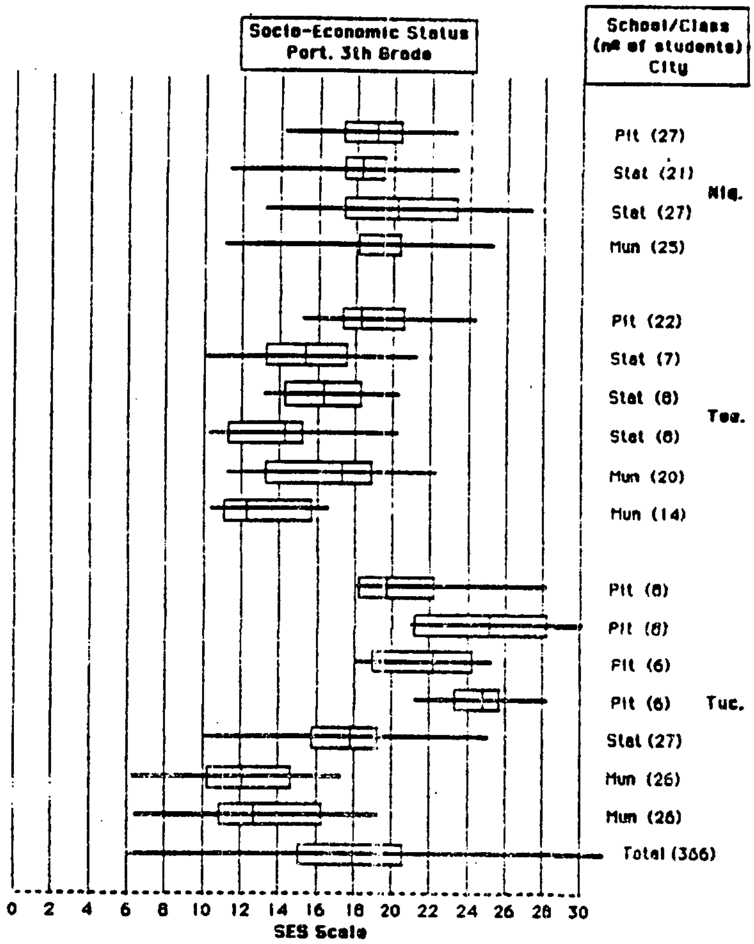
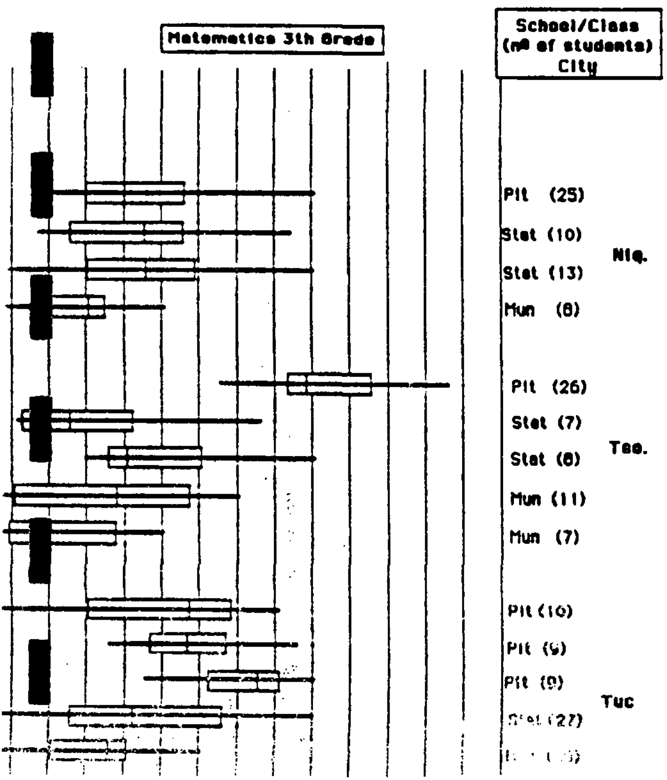

Socio-Economic Status Mat. 1th Grade
School/Class (nº of students) City
Pit (24)
Stat (28) Niq.
Mun (26)
Pit (20)
Stat (16)
Tee.
Mun (18)
Mun (9)
Pit (7)
Stat (20) Tuc.
Mun (17)

Table 2b


Portuguese 3th Grade
School/Class (nº of students) City
Pit (27)
Stat (11)
Stat (14)
Mun (19)
Niq.
Pit (22)
Stat (7)
Stat (8)
Stat (8)
Mun (7)
Mun (7)
Tea.
Pit (8)
Pit (8)
Pit (6)
Pit (3)
Tuc.
4 6 8 10 12 14 16 18 20 22 24 26 28 30
Scores


Socio-Economic Status Port. 3th Grade
School/Class (nº of students) City
Pit (27)
Stat (21)
Stat (27)
Mun (25)
Niq.
Pit (22)
Stat (7)
Stat (8)
Stat (8)
Mun (20)
Mun (14)
Tea.
Pit (8)
Pit (8)
Pit (6)
Pit (6)
Stat (27)
Mun (26)
Mun (26)
Tuc.
Total (366)
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
SES Scale


Matematics 3th Grade
School/Class (nº of students) City
Pit (25)
Stat (10)
Stat (13)
Mun (8)
Niq.
Pit (26)
Stat (7)
Stat (8)
Mun (11)
Mun (7)
Tea.
Pit (10)
Pit (9)
Pit (9)
Stat (27)
Tuc

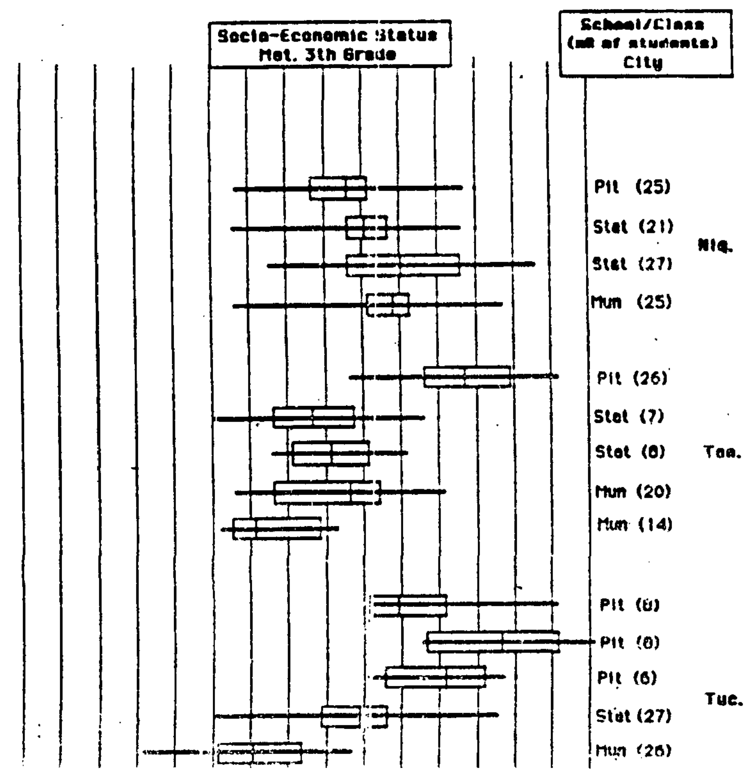

Socio-Economic Status Mat. 3th Grade
School/Class (nº of students) City
Pit (25)
Stat (21)
Stat (27)
Mun (25)
Niq.
Pit (26)
Stat (7)
Stat (8)
Mun (20)
Mun (14)
Tea.
Pit (8)
Pit (8)
Pit (6)
Stat (27)
Mun (26)
Tuc.

Table 3

ANOVA test for equality of the means.

SES Level		Probability of Significance			
		Port. 1th G.	Mat. 1th G.	Port. 3th G.	Mat. 3th G.
Lower	Pitag./State	0.001	n/s	0.02	n/s
Higher	Pitag./State	0.027	n/s	0.01	0.0001

SES Level		Probability of Significance			
		Port. 1th G.	Mat. 1th G.	Port. 3th G.	Mat. 3th G.
Lower	Niquelândia	n/s	n/s	n/s	n/s
	Teofilândia	-	n/s	0.0003	0.007
	Tucuruí	0.008	n/s	0.003	n/s
Higher	Niquelândia	n/s	n/s	n/s	n/s
	Teofilândia	-	n/s	n/s	0.001
	Tucuruí	0.004	n/s	0.02	0.002

A Note on

BRAZILIAN SCHOOLING INVESTMENTS IN AN INTERNATIONAL PERSPECTIVE: WHERE DOES BRAZIL FIT?

by

Jere R. Behrman and Ryan Schneider*

March 22, 1991

Prepared for conference on

Education, Growth and Inequality in Brazil

Rio de Janeiro

25-26 March 1991

*Behrman is the William R. Kenan, Jr. Professor of Economics at the University of Pennsylvania, Philadelphia, PA 19104-6297 USA and, for the 1990-1 academic year, Arnold Bernhard Visiting Professor of Economics at Williams College, Williamstown, MA 01267 USA. Schneider is a senior at Williams College.

Schooling is widely thought to be an important human capital investment with significant potential for increasing growth and equity in developing countries. While there may be some disagreements about just how effective schooling is for these purposes and how well empirical estimates of the impact of schooling control for factors such as ability, motivation, schooling quality, and family background, there seems to be widespread agreement that the effects of schooling are likely to be substantial and important in many developing country contexts.[1]

This conference is devoted to examining a number of dimensions of the Brazilian experience with schooling and economic goal attainment in Brazil. The purpose of this note is to provide some very simple descriptive cross-country perspective regarding how Brazilian schooling investments compare with those in other developing countries based on data in World Bank (1990). Such description is limited by the data that are available and by measurement problems for those data. For example, only enrollment rates and literacy rates are available to represent per capita income, without control for grade repetition nor for the quality of the schooling,[2] and there are differences in definitions of variables such as literacy across countries. Also only per capita GNP at official exchange rates are available for many of the countries, so these data are used instead of purchasing-power-parity measures of income per capita. Even aside from such data problems, such description in itself does not answer any very profound questions, nor does it provide a very firm basis for comments on causality. But it may provoke some questions about why the Brazilian experience differs, to the extent that it does, from that of other developing countries.

Section 1 presents the enrollment and illiteracy rates for Brazil and, for comparison, for all low-income and middle-income countries, for 1965 and 1987. Section 2 summarizes where Brazil stands with regard to total schooling enrollment rates and the implied expected years of schooling for a synthetic cohort in comparison with cross-country regressions that control for GNP per capita in each of these two years, as well as for a indicator of the schooling price for 1987. Section 3 describes similar relations for males and females and for the differences between them. Section 4 summarizes where Brazil stands ~~with regard to changes in these variables between~~ 1965 and 1987 in comparison with cross-country regressions that control for changes in real GNP per capita during this period and for all country-specific fixed effects.

Section 1. Brazilian Enrollment and Illiteracy Rates in 1965 and 1987

Table 1 gives primary, secondary and tertiary school enrollment rates and adult illiteracy rates

---

[1]For surveys of the impact of schooling on growth and distribution, see Behrman (1990a,b,c), Colclough (1982), Eisemon (1988), Haddad, Carnoy, Rinaldi and Regel (1990), King (1990), King and Hill (1991), Pscharopoulos (1985, 1988), Schultz (1988, 1991), and World Bank (1980, 1981, 1990, 1991).

[2]Some data are given on primary net enrollment rates and primary pupil-teacher ratios, but data are missing for many countries in each case (including Brazil in the second case) so I do not include them in this note.

for Brazil and for low- and middle-income country groups.[3]

Enrollment rates reflect one important component of the current investment in schooling. In 1965 Brazilian total and female primary schooling enrollment rates were over 100 per cent of the age group, with no gender gap. These rates were high relative to the means of 92 per cent for the total and 86 percent for females for current middle-income countries. The secondary enrollment rates were 16 per cent both for the total and for females, again with no gender gap. These rates were low, more so for males, relative to the means for current middle-income countries of 26 per cent for the total and 22 per cent for females. The total tertiary enrollment rate of 2 per cent also was low relative to the 6 per cent mean for current middle-income countries. Thus in 1965 Brazil seemed to have relatively high primary school enrollment rates and a relatively small gender gap at both the primary and the secondary levels, but relatively low post-primary enrollment rates.

Table 1. Brazilian Country Mean Enrollment and Literacy Rates in 1965 and 1987

	Percentage of age group enrolled in education										Adult illiteracy rates in 1985	
	Primary				Secondary				Tertiary			
	Total		Female		Total		Female		Total		Total	Female
	1965	1987	1965	1987	1965	1987	1965	1987	1965	1987		
Brazil	108	103	108	..	16	39	16	45	2	11	22	24
Low-income	73	104	..	95	20	37	..	29	2	..	44	58
Middle-income	92	104	86	101	26	54	22	54	6	17	26	31

In 1987 Brazilian total primary schooling enrollment rates were slightly lower (presumably reflecting fewer enrollers outside of the normal primary-school age range), but still over 100 per cent and at about the same level as the means for all developing countries. The Brazilian total secondary school enrollment rate had increased substantially from 16 to 39 per cent by 1987, but the latter level and the change both were below those for the means for middle-income countries. The Brazilian female secondary school enrollment rate had increased more than the total, from 16 to 45 per cent.

[3]According to the World Bank (1990), in 1987 Brazil was high in the lower-middle-income country group, and was 45th from the top (or 76th from the bottom) in terms of GNP per capita in that year (out of a total of 121 countries). In terms of annual growth rates in GNP per capita for the 1965-1987 period, Brazil (with 3.6 per cent) tied Egypt for 14th place among the 101 countries for which such rates are given, and had the highest growth rate in the Western Hemisphere. The countries with higher reported rates are: Botswana (8.6 per cent), Singapore (7.2), Hong Kong (6.3), Republic of Korea (6.8), Oman (6.4), China (5.4), Lesotho (5.2), Hungary (5.1), Japan (4.3), Thailand (4.0), Malaysia (4.0), Saudi Arabia (3.8), and Cameroon (3.7)

This implied the opening up of a gender gap favoring females at this level of schooling, a relative rarity among developing countries.[4] But despite the relatively large increase in female as compared with male Brazilian secondary school enrollment rates, Brazilian female secondary school enrollment rates in 1987 and their changes between 1965 and 1987 were smaller than the changes in the means for middle-income countries. Brazilian tertiary enrollment rates also increased substantially from 2 to 11 per cent between 1965 and 1987, but also both the change between these years and the absolute level in the latter year were below the means for middle-income countries.

Adult illiteracy rates reflect (inversely) one possibly important component of the stock of human capital. In 1985 the Brazilian total adult illiteracy rate was 22 per cent, somewhat below the 26 per cent for the mean for middle-income countries. That for females was slightly higher at 24 per cent, but somewhat more below the 31 per cent for the mean for middle-income countries. Thus, the relatively strong investment in basic schooling in Brazil apparently resulted in somewhat below average illiteracy by the mid 1980s, with a smaller than average gender gap in illiteracy. However the illiteracy criteria probably refers only to basic -- probably primary -- schooling, not to the post-primary schooling levels at which Brazilian schooling investments long have been less than the means for middle-income countries.

Section 2. Comparisons of Total Brazilian Schooling Investments with Cross-Country Experience

We now turn to comparisons of Brazilian schooling investments in 1965 and 1987 with cross-country regressions based on all countries in the World Bank (1990) that have the necessary data. The dependent variables are the total enrollment rates for each of the three schooling levels and the expected years of schooling implied by those rates for a synthetic cohort that experience such schooling enrollment rates.[5] The cross-country regressions control for per capita income in the relevant year, both with a dichotomous variable for the country group (i.e., lower-middle-income, upper-middle-income, and high-income, with low-income as the reference group) into which each country falls and with a quadratic in per capita GNP to capture variations within these broad country groups. The income effect can be interpreted as representing the income demand effect for school investments, under the presumption that such effects are transmitted through the political system for public schools

[4]Gender gaps favoring females in secondary school enrollment levels also are reported in World Bank (1990) only for Sri Lanka, the Philippines, the Dominican Republic, El Salvador, Botswana, Jamaica, Ecuador, Chile, Costa Rica, Poland, Panama, Nicaragua, Argentina, Venezuela, Trinidad and Tobago, and Romania. The magnitude of this gap is larger than in Brazil only for the Dominican Republic and Nicaragua.

[5]The numerators for the enrollment rates for the three schooling levels are respectively the number of children 6 to 11, 12 to 17 and 20 to 24 years old. Therefore the expected years of schooling for a synthetic cohort that faced the enrollment rates for these three schooling levels recorded for Brazil in 1987, for example, is $9.07 = 1.03*6 + 0.39*6 + 0.11*5$, given the enrollment rates for that year in Table 1.

in addition to any more direct demand effects for private schools.[6] For 1987, in addition, the regressions include a proxy for the relative price of schooling in the form of the adult literacy rate.[7] The argument for interpreting this variable as a price variable has three components: First, that the relative price of sufficiently skilled labor to staff the school levels that most students attend is inversely related to the share of the adult population that has such skills. Second, that the adult literacy rate is a good proxy for the relative size of the adult population with such skills.[8] Third, once again, that through the political process for public schools in addition to more direct market effects for private schools, the relatively scarcity of potential staff ("prices") affects the demand for schooling.[9]

Table 2 gives the cross-country regression estimates both for 1965 and for 1987 for the total enrollment rates for each of the three school levels and for the expected years of schooling for a synthetic cohort.

For 1965 the regressions are consistent with from half to two-thirds of the sample variance. For primary school enrollments in that year, the indicators of the country group by income have a substantial effect, without an added significant impact of the quadratic in per capita income. But the coefficients for the country groups indicate that the basic significant difference is between the average enrollment rates of the low-income countries and all other countries. For expected years of schooling, in addition to significant differences between low-income and other country groups, there is a significant quadratic positive (at a declining rate) effect of schooling. For post-primary schooling enrollments, in addition to significant country group effects between low-income and other countries and the quadratic in income, there is a significant difference between lower-middle-and upper-income countries.

For 1987 the regressions are consistent with more (than for 1965) of the cross-country variance -- from three-fifths to seven eights of those variances. This greater consistency with the data is due largely to the inclusion of the literacy rate, which appears to be highly significant in each of the

[6]Schultz (1987, 1988) and Behrman (1987), for example, give such an interpretation. From a micro perspective, however, if schooling were purely an investment and if capital markets were perfect, income would not enter into the schooling investment decision though interest rates would. To date we have not explored the possibility that interest rates enter into these relations.

[7]We actually use the literacy rate for 1985 because that is what is available in World Bank (1990). Since adult literacy refers to a stock concept, we do not think that the two-year difference between 1985 and 1987 has much effect on the estimates. In fact, to the extent that there is some lag in adjustment to relative scarcities, it may be preferable to use a lagged value for the relative stock of literate adults. We do not use this variable for 1965 because it is not available in World Bank (1990).

[8]Presumably how good a proxy it is depends in substantial part on the levels of schooling which most students attend. That is, it probably is a better proxy for primary and lower middle schooling than for upper secondary and tertiary schooling.

[9]Behrman (1987) makes all three of these assumptions as well. Schultz (1987, 1988) makes the last, but uses expenditure data per teacher to represent the price effect.

Table 2. Cross-Country Regressions for Total Schooling Investments 1965 and 1987

Depend. Var.	Const.	Inc.	Inc^2	Country Group: Lower-Middle	Country Group: Upper-Middle	Country Group: High	Adult Lit. Rate	R^2	N	F	Residual for Brazil
1965											
Enroll. Rates											
1 Primary	47.0	0.0007	-6.25*10-9	39.62	48.76	47.71		.50	96	20.0	20.7
		(0.37)	(-0.13)	(7.05)	(6.06)	(4.00)					
2 Second.	6.17	0.0031	-7.99*10-8	9.62	17.42	31.34		.67	94	39.2	-2.6
		(2.80)	(-2.66)	(2.84)	(3.45)	(4.30)					
3 Tertiary	0.30	0.0012	-4.27*10-8	2.51	4.06	6.74		.53	88	20.7	-1.9
		(1.91)	(-1.26)	(1.88)	(2.02)	(2.23)					
4 Years of School	3.33	0.00071	-3.5*10-8	2.84	3.10	3.66		.63	85	29.9	0.7
		(2.78)	(-2.57)	(5.432)	(3.75)	(3.06)					
1987											
Enroll. Rates											
5 Primary	35.8	0.00072	-5.45*10-8	8.34	-3.06	-8.03	0.80	.60	83	21.4	3.5
		(0.157)	(-0.331)	(1.43)	(-0.3)	(-0.27)	(7.57)				
6 Second.	-1.64	0.0036	-9.02*10-8	8.20	22.1	18.43	0.459	.85	82	79.1	-2.3
		(1.01)	(-0.72)	(1.81)	(2.3)	(0.82)	(5.61)				
7 Tertiary	-6.10	0.00022	-3.23*10-9	6.97	9.13	18.47	0.182	.64	80	24.2	2.5
		(0.13)	(-0.065)	(2.23)	(1.6)	(1.4)	(3.14)				
8 Years of School	1.59	0.00002	7*10-10	2.12	2.76	3.40	0.082	.86	75	76.3	-1.0
		(0.060)	(0.053)	(3.80)	(2.5)	(1.27)	(8.25)				

Data source: World Bank (1990). Enrollment rates (and therefore residuals for enrollment rates) are in percentages. Years of school (i.e., expected years of school for a synthetic cohort as defined in note 5) are in years (and therefore so are the residuals for these regressions).

regressions.[10] If the price interpretation is correct, there are strong price effect in these demand relations. The income effects are relatively imprecisely estimated, though the coefficients for the middle-income country groups are significantly positive at standard levels for post-primary school enrollments and for the expected schooling for a synthetic cohort.

How do Brazilian schooling investments compare with these cross-country regressions? The last column in Table 2 gives the estimated residuals for Brazil. In 1965 Brazil had a primary school enrollment rate that was 20.7 per cent above the regression line, but secondary and tertiary enrollment rates that were 2.6 and 1.9 per cent below the respective regression lines. The net result was an expected years of schooling 0.7 years above that predicted by cross-country experience. In 1987 Brazil had primary and tertiary-school enrollment rates that were 3.5 and 2.2 per cent above the respective regression lines, but a secondary enrollment rate that was 2.3 per cent below the regression line. The net result was an expected years of schooling 1.0 years below the regression line.

Section 3. Comparisons of Male and Female Brazilian Schooling Investments with Cross-Country Experience

We note in Section 1 that there appear to be some interesting gender patterns in the Brazilian schooling investments, particularly with larger investments in the secondary schooling of females than of males in 1987. In Section 2 we consider total schooling investments relative to cross-country regressions for 1965 and 1987. Now we turn to similar considerations, but with separate relations for males than for females.[11] Table 3 gives relations similar to those in Table 2, but with separate relations for males versus females. In addition this table gives estimates of the differences between the male and female primary and secondary enrollment rates and the expected years of schooling for synthetic cohorts.

Many of the patterns in these regressions are similar to those that are discussed in Section 2, without significant differences between the estimates for females versus males. Therefore, to avoid repetition, we here summarize only the results that indicate some significant difference between males

[10]Also, in preliminary regressions that are not presented here, female adult literacy rates has more consistency with the cross-sample variances and have larger point estimates than do total or male adult literacy rates. This suggests that the relative concentration of women in education results in a greater price-reducing effect for future generations of increasing the schooling of females in this generation than of increasing equally the schooling of males in this generation.

[11]Data are given in World Bank (1990) only for the female enrollment rates for primary and for secondary school in addition to the total enrollment rates for all three schooling levels. For the estimates that we present in this section we assume that the sizes of the relevant age cohorts are the same for males as for females in order to be able to estimate male primary and secondary enrollment rates. To calculate the expected years of schooling for synthetic cohorts we assume that there are no gender differences in the tertiary enrollment rates. This last assumption probably results in an underestimate of the gender gap for many countries, but has a limited impact on the overall estimates because of the relatively low enrollments for the tertiary (as opposed to the primary and the secondary) schooling levels in most developing countries (including Brazil).

Table 3. Cross-Country Regressions for Male and Female Schooling Investments 1965 and 1987

Depend. Var.	Const.	Inc.	Inc^2	Country Group: Lower-Middle	Country Group: Upper-Middle	Country Group: High	Adult Lit. Rate	R^2	N	F	Residual for Brazil
1965											
Enroll. Rates											
Primary											
1 Female	34.8	0.00061	-1.40*10-8	43.37	55.52	59.48		.52	94	21.2	29.3
		(0.30)	(-0.25)	(6.91)	(6.18)	(4.45)					
2 Male	56.2	0.0009	-4.2*10-9	38.67	44.43	37.79		.44	94	16.1	12.3
		(0.51)	(-0.08)	(6.78)	(5.44)	(3.11)					
3 M - F	21.4	0.0003	9.79*10-9	-4.70	-11.10	-21.70		.12	94	3.56	-17.0
		(0.24)	(0.26)	(-1.112)	(-1.183)	(-2.41)					
Second.											
4 Female	2.88	0.0032	-8.81*10-8	9.80	16.60	32.06		.67	93	39.6	0.4
		(2.90)	(-2.93)	(2.87)	(3.27)	(4.38)					
5 Male	8.24	0.003	-7.28*10-8	10.62	19.34	31.56		.64	93	34.4	-5.6
		(2.56)	(-2.26)	(2.90)	(3.56)	(4.03)					
6 M - F	5.35	-0.0002	1.53*10-8	0.82	2.74	-0.50		.00	93	0.87	-6.0
		(-0.26)	(0.86)	(0.41)	(0.91)	(-0.12)					
Years of School											
7 Female	2.40	0.00072	-3.57*10-8	3.04	3.47	4.39		.62	84	28.8	1.5
		(2.52)	(-2.38)	(5.18)	(3.77)	(3.30)					
8 Male	3.96	0.00073	-3.47*10-8	2.92	2.97	3.14		.62	84	28.3	0.0
		(2.94)	(-2.67)	(5.74)	(3.72)	(2.72)					
9 M - F	1.56	0.00000	1.02*10-9	-0.12	-0.50	-1.25		.06	84	2.15	-1.4
		(0.055)	(0.113)	(-0.332)	(-0.90)	(-1.56)					
1987											
Enroll. Rates											
Primary											
10 Female	31.6	0.00050	-5.76*10-8	8.71	-7.74	-10.3	0.92	.71	77	33.1	nd
		(0.109)	(-0.347)	(1.52)	(-0.6)	(-0.35)	(9.89)				
11 Male	37.4	-0.0010	1.37*10-8	5.43	-4.72	-5.57	0.85	.52	77	15.2	nd
		(-0.208)	(0.082)	(0.939)	(-0.38)	(-0.188)	(7.29)				
12 M - F	3.72	-0.0023	1.0*10-7	-4.54	1.51	6.25	0.84	.61	77	21.0	nd
		(-0.99)	(1.17)	(-1.56)	(0.24)	(0.41)	(8.068)				
Second.											
13 Female	-2.49	0.0025	-5.99*10-8	9.87	26.9	26.8	0.49	.89	74	101.8	5.3
		(0.73)	(-0.49)	(2.18)	(2.9)	(1.23)	(6.68)				
14 Male	-0.10	0.0027	-5.34*10-8	9.84	25.0	21.32	0.436	.82	74	55.9	-7.2
		(0.70)	(-0.385)	(1.92)	(2.4)	(0.861)	(4.32)				

15 M - F	-7.30	0.0008	-7.9*10-9	0.34	-1.23	-6.99	0.83	.67	74	26.6	-7.2
		(0.51)	(-0.14)	(0.16)	(-0.28)	(-0.68)	(10.12)				
Years of <u>School</u>											
16 Female	1.44	-0.00004	2.98*10-9	2.20	2.37	3.43	0.09	.90	66	105.3	nd
		(-0.122)	(0.208)	(4.06)	(2.1)	(1.31)	(10.50)				
17 Male	1.70	-0.00008	5.73*10-9	1.99	2.34	3.35	0.08	.83	66	53.3	nd
		(-0.187)	(0.360)	(3.35)	(1.9)	(1.15)	(7.43)				
18 M - F	-0.11	-0.00008	4.42*10-9	-0.29	-0.10	0.02	0.094	.70	66	26.1	nd
		(-0.401)	(0.602)	(-1.09)	(-0.2)	(0.014)	(8.55)				

Data source: World Bank (1990). Enrollment rates (and therefore residuals for enrollment rates) are in percentages. Years of school (i.e., expected years of school for a synthetic cohort as defined in note 5) are in years (and therefore so are the residuals for these regressions). "nd" means that the date are not available for Brazil. For the M - F regressions for 1987 the adult literacy variable is the difference between that for male and that for female.

and females.

For 1965, for the low- and middle-income country groups primary school enrollments tend to be higher by 21.4 per cent for males than for females, but there is no significant difference for high-income countries (regression 3).[12] For secondary school enrollment, the rates tend to be about 5 per cent higher for males than for females, with no significant differences across country groups (regression 6). The net effect is an average difference in expected years of schooling for a synthetic cohort of 1.6 years favoring males, though for high-income countries there is no significant difference at the 15 per cent level (regression 9).

For 1987, primary school enrollment rates are still higher for males than for females, but the gap has fallen substantially since 1965 to 3.7 per cent (8.3 per cent at the 15 per cent level for lower-middle-income countries, regression 12). For secondary school enrollment rates, the gap is reversed to -7.3 percent (regression 15). For expected years of schooling the net effect is an estimated gender gap that is virtually zero at -0.1. However these estimates for 1987 must be qualified as underindicators of the gender gap favoring males since also included in the estimates is the male-female difference in adult literacy rates. For the revision of this paper we will undertake estimates without the difference in adult literacy rates.

How does Brazilian schooling investment compare with the cross-country experience? The last column in Table 3 gives the residuals for Brazil. For 1965, Brazil had female and male primary schooling enrollment rates 29.3 and 12.3 per cent above the regression lines, with the male - female differential 17.0 percent below the regression line. The female secondary schooling enrollment rate was 0.4 per cent above the regression line and the male rate was 5.6 per cent below, with the gender gap at the secondary school level 6.0 per cent below. The net result was an expected years of schooling for females 1.5 years above the regression line and an estimate on the regression line for males, so the gender gap was 1.4 years below the regression line. For 1987, there are not data for the gender differences in the Brazilian primary schooling enrollment rates (and therefore in the expected years of schooling). For secondary school enrollment rates, Brazil was 5.3 per cent above the regression line for females and 7.2 per cent below the regression for males, with the net effect being 7.2 per cent below the regression line for the gender gap.

Section 4. Comparisons of the Changes in Brazilian Schooling Investments with Cross-Country Experience

The previous two sections have focused on how Brazilian schooling investments compared with the cross-country experience in 1965 and 1987. A related, but distinct, question is how did the changes in Brazilian schooling investments compare with changes in the cross-country experience. This question is of interest because of its dynamic nature and because the differenced estimates control

[12]The gap is given by the constant in this regression for all but the high-income countries because no other variables have significantly nonzero coefficient estimates. But the coefficient estimate for the high income country group dichotomous variable basically offsets the constant so that for this country group the gender gap is zero.

for unobserved additive country-specific fixed effects.[13]

Table 4 gives such estimates. For primary school enrollment rates they indicate an increase of 24.9 per cent on the average for the low-income countries, though significantly smaller increases for the middle- and high-income country groups.[14] These increases tend to be larger for females than for females, particularly for the lower-middle-income group. For secondary school enrollment rates the patterns are somewhat different. There are increases on the average are 12.0 per cent for low-income countries, but there are significantly larger increases for lower-middle-income countries (25.4 per cent) and still larger increases for upper-middle- and high-income countries (over 36 per cent on the average). Also for the low-income countries the increases are larger for males than for females on the average, though the opposite is the case for the middle- and upper-income country groups. For tertiary school enrollments, the average increase is relatively low for low-income countries (1.6 per cent), but significantly higher (10.3 per cent) for lower-middle-income countries and still higher (above 13.9 per cent) for the upper-middle- and high-income countries. This implies an gain in expected years of schooling of 2.0 for the low- and high-income countries, and of 3.4 years for the middle-income countries.

Once again, how does Brazil compare? Once again, the last column gives the residuals for Brazil. They suggest that Brazil was 20.3 per cent below the regression line for changes in primary school enrollments, 1.8 per cent below for changes in secondary school enrollments, 2.1 per cent below for changes in tertiary enrollments, and 1.5 years below the line for changes in expected years of schooling for a synthetic cohort. For secondary school enrollments, Brazil had a strong shift towards females in the gender gap relative to other country experience, with an increase for females 2.2 per cent above the regression line and one for males 7.7 per cent above the regression line. Thus the two characteristics of the Brazilian changes, relative to the cross-country changes, that stand out are: (1) small increases in schooling investment and (2) more of a shift towards females. The first point at the primary school level may reflect in part the relatively high primary school enrollments in 1965 (see Tables 1-3), but it also reflects relatively smaller increments in post-primary schooling.

[13]That is, these estimates, by controlling for such fixed effects, have less in the way of omitted variable bias (though perhaps more in the way of random measurement error bias) than do the estimates in Tables 2 and 3.

[14]There is an inverse relation between the point estimates of for the increases and the income levels, but the standard errors are large enough that they do not indicate significant differences among the lower-middle-income, upper-middle-income and high-income groups.

Table 4. Cross-Country Regressions for Changes Between 1965 and 1987 in Total, Male and Female Schooling Investments

Depend. Var.	Const.	Inc.	Inc^2	Country Group: Lower-Middle	Country Group: Upper-Middle	Country Group: High	R^2	N	F	Residual for Brazil
1987 -1965										
Enroll. Rates										
Primary 1 Total	23.86	-0.0000	-5.8*10-8	-8.49	-15.68	-19.39	.23	89	6.40	-20.3
		(-0.084)	(-1.175)	(-2.11)	(-2.74)	(-2.63)				
2 Female	24.56	-0.0002	-4.7*10-8	-3.46	-15.69	-18.24	.22	83	5.70	nd
		(-0.319)	(-0.945)	(-0.810)	(-2.649)	(-2.361)				
3 Male	20.18	0.0002	-6.4*10'8	-9.22	-12.73	-18.04	.18	83	4.62	nd
		(0.207)	(-1.227)	(-2.058)	(-2.048)	(-2.226)				
Second. 4 Total	12.00	-0.0005	-4.2*10-8	13.40	24.05	25.88	.38	86	11.79	-1.8
		(-0.949)	(-1.235)	(4.657)	(5.668)	(4.888)				
5 Female	9.26	-0.0005	-3.8*10-8	18.16	34.02	32.15	.52	79	18.01	2.2
		(-0.890)	(-1.06)	(5.79)	(6.96)	(5.77)				
6 Male	12.50	-0.0005	-4.6*10-8	12.86	24.24	21.73	.32	79	8.50	-7.7
		(-0.925)	(-1.325)	(4.238)	(5.127)	(4.037)				
Tertiary 7 Total	1.64	0.00064	8*10-10	8.70	12.44	12.23	.45	78	13.64	-2.1
		(0.776)	(0.012)	(4.463)	(5.029)	(3.286)				
Years of School 8 Total	1.99	0.0001	-1.7*10-8	0.96	1.42	0.85	.10	71	2.56	-1.5
		(0.464)	(-0.764)	(2.46)	(2.49)	(0.85)				
9 Female	1.98	0.0003	-3.09*10-8	1.40	1.50	0.65	.24	63	3.76	nd
		(1.00)	(-1.28)	(3.29)	(2.34)	(0.60)				
10 Male	1.80	0.0001	-1.1*10-8	0.83	1.37	0.92	.04	63	1.51	nd
		(0.20)	(-0.44)	(1.83)	(2.03)	(0.80)				

Data source: World Bank (1990). Enrollment rates (and therefore residuals for enrollment rates) are in percentages. Years of school (i.e., expected years of school for a synthetic cohort as defined in note 5) are in years (and therefore so are the residuals for these regressions). "nd" means that the date are not available for Brazil. The per capita GNP variables are in the form of differences between 1965 and 1987.

References

Behrman, Jere R., 1987, "Schooling in Developing Countries: Which Countries are the Under- and Overachievers and What is the Schooling Impact?" Economics of Education Review 6:2, 111-128.

Behrman, Jere R., 1990a, Human Resource Led Development? New Delhi, India: ARTEP/ILO.

Behrman, Jere R., 1990b, The Action of Human Resources and Poverty on One Another: What We Have Yet to Learn, Washington, D.C.: Population and Human Resources Department, World Bank.

Behrman, Jere R., 1990c, "Women's Schooling and Nonmarket Productivity: A Survey and A Reappraisal," Philadelphia: University of Pennsylvania, mimeo (prepared for the Women in Development Division of the Population and Human Resources Department of the World Bank).

Colclough, C., 1982, "The Impact of Primary Schooling on Economic Development: A Review of the Evidence," World Development 10, 167-185.

Eisemon, Thomas Owen, 1988, "The Consequences of Schooling: A Review of Research on the Outcomes of Primary Schooling in Developing Countries," Cambridge, MA: Harvard University, mimeo.

Haddad, Wadi D., Martin Carnoy, Rosemary Rinaldi, and Omporn Regel, 1990, Education and Development: Evidence for New Priorities. Washington, DC: World Bank Discussion Paper 95.

King, Elizabeth M., 1990, Educating Girls and Women: Investing in Development, Washington, DC: World Bank.

King, Elizabeth M. and M.Anne Hill, eds, 1991. Women's Education in Developing Countries, Washington, D.C.: World Bank.

Psacharopoulos, G., 1985, "Returns to Education: A Further International Update and Implications," Journal of Human Resources 20, 583-597.

Psacharopoulos, G., 1988, "Education and Development: A Review," The World Bank Research Observer 3:1 (January), 99-116.

Schultz, T. Paul, 1987, "School Expenditures and Enrollments, 1960-1980: The Effects of Income, Prices and Population Growth," in D. Gale Johnson and Ronald D. Lee eds., Population

Growth and Economic Development: Issues and Evidence, Madison, WI: University of Wisconsin Press, 413-476.

Schultz, T. Paul, 1988, "Education Investments and Returns," in Hollis Chenery and T.N. Srinivasan eds., Handbook of Development Economics, Amsterdam: North-Holland Publishing Company, 543-630.

Schultz, T. Paul, 1991, "Returns to Women's Education," in E.M. King and M.A. Hill, eds., Women's Education in Developing Countries, Washington, D.C.: World Bank.

World Bank, 1980, World Development Report, 1980, Washington: World Bank.

World Bank, 1981, World Development Report, 1981, Washington: World Bank.

World Bank, 1988, "The World Bank's Support for the Alleviation of Poverty," Washington, D.C.: The World Bank.

World Bank, 1990, World Development Report: Poverty, Oxford: Oxford University Press for the World Bank.

World Bank, 1991, World Development Report, 1991, Oxford: Oxford University Press for the World Bank.

A PÓS-GRADUAÇÃO NO BRASIL: **uma análise do período 1970-90**

Ricardo C. de Rezende Martins

A atividade de pesquisa no Brasil, a literatura especializada tem contemplado inúmeros aspectos do sistema educacional no País , nos diferentes graus de ensino. É volumosa a bibliografia disponível versando sobre questões fundamentais como as relações entre educação , escola e sociedade; os determinantes e condicionantes socio-político - econômicos da educação; a dimensão filosófica da educação; questões mais específicas de métodos e técnicas de ensino; avaliação de aprendizagem; custos e financiamento da educação.

Dentro deste quadro de investigações, porém, o segmento referente à pós-graduação, talvez por sua recente implantação - cerca de 20 anos - tem sido pouco estudado, configurando uma lacuna que urge ser preenchida, no que diz respeito às políticas do Estado para elas voltadas, sua função social, suas condições de funcionamento, sua organização, sua produção, seu financiamento, sua contribuição para o desenvolvimento econômico e social.

São ainda bastante raros os estudos que, de modo consistente e sistemático, procuram analisar a política nacional de pós-graduação, o movimento do Estado visando a implantação do sistema de cursos de mestrado e doutorado, as motivações que presidiram as diferentes determinações governamentais, as articulações deste movimento com a dinâmica interna das instituições universitárias, isto tudo considerado no contexto concreto do movimento socio-econômico-político nacional.

Se, por um lado, são bem conhecidas as características básicas do projeto de pós-graduação desenhado para o País através dos diferentes documentos elaborados pelo Conselho Federal de Educação, resta ainda desvendar de modo vertical as determinações de tal projeto, suas condições de implantação e os embates aí presentes. Pouco se elaborou ainda, por exemplo, sobre os conflitos decorrentes da distância das características do novo projeto em relação à estrutura de pós-graduação que se desenvolvia autonomamente em algumas instituições, dentre as quais destaca-se a Universidade de São Paulo.

O desenvolvimento do sistema nacional de programas de mestrado e doutorado tem seu ponto de partida de aceleração em fins da década dos anos 60, dentro de um quadro que, moldado pelo movimento revolucionário de 1964, estabeleceu uma nova ordem política que, por ações rápidas e efetivas, pretendia apresentar resultados que justificassem sua implantação, particularmente no domínio econômico.

Não constitui novidade afirmar que foi então escolhido um caminho que propiciasse o crescimento econômico acelerado, baseado numa política de investimentos financiados parcialmente por capitais nacionais e fortemente por recursos externos, aos quais foram oferecidas vantajosas condições de retorno.

A esta política concreta aberta ao exterior, corresponderam um discurso e procedimentos de planejamento cujo tom era bastante nacionalista. Neste sentido, pode-se compreender a política de formação de recursos humanos de alto nível. Se a política econômica enfatizava a necessidade de recursos externos para investimentos básicos no País, oferecendo inclusive grandes facilidades para importação de tecnologia, a política de formação de recursos humanos obedecia a duas linhas fundamentais: a necessidade futura de mão-de-obra especializada para preencher os novos empregos criados pelo desenvolvimento econômico previsto e a necessidade de cientistas, pesquisadores e técnicos aptos a desenvolver a pesquisa indispensável para a mudança, ao longo dos anos, do eixo de origem e de sustentação do desenvolvimento, do exterior em direção ao próprio País.

Estas duas linhas de atuação, aliadas a uma forte inspiração na teoria do capital humano, estão sem dúvida na base da formulação da política que resultou na criação do sistema nacional de pós-graduação e no reforço de certos órgãos oficiais voltados para assegurar seu funcionamento.

Assim, a reconstrução do contexto de gênese do sistema nacional de pós-graduação; a ecologia deste sistema, entendida como o conjunto de determinantes e condicionantes do surgimento dos cursos nas diversas instituições; a organização do Estado visando a implantação de mecanismos de formação de recursos humanos de alto nível e centros produtores de pesquisa científica e tecnológica; os objetivos subjacentes a esta ação deliberada do Estado; enfim, a reconstrução deste quadro constitui tema que ainda desafia uma pesquisa de porte.

A análise sistemática dos Planos Nacionais de Pós-Graduação, descortinando concepções e posturas neles presentes, o contraste entre seus objetivos e metas com resultados obtidos, as condições de sua implementação, bem como sua articulação com a política nacional de Educação e Ciência e Tecnologia e estas, dentro do contexto mais amplo da política econômica e social, constitui domínio pouco explorado.

A atuação dos diferentes organismos do Estado voltados direta ou indiretamente para a ordenação e financiamento da pós-graduação, carece também de análise mais detida. A criação, desenvolvimento, estruturação e, sobretudo, a ação de órgãos como a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Finaciadora de Estudos e Projetos (Finep), a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) e vários outros, neste quadro, merecem atenção especial.

A articulação do sistema de pós-graduação com o de graduação, com o setor produtivo (público e privado) e outros segmentos da sociedade também constitui objeto de análise necessária, tanto sob o ponto de vista de mercado empregador dos egressos dos programas de mestrado e doutorado, como sob a ótica de fontes de financiamento de produção acadêmica, científica e técnica.

A relação entre o Estado e a chamada comunidade científica, estabelecida no âmbito de tais organismos, através de mecanismos de participação, tais como as Comissões de Consultores Científicos e Comitês Assessores, por si só configuraria objeto de estudo específico. Nesta direção, torna-se importante analisar um dos grandes paradoxos do quadro da pós-graduação no País: se, por um lado, o conjunto de professores e pesquisadores dispõe de um considerável espaço de participação nos órgãos de governo que definem políticas e apoiam a pós-graduação, por outro lado esta se ressente de uma institucionalização mais efetiva nas universidades e escolas que oferecem cursos de mestrado e doutorado.

Este breve levantamento de questões, certamente não exaustivo, parece demonstrar que os aspectos referentes ao sistema nacional de pós-graduação, em si mesmo, como integrante do sistema universitário e do sistema nacional de Ciência e Tecnologia, permanecem, em boa medida, requerendo análise mais detida.

O Brasil vem despendendo um volume substancial de recursos no desenvolvimento e aperfeiçoamento do sistema de cursos de mestrado e doutorado. Torna-se, pois, imperiosa a realização de um balanço dos resultados obtidos com tais investimentos em formação de recursos humanos, seja no que tange aos processos - os programas em funcionamento - como aos seus produtos - egressos titulados, sua contribuição e a produção técnico-científica.

Todas estas questões constituem, na realidade, uma linha de pesquisa a ser desenvolvida durante um vasto período de tempo, configurando talvez uma base-diretriz para estudos de acompanhamento e avaliação de políticas voltadas para a pós-graduação.

O presente estudo certamente não pode abarcar todas as indagações levantadas, mas volta-se para uma análise diagnóstica de alguns pontos selecionados sobre o sistema de programas de mestrado e doutorado, buscando oferecer algumas respostas mais imediatamente necessárias, de modo a permitir um primeiro balanço deste amplo sistema em operação.

A política de pós-graduação: uma rápida visão através dos planos

O I Plano Nacional de Pós-Graduação (I PNPG) teve vigência no período 1975-79 e apresentou fortemente marcado pelo caráter da política educacional de então, que visava um reforço substantivo do sistema universitário, a fim de propiciar a formação de recursos humanos qualificados, tidos como necessários ao acelerado processo de desenvolvimento econômico, que constituiu o espírito da década brasileira dos anos 70.

Apresentando metas quantitativas precisamente delimitadas , o I PNPG pode ser ainda situado dentro do quadro de planejamento global que aproveitava importantes conceitos do "manpower approach", embora esta abordagem já estivesse sendo progressivamente abandonada nos planos e programas nacionais. De todo modo, a implementação do I PNPG deixou marcas extremamente fortes e positivas no sistema educativo brasileiro, podendo ser mencionadas duas ou três de suas importantes decorrências. Em primeiro lugar, cabe destacar a implantação do Programa Institucional de Capacitação de Docentes (PICD) que,desde 1976, visa, através da concessão de bolsas de estudos, propiciar aos docentes universitários a possibilidade de realizar cursos de mestrado ou doutorado em bons centros de pós-graduação no País. A necessidade de identificar estes centros de qualidade, aliada ao imperativo de acompanhar, regular e coordenar a expansão da pós-graduação, levou à criação de um sofisticado sistema de avaliação dos cursos, que constitui experiência original desenvolvida pelo Brasil e cujo sucesso pode ser atestado pela sua permanência e evolução até os dias de hoje, respaldado pelo reconhecimento da comunidade científica nacional.

A vigência do I PNPG coincidiu com um período em que os recursos governamentais para o sistema universitário foram bastante abundantes. Assim, a segunda metade dos anos 70 corresponde a uma significativa expansão do número de cursos de mestrado e doutorado, acoplada a uma expressiva implantação de laboratórios e centros de pesquisa.

O surgimento de tais cursos, porém, em boa medida se fez de modo exógeno ao conjunto das instituições em que eram implantados, dado que os recursos necessários ao seu desenvolvimento provinham de órgãos governamentais não diretamente comprometidos com o orçamento básico das universidades. Deste forma, os centros de pós-graduação poderiam ser tidos como "ilhas" dentro das instituições, formados por docentes e pesquisadores com sofisticada formação no exterior ou no País, cujas atividades, exceção feita aos salários, eram financiadas por órgãos externos.

Esta situação, facilmente administrável em tempos de abundância de recursos, passa a representar preocupação relevante quando a década dos anos 80 mostra sinais de retração econômica e contenção orçamentária. Este novo momento corresponde ao período do II Plano Nacional de Pós-Graduação (II PNPG), que abandona o discurso expansionista do primeiro plano e passa a articular duas diretrizes básicas : a primeira, voltada para a consolidação do sistema já implantado, através do reforço nos mecanismos de acompanhamento e avaliação, visando a melhoria da qualidade dos programas e a racionalização dos investimentos no setor.

A segunda diretriz voltou-se para a institucionalização da pós-graduação, com claro reconhecimento do caráter bastante exógeno com que se deu sua implantação. Neste sentido, o surgimento de um programa de apoio à infra-estrutura dos cursos de pós-graduação, com cunho orçamentário, no âmbito da CAPES, visava a garantia de recursos mínimos institucionais para subsistência e desenvolvimento dos cursos. A concepção original deste programa, iniciado em 1981/82, previa uma etapa inicial em que os recursos seriam diretamente distribuídos pela CAPES, de acordo com os dados da avaliação por ela promovida. Ao longo do tempo e na medida em que o volume de recursos fosse se tornando mais significativo - estimava-se um crescimento real de 10 a 15% ao ano - o programa de apoio seria paulatinamente incorporado ao orçamento das universidades, viabilizando, desta forma, a institucionalização da manutenção dos programas de pós-graduação, de acordo com critérios de qualidade.

As dificuldades orçamentárias e financeiras dos anos 80 encontram-se bem espelhadas na drástica redução sofrida por este programa que, se em 1984 chegou a representar cerca de 34% do orçamento global da CAPES, em 1990 atingiu um ponto mínimo de 5%, menos de 6 milhões de dólares. Esta reduçãao inviabilizou a idéia inicial de repassar ao orçamento das universidades as dotações do programa.

Dentre as muitas explicações para esta redução, que constitui apenas um exemplo dentre vários, encontram-se argumentos que sugerem a existência de inúmeras outras fontes de financiamento para a pós graduação, não se justificando, assim, que o Ministério da Educação, além das dotações orçamentárias normais para as universidades, ainda acrescentasse este programa mantido pela CAPES. Por outras fontes de recursos, citam-se, por exemplo, o CNPq e a Finep, órgãos financiadores de projetos de pesquisa, em termos de custeio e capital. Esta discussão reflete, em boa medida, uma larga incompreensão sobre o papel de cada uma destas agências e de seus respectivos programas.

O III Plano Nacional de Pós-Graduação (III PNPG) retrata de modo claro este estado de coisas, alertando enfaticamente para a necessidade de articulação cuidadosa das ações voltadas para a pós-graduação, oriundas de diferentes setores governamentais e privados. O novo plano, além de, pela primeira vez, indicar a necessidade de contemplar a relação entre universidade(pós-graduação) e setor produtivo, assinalou ser imperativo integrar os agentes do sistema de pós-graduação e do sistema de ciência e tecnologia. Com zonas de interseção bem evidentes, a articulação entre ambos, particularmente entre órgãos definidores de políticas e de investimentos, era bastante tênue, gerando por diversas vezes ações concorrentes, quando não contraditórias. Com efeito, os Planos Básicos de Desenvolvimento Científico e Tecnológico e os Planos Nacionais de Pós-Graduação jamais forma elaborados de modo conjunto ou sequer coordenado. As convergências observadas se deram menos por relações institucionais do que por articulações eventuais.

O III PNPG apresenta traços de contemporaneidade bastante interessantes, por vezes antecipando questões que hoje, nos anos 90, encontram-se decididamente na ordem do dia das discussões no meio universitário brasileiro e órgãos de governo. A preocupação com a articulação com o setor produtivo, seja como fonte de recursos adicionais (questão especialmente relevante em tempos em que os recursos púublicos se tornam mais e mais escassos), seja como aplicação de pesquisas e mobilização de estudos aplicados, caracteriza exemplo marcante.

Outra questão importante é a sinalização no sentido de um repensar do perfil da pós-graduação brasileira. Implantada de acordo com moldes norte-americanos, sua evolução denota um distanciamento progressivo do modelo inicialmente adotado, em particular no que tange ao mestrado que, no Brasil, guarda muito pouca semelhança com o "master" americano.

O caminho da pós-graduação brasileira apresenta dois traços bem nítidos: de um lado, um alongamento na duração dos programas de mestrado e doutorado; de outro, uma clara desvalorização da chamada vertente "lato sensu", constituída por cursos de especialização e aperfeiçoamento considerados, ainda que involuntariamente, cursos menos nobres e pouco valorizados no cartorial "mercado de títulos e certificados acadêmicos" do País. Um claro efeito deste quadro pode ser observado, como o III PNPG indicava, no fato de que um contingente significativo da demanda dos cursos de mestrado, especialmente em áreas profissionais, como as Engenharias, não busca o título de mestre, mas uma especialização através da obtenção de créditos em disciplinas, deixando-se de lado a elaboração de dissertação ou trabalho terminal, que conduziria à titulação final.

Partindo de um diagnóstico correto, já antevisto pelo II PNPG, o terceiro plano propugnava a valorização da pós-grsduação "lato sensu" como instrumento de formação de recursos humanos em áreas específicas, especialmente aquelas ligadas ao setor produtivo ou ao mercado profissional não-acadêmico. Esta é uma questão bem atual, mais até que à época do III PNPG, que não produziu concretamente nenhuma iniciativa de porte nesta direção. Este é pois um segmento da pós-graduação que se ressente, até o presente, de uma política consistente que favoreça o seu desenvolvimento.

É possível, portanto, afirmar, com as ressalvas apresentadas, que, caso raro na história da educação brasileira, os planos nacionais de pós-graduação constituiram de fato instrumento de política, isto é, as ações de governo guardaram suficiente coerência com os objetivos e metas declarados nos planos. É um dos poucos exemplos em que o planejamento, além de significar uma síntese concreta de um projeto de governo, apresentou conseqüências palpáveis na realidade do sistema de ensino brasileiro. As tradicionais contradições apontadas entre o discurso do planejamento educacional e a realidade do sistema de ensino, não se observam com respeito à pós-graduação.

As oportunidades para estudos de pós-graduação, ao menos em relação à demanda, são hoje bastante amplas, embora ainda longe de configurar uma infra-estrutura de formação de cientistas e pesquisadores compatível com a de países do hemisfério norte.Em 1990 , o Brasil contava com cerca de 52.000 estudantes de pós-graduação , dos quais mais de 80% no nível de mestrado. A relação candidatos / vagas situava-se dentro de limites bastante aceitáveis, variando entre 1 e 2,7 , favorecendo uma adequada seleção acadêmica, sem configuar qualquer tipo aparente de gargalo.

É importante notar que o suporte oferecido aos estudantes não é desprezível. Tomando-se o orçamento aprovado para 1991, e abstraindo-se os contingenciamentos determinados pelo Governo Federal desde fevereiro deste ano, e mantidos os atuais valores das bolsas, seria possível, em tese, contemplar quase 60% dos alunos com bolsas de estudos no País. Com relação à formação de recursos humanos no exterior, as oportunidades são também bastante amplas e acessíveis.

Além do valor das bolsas situar-se dentre os mais elevados do mundo (cerca de US$ 1.200 mensais, em média, mais pagamento de taxas escolares, seguro-saúde, passagens para o bolsista), o número de bolsas oferecidas a cada ano contempla, em geral, toda a demanda qualificada, selecionada de um volume de candidaturas que situa a relação candidatos/bolsas disponíveis num patamar inferior a 3.

Destes números, contudo, não se deve extrair uma precipitada conclusão de que o País, em relação à formação de recursos humanos de alto nível, se encontra em situação privilegiada. O Brasil está longe de alcançar um patamar ao menos aceitável, em termos de volume e velocidade de formação de cientistas, pesquisadores e demais profissionais de elevado nível. Comparando-se com outros países, como EUA, França e Japão, o Brasil apresenta um deficit de cerca de 400.000 mestres e doutores, em relação à sua densidade populacional.

A necessidade de continuar a investir em pós-graduação, ciência e tecnologia é inquestionável, ainda que suas atividades sejam bastante caras. A natureza dos investimentos, porém , sua diversidade e rentabilidade são questões que talvez mereçam hoje uma reavaliação que possa contribuir´ara uma eventual reorientação na alocação dos recursos. Para isso, cabe uma análise do desempenho da pós-graduação brasileira, iniciando pelo perfil do modelo efetivamente praticado na realidade das instituições universitárias e de pesquisa.

O "modelo" brasileiro de pós-graduação

Até os anos 60, o desenvolvimento da pós-graduação brasileira, além de bastante modesto, fez-se de forma quase expontânea , segundo tendências ou padrões endógenos às próprias instituições universitárias. A partir de 1964/65, sente-se a ação do Governo Federal no sentido de implantar um determinado modelo de pós-graduação, que vai buscar inspiração nos moldes norte-americanos. O primeiro parecer do Conselho Federal de Educação, o de nº 977/65, é bem claro neste sentido.

A adoção deste modelo certamente não se fez sem problemas, particularmente nas instituições que já haviam desenvolvido modelos outros, como, por exemplo, o da Universidade de São Paulo, sob nítida orientação francesa que, por sinal, encontra-se nas origens da própria universidade. No período 1970/75 praticamente todos os processos de solicitação de credenciamento de cursos da USP junto ao Conselho Federal de Educação foram objeto de diligências e diferentes polëmicas.

Durante longo período, as normas estabelecidas pelo CFE, o mecanismo de credenciamento dos cursos inegavelmente moldaram, ainda que por vezes, de modo apenas formal, como no caso da USP, o perfil dos programas de pós-graduação, inclusive de forma indistinta para as diversas áreas do conhecimento, em requisitos tais como qualificação do corpo docente, produção científica, tradição de ensino e pesquisa do grupo, disponibilidade de recursos materiais adequados (instalações, equipamentos e biblioteca), coerência de organização e regime didático-científico, estrutura curricular, áreas de concentração e domínio conexo, etc.

O espírito da legislação, em princípio, era bem flexível, dando aos programas ampla liberdade de organização acadêmica e curricular.Esta, aliás, é uma característica até hoje presente na legislação brasileira, tornada ainda mais flexível nas normas adotadas a partir de 1983.

Não obstante tal espírito, a implementação de seus princípios nem sempre a ele, correspondeu. De um lado, as interpretações dos órgãos de governo freqüentemente apegaram-se mais a questões de ordem formal do que substantiva, tendentes a estipular determinados moldes de organização e funcionamento da pós-graduação. De outro lado, a estrutura burocratizada e normativa das próprias universidades, colocando inúmeras exigências organizacionais e acadêmicas, transformando o que deveria ser leve em pesadas estruturas acadêmicas; o que deveria ser diversificado em estruturas curriculares obrigatórias; o que deveria ser objetivo e criativo, em insípidas corridas pela creditação acadêmica. A pós-graduação passou, portanto, por um longo período de enrijecimento de sua estrutura, resultando, por exemplo, em mestrados intermináveis, com 5 ou mais anos de duração.

Este estado de coisas vem sendo hoje questionado por diferentes segmentos ligados à pós-graduação.

Chega-se a propor um "novo" sistema de pós-graduação, norteado por uma flexibilidade total no que diz respeito ao número de créditos em disciplinas que, a juízo do orientador e do próprio estudante, seriam aqueles que contribuíssem substancialmente para seu trabalho de pesquisa.

As disciplinas a que os alunos, de acordo com esta sistemática, deveriám cursar deixariam de ser privativas de determinados professores ou departamentos, abrindo-se uma flexível composição curricular para cada aluno, que poderia buscar seus cursos em diferentes segmentos de uma dada universidade ou mesmo em outras universidades.

As figuras do orientador de tese e de programa, hoje existentes em muitos programas, se fundiriam numa única - a do primeiro - cuja disponibilidade e aceitação dos candidatos constituiriam os mecanismos básicos de seleção , a qual consideraria critérios de motivação e qualidade para a pesquisa, não enaltecendo as provas de conhecimento.

Isto permitiria modificar a atual configuração da pós-graduação brasileira rígida academicamente, caracterizando-se, pelo menos no nível de mestrado, como aprofundamento de conteúdo de graduação, o que inibe a atuação dos programas como geradores de pesquisas.

A rigidez curricular acaba se constituindo num problema para as instituições e para os estudantes. O fato de que os cursos muitas vezes não conseguem oferecer de modo regular seus respectivos elencos de disciplinas, leva a que os estudantes partam na busca, por vezes indiscriminada, de disciplinas com a única finalidade de completar créditos. Esta corrida às disciplinas subtrai aos estudantes importantes momentos de seus programas de pós-graduação, nos quais já poderiam estar desde logo e sempre voltados para seus trabalhos de dissertação ou tese.

Finalmente, no próprio processo de elaboração de tese, são vistos problemas extremamente importantes. Desde um certo dirigismo dos orientadores na escolha dos temas e dos resultados esperados, como uma prática paternalista na elaboração do trabalho terminal,no qual o orientador ou co-orientador tanto intervém que até o estilo destes é incorporado à redação final. Além disso, o processo de aprovação do trabalho até a chegada à defesa pública, com a realização de diversos seminários e revisões críticas de professores, é garantia quase certa de aprovação da tese.

Assim, como afirma Beiguelman, "para a elaboração de uma tese em nossos atuais cursos de pós-graduação, a inteligência criativa é menos exigida do que a dedicação ao trabalho, a perseverança, a disciplina, a paciência, a docilidade e uma certa dose de inteligência receptiva a um tema específico". Este esquema de funcionamento, portanto, leva a uma ineficácia dos cursos de pós-graduação, tornando menos provável o aparecimento de trabalhos de pesquisa que se destaquem por sua criatividade.

Trabalhando-se com dados relativos à idade média dos estudantes da pós-graduação, tempos médios de titulação e taxas de conclusão,reforçam-se as preocupações com o longo prazo necessário à formação de pesquisadores na pós-graduação brasileira e como esta demora se traduz em prejuízo para a criatividade da Ciência no País.

São estes alguns dos principais argumentos para a proposição de um "novo" modelo de pós-graduação, cuja viabilidade, contudo, segundo alguns, depende da superação de obstáculos internos e externos à universidade. Dentre estes últimos, arrola-se,por exemplo, o mecanismo de credenciamento pelo Conselho Federal de Educação,que fere a autonomia das universidades. Os mecanismos de credenciamento pelo CFE induzem ao formalismo e à rigidez da pós-graduação , quando exigem que se demonstre uma infraestrutura acadêmica típica de cursos de aprofundamento de conteúdo da graduação (currículo estruturado sob a forma de cursos, espaço físico e corpo docente bem especificado) e demonstração de recursos financeiros.

São legítimas várias destas preocupações quanto à formação de pesquisadores brasileiros. O diagnóstico das causas dos problemas levantados, porém, em particular quanto à natureza dos obstáculos(externos ou internos) merece uma análise mais detida.

De início, cabe examinar cômo se encontra apresentada, na legislação específica, a proposta de pós-graduação consagrada pelo Conselho Federal de Educação. Tome-se, pois, como referência a Resolução nº 5, de 10 de março de 1983, que fixa normas de funcionamento e de credenciamento dos cursos de pós-graduação "stricto sensu".

No artigo 2º, inciso IV, lê-se: "Além das atividades didáticas e acadêmicas, exigir-se-á do candidato ao grau de mestre a apresentação de dissertação du de outro tipo de trabalho terminal compatível com as características da área do conhecimento".

No parágrafo 4º desse mesmo artigo, encontra-se:"O doutorado será organizado em forma de programas de trabalho, com o fim de proporcionar formação científica ou cultural ampla e aprofundada, desenvolvendo a capacidade de pesquisa e o poder criador nos diferentes ramos do saber".

Do artigo 6º., que trata das informações necessárias ao pedido de credenciamento, cabe destacar três incisos:

"III - Relação dos docentes resposáveis pela orientação de dissertações, teses ou trabalhos equivalentes, cuja qualificação será comprovada pela formação acadêmica, com a titulação correspondente, e pela produção científica ou atividade criadora, devendo ser explicitadas as linhas de pesqúisa em que atua cada orientador.

IV - Experiência de pesquisa do grupo, demonstrada mediante a descrição da atividade criadora específica dos membros do corpo docente e a produção de trabalhos originais.

V - Estrutura curricular do curso, docentes responsáveis e caráter obrigatório ou optativo das disciplinas que são ministradas."

A leitura destes excertos da Resolução traz à tona algumas informações importantes. Em primeiro lugar, pode-se retirar da discussão o nível de doutorado. Como citado, o doutorado deve se organizar sob a forma de programas de trabalho, o que desde logo sugere que não deva ou ao menos não precise ele ser estruturado sob a forma de cursos, disciplinas e outras atividades de natureza semelhante. Embora se tenha notícia de que diversos programas de doutorado se estruturam com uma considerável carga de créditos em disciplinas, não se pode pois afirmar que se trata de uma exigência da legislação.

Os requisitos referentes ao mestrado, porém, sugerem efetivamente que ele deva receber uma estruturação mais acadêmica, sob a forma de cursos, que permitam aprofundar a formação dos candidatos ao título. Isto, porém, não significa que o estudante deva ser assoberbado de créditos a serem cursados, mas que a ele seja fornecida a necessária base de conhecimentos que permita sua iniciação ou aprimoramento no desenvolvimento da pesquisa ou da atividade profissional a que ele venha se dedicar. Quanto a este ponto, deve ser lembrado que, se para todas as áreas a pesquisa científica é elemento central inquestionável no nível de doutorado, isto não é questão pacífica no nível de mestrado. Há áreas, como as Engenharias, por exemplo, em que se discute a conveniência de um mestrado por vezes chamado "profissional", conhecido como o mestrado sem tese. Em outras áreas, como a Medicina, encontram-se defensores de que a vocação do mestrado deva ser a formação do professor de nível superior, com ênfase na formação didático-acadêmica.

Desta forma, abordar indiscriminadamente a vocação da pós-graduação, sem contemplar a especificidade das diferentes áreas do conhecimento , parece ser uma generalização temerária. Por outro lado, há a questão da distinção entre os níveis de mestrado e doutorado. Se ambos devem apresentar configuração semelhante, onde residiria a distinção: no mestrado haveria uma prática "menor" de pesquisa e no doutorado uma prática "maior"? Quais seriam os limites do menor para o maior? Ou o mestrado seria uma espécie de "prêmio de consolação" para aqueles que não merecessem alcançar o título de doutor? Quais seriam os critérios para tanto?

Se em certas áreas não se justificam propostas diferenciadas de pós-graduação por níveis, talvez o recomendável fosse eliminar um deles- o mestrado - passando a oferecer apenas o doutorado. Isto, por sinal, é possível, pois a legislação atual não coloca o mestrado como pré-requisito para o doutorado. Do contrário, estar-se- ia correndo o risco de transformar mestrados em doutorados, sem dar a eles o "status" dos últimos. Em algumas áreas, chega-se a afirmar que o longo tempo médio de titulação se explica, em vários casos, pelo fato de que, no mestrado, as exigências feitas são, na realidade, exigências em nível de doutorado. Assim, ou se tem uma pós-graduação claramente vocacionada, com níveis e com finalidades distintas, ou abolem-se os níveis para os quais os limites não são definíveis. Não se pode, porém, querer manter uma estruturação de pós-graduação diferenciada sem características diferenciadoras.

Deve ficar claro que aqui não se postula uma posição de que os mestrados sejam "aulísticos", mas tão somente que, dadas as características de certas áreas do conhecimento, uma certa dose de estudos sob a forma de disciplinas organizadas de acordo com um currículo estabelecido parece ser aconselhável. Isto não significa que o currículo de mestrado deva ser extenso ou rígido. Não quer dizer que a estrutura curricular deva ser uma "camisa de força" ou que para um determinado programa de pós-graduação não se possam compor programas individuais de estudos para os alunos.

A tal ponto a legislação de 1983 tentou avançar em termos de flexibilidade da pós-graduação que, na Resolução nº 5, não se encontra mencionada a necessidade dos cursos se organizarem sob a forma de áreas de concentração/domínios conexos.

A legislação não menciona currículos mínimos, nem número mínimo de créditos e tampouco que o programa deva ser mantido por um determinado departamento ou segmento universitário.Fixa-se apenas que mestrado e doutorado devem ter duração mínima de um e dois anos, respectivamente . Quando se menciona a necessidade de comprovação de recursos materiais e financeiros, não se afirma que estes devem ser vinculados a determinados e exclusivos centros de custos da universidade, mas refere-se à existência de condições para manutenção do programa. Se, para o pedido de credenciamento de um programa de pós-graduação, é solicitada uma relação de docentes/pesquisadores que sustentarão as atividades do programa, o objetivo maior é identificar a existência da massa crítica indispensável para a formação pós-graduada na área em que a instituição se propõe atuar. Não é, pois, nesta direção que se pode inferir que as exigências do credenciamento levam ao atrelamento dos cursos de pós-graudação a determinadas unidades, impedindo iniciativas para as quais concorram diferentes segmentos das instituições de ensino superior.

A flexibilidade da legislação também pode ser observada em outros pontos. O Artigo 15 abre a possibilidade de convênios entre instituições para ministrar um mesmo curso de pós-graduação e, em seu parágrafo único, prevê até mesmo que um estudante possa realizar estudos fora da sede do curso, no País ou no exterior, desde que assegurada a existência de orientador individual qualificado.

Finalmente, quanto à questão da seleção de alunos, a Resolução nº 5 preconiza: "Art. 9º - A admissão de estudantes aos cursos de pós-graduação deverá estar condicionada à capacidade de orientação de cada curso, comprovada através da existência de orientadores com disponibilidade de tempo para esse fim".

Assim, ao ler a Resolução do CFE ou os documentos da Comissão cujos estudos a geraram, não se conseguiu perceber a indução legal à rigidez curricular ou a que os cursos de pós-graduação devam se constituir em compartimentos estanques, à semelhança de feudos medievais.

Não há dúvida de que a prática de pós-graduação nas instituições apresenta muitas das anomalias criticadas. E também é certo que, no passado, muitas delas foram induzidas pela prática de órgãos centrois do governo, e que seus efeitos ainda se fazem presentes.

A questão curricular nos programas de pós-graduação, porém, parece enfrentar hoje dilemas cujas raízes se encontram muito mais na própria prática das universidades do que nas normas ou na prática dos órgãos de governo. Assim, se mudanças são necessárias, elas podem ser desenvolvidas dentro das instituições de ensino superior, posto que a legislação é suficientemente flexível para abrigá-las.

Enfim, o mecanismo de credenciamento dos cursos de pós-graduação é algo cuja existência pode ser discutida, mas dentro de um quadro realista, colocando de modo preciso todos os aspectos envolvidos.

A prática da pós-graduação nas instituições é bastante elucidativa a este respeito. Embora haja praticamente consenso de que o modelo de pós-graduação brasileiro é voltado para a pesquisa, percebe-se que a prática dá ênfase acentuada às atividades de ensino, dentro de sua modalidade didática mais tradicional- a aula expositiva- exceção feita talvez às Ciências Biológicas, onde parece mais forte a tendência à atividade laboratorial. A atividade de pesquisa apresenta-se também em sua roupagem mais tradicional, marcada pelo individualismo das investigações. É bastante raro encontrar projetos coletivos de pesquisa, embora a larga maioria dos cursos declarem atuar em torno de linhas programáticas de pesquisa. Os conceitos de núcleos temáticos, linhas institucionais de pesquisa e outros assemelhados constituem realidade pouco difundida ou praticada.

Há , pois, uma contradição fundamental na estrutura dos cursos de pós-graduação, interna à dinâmica das instituições de ensino e pesquisa, Há uma estrutura de disciplinas obrigatória e coletiva para os alunos, que permite pouco direcionamento a projetos individuais de investigação. E, na etapa de elaboração da dissertação ou tese, requisito universal na pós-graduação brasileira, inclusive nos mestrados, passa-se a um trabalho específico de investigação, ligado a um indivíduo orientador, para o qual nem sempre os cursos seguidos terão fornecido embasamento suficiente ou sequer coerente. Ademais, ressalvadas algumas áreas experimentais, a iniciação à pesquisa raramente se faz pela integração do aluno em grupos de pesquisa, estágios ou atividades assemelhadas. É freqüente que a etapa de cursos e disciplinas seja algo à parte, sendo depois o aluno "jogado" em sua atividade de elaboração de tese, talvez após uma ou duas cadeiras de metodologia da pesquisa científica. Não seria demais afirmar que, em muitos casos, a prática da pós-graduação constitui uma refinada reprodução da minuciosamente delimitada organização dos cursos de graduação, no tocante ao ensino e uma " liberação individual" com respeito à atividades de pesquisa.

O SISTEMA BRASILEIRO DE PÓS-GRADUAÇÃO: uma visão geral

O Brasil conta hoje com mais de 1.000 programas de mestrado e 400 de doutorado, constituindo o maior sistema de pós-graduação existente no continente latino-americano. Trata-se, com certeza, do resultado de um significativo esforço de crescimento, haja vista que, até o início de 1970, contavam-se apenas 93 mestrados e 32 doutorados.

O período de expansão mais acelerada coincidiu com os anos 70, quando o sistema praticamente aumentou oito vezes de tamanho. O início da década de 80, quando se agravou a crise econômica, marcou uma retração na criação de novos cursos, que voltou a tomar impulso a partir de 1985, sobretudo em nível de doutorado. Hoje as taxas de expansão situam-se em torno de 3% para mestrado e 6% para doutorado.

Importa notar que nem sempre a expansão no número de cursos vem acompanhada de um proporcional aumento do alunado. Ao contrário, no nível de mestrado, o número médio de alunos por curso decresceu de 53,5, em 1982, para 45,6, em 1989/90. No doutorado, a relação não sofreu variação, situando-se em torno de 24 estudantes por programa. Se o doutorado, portanto, se expande de forma linear, o mestrado parece perder fôlego na captação de estudantes, o que se apresenta consistente com o longo período de perda no poder aquisitivo das bolsas(especialmente 1981 a 1986) , a falta de incentivo na carreira docente das universidades federais para obtenção deste título, a excessiva complexidade dos cursos, que os torna muito prolongados, constituindo desestímulo aos candidatos, sobretudo aqueles não vinculados à docência. O alunado de mestrado sistematicamente decresceu, no período de 1982 a 1986, de 40.690 para 37.825 matrículas, voltando a crescer a partir de então, situando-se hoje em torno de 42.000.

A distribuição dos programas segundo as áreas do conhecimento demonstra uma significativa constância no desenvolvimento do sistema. Comparando-se os anos de 1980 e 1990, a distribuição percentual dos cursos de mestrado pelas grandes áreas do saber praticamente permanece inalterada, observando-se maior concentração relativa na área de Ciências Humanas e Sociais(cerca de 30%), seguida das Profissões da Saúde (21,5%), Ciências Exatas e da Terra (entre 13 e 14%), configurando-se então um terceiro grupo (cada uma entre 10 e 12%), formado pelas áreas de Ciências Biológicas, Engenharias e Ciências Agrárias.Em nível de doutorado, o perfil se modifica, passando a uma maior importância relativa das Profissões da Saúde (26%), seguidas pelas Ciências Humanas e Sociais (22%), Ciências Exatas e da Terra(16%), Ciências Biológicas (14%), Engenharias (12%), Ciências Agrárias(8,2%).

Esta constância na distribuição percentual retrata o fato de que, de modo geral, a política nacional de pós-graduação, no período, não privilegiou nenhuma área do conhecimento em particular, na utilização de seus instrumentos de fomento.

Chama a atenção a maioria relativa dos programas de Ciências Humanas e Sociais e Profissões da Saúde na pós-graduação. Resultado de um desenvolvimento quase que espontâneo dos departamentos das instituições, entende-se a expressiva presença dos cursos de Ciências Humanas e Sociais, não só por serem tradicionalmente cursos menos caros, mas também pelo fato de que o mercado acadêmico é fundamental para os profissionais dessas áreas, sustentando, assim, uma demanda bastante estável pela titulação e uma natural propensão à criação de tais cursos. Ademais, já eram eles preponderantes na pós-graduação pré-existente à reforma do ensino superior encetada em 1968, caracterizando-se, portanto, uma tendência histórica não modificada.

A presença de um número importante de cursos ligados às Profissões da Saúde (Medicina e Odontologia, basicamente) também já se faz sentir antes dos anos 70: cerca de 20% dos cursos de então ligavam-se, por exemplo, as áreas de Higiene e Saúde Pública. Não se pode esquecer, inclusive, que uma das mais antigas tradições do ensino superior no Brasil encontra-se exatamente na área médica, com a criação, em 1808, das aulas de Anatomia e Cirurgia, das quais resultaram, duas décadas adiante, as Faculdades de Medicina da Bahia e do Rio de Janeiro. Além disso, a permanente especialização nas áreas médica e odontológica foi desdobrando, gradativamente, os antigos cursos gerais segundo as especialidades, gerando assim uma multiplicação de cursos de tamanho relativamente reduzido em termos de docentes e alunos, porém com focos bastantes delimitados.

Não há dúvida de que são áreas relevantes e que incumbe ao sistema de ensino superior delas cuidar. Mas certeza cabe indagar se, tendo em vista as necessidades de formação de uma base científica e tecnológica nacional, outras áreas não deveriam ter sido atendidas ou estimuladas com maior destaque. Neste sentido, embora a sinalização dos dados seja tênue, causa preocupação observar um decréscimo na participação das Ciências Exatas e da Terra (insignificante no mestrado mas de quase 4% nos doutorados) e o não aumento relativo da presença de cursos das áreas biológicas. Cabe, porém, observar que são exatamente estes dois conjuntos de áreas que atingiram maiores níveis de "dureza" científica e reconhecimento internacional. Verticalizaram-se qualitativamente, mas o potencial multiplicador encontra-se limitado pela modéstia de sua expansão quantitativa. Inegavelmente são áreas fundamentais para um salto em termos de Ciência e Tecnologia, mas são também aquelas que mais exigem em termos de investimentos em laboratórios, equipamentos e materiais. Explica-se, assim, em boa parte, a modéstia com que tem crescido o número de cursos nestas áreas em função da menor disponibilidade de recursos para investimentos no ensino superior. A necessidade de recuperação deste quadro parece já ter sido sentida, como atesta a existência do Programa de Apoio ao Desenvolvimento Científico e Tecnológico (PADCT), com recursos do Governo brasileiro e do Banco Mundial.

Deve ser ainda lembrado o expressivo crescimento, em termos absolutos e relativos, dos programas de doutorado em Ciências Agrárias, área em que se avolumam as contribuições da pesquisa científica à melhoria das atividades de cultivo e produção agrícola, observando-se inclusive, em vários centros, uma oportuna tendência de associação das universidades com os centros de pesqüisa da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA). Se, em 1980, havia apenas 13 doutorados na área, representando 4,7% do total de cursos, os números de 1990 são muito mais significativos: 33 doutorados, 8,2% do total.

A distribuição do alunado segundo as áreas do conhecimento desvela nitidamente as diferenças de tamanho e escala de operação dos cursos. Cerca da metade dos alunos encontram-se matriculados em programas de Ciência Humanas e Sociais. Esta desproporção pode ser analisada ao inverso. Por exemplo, se as Ciências Biológicas respondem por 11% dos mestrados e 14% dos doutorados brasileiros, o respectivo corpo discente representa apenas 6 ou 7% do total. Nas Profissões da Saúde, a distância dos números é ainda maior: mantendo 21% dos mestrados e 26% dos doutorados, seu alunado não chega a 12%.

Este perfil dos programas e do alunado de pós-graduação diferencia de modo significativo o Brasil de outros países industrializados, onde a participação do alunado das áreas de Exatas, Biológicas e Engenharias é maior, refletindo uma articulação estrutural entre pesquisa básica e aplicada e sua utilização nos diferentes setores produtivos nestes países.

Algumas evidências, no Brasil, sugerem uma tendência ao estabelecimento desta articulação, contando-se algumas experiências muito bem sucedidas. Como exemplos, podem ser citadas a área de Engenharia Mecânica em Santa Catarina, as Engenharias em geral no Rio de Janeiro e em São Paulo, a Física em São Paulo, a Metalurgia em Minas Gerais, a Biotecnologia no Rio de Janeiro e em São Paulo, entre outros.

A quase totalidade deste tipo de iniciativa encontra-se concentrada nos estados da Região Sudeste, onde precisamente se concentram os programas de pós-graduação brasileiros. Mais de 70% dos cursos estão aí situados, sendo ainda mais elevada a proporção dos doutorados, em torno de 90%. Este quadro, embora não seja surpreendente, dadas as diferenças regionais de desenvolvimento econômico e educacional, apresenta algumas dificuldades bastante importantes, afora a questão mais geral de desequilíbrio nas oportunidades de acesso à pós-graduação.

Há problemas de pesquisa suficientemente específicos, bem como uma série de sítios de investigação que não dispõem proximamente de centros de formação de recursos humanos e de pesquisa, como seria desejável. É o caso do Norte do País, com a vasta região amazônica, apresentando enormes desafios de pesquisa nos campos da biologia, aproveitamento de recursos hídricos, florestas, etc. Menos de 1% dos cursos de pós-graduação encontram-se aí localizados. Agrava a situação perceber que a pequena tendência de crescimento de programas nas universidades locais direciona-se para outras áreas que não aquelas mais estreitamente ligadas às peculiaridades locais. Recentemente, contudo, as discussões do Projeto Norte de pós-Graduação, congregando órgãos de governo (particularmente a CAPES e CNPq) e as universidades da região, tentam redirecionar este movimento, visando uma maior aderência entre ensino, pesquisa e realidade.

A questão de aproveitamento e recuperação de solos, irrigação e drenagem é fundamental para o Nordeste que, embora conte com um certo número de cursos de mestrado (11% do total de mestrados do País, em diversas áreas), não dispõe de um único doutorado nestas áreas.

Outro grande vazio da pós-graduação brasileira se encontra na região Centro-Oeste, onde retirando-se o Distrito Federal (leia-se Universi dade de Brasília), praticamente não existem programas de mestrado e doutorado).

A região Sul começa a tomar impulso a partir dos anos 80, mantendo hoje quase 15% dos cursos de pós-graduação.

A distribuição regional dos programas indica que as oportunidades de acesso à pós-graduação são extremamente concentradas. Isto apresenta algumas decorrências graves, na medida em que a concorrência por vagas nos melhores programas, normalmente situados na região Sudeste, conta com elevadíssima participação da população estudantil local que também em geral é a melhor preparada para a disputa de vagas. Assim, as universidades dos locais periféricos, necessitadas de qualificação de seu corpo docente, tem por vezes dificuldades em enviar seus docentes para tais cursos. Não é incomum que docentes de universidades do Nordeste e do Norte consigam aceitação, com mais facilidade, em universidades estrangeiras do que em instituições brasileiras. Maior rigor de seleção aqui do que lá? Interesse mercadológico das instituições estrangeiras? Estas e outras hipóteses talvez sejam plausíveis, mas não elidem o fato de que, muitas vezes, para firmar a competência de investigação e criação é preciso localizar os seus centros geradores no âmago dos locais onde se apresentam os problemas ou desafios de pesquisa, por razões de natureza acadêmica, científica e econômica.

A análise dos dados referente à vinculação institucional informa que quase 80% dos programas de pós-graduação são mantidos por universidades públicas, federais e estaduais, destacando-se dentre estas últimas as paulistas. Cerca de 10% dos cursos são oferecidos por escolas isoladas federais, 8% por universidades privadas e 3% por estabelecimentos isolados particulares. Quase toda a pós-graduação brasileira é assim mantida pelo Poder Público. Trata-se de um quadro que resulta da associação da histórica tendência do Estado brasileiro assumir o ônus dos segmentos educacionais de elevado custo e voltados para a formação sofisticada de recursos humanos, com a sua característica, da pós-graduação, de natural desdobramento das instituições de ensino superior de melhor qualidade que , no Brasil, também se identificam, em geral, com as escolas públicas de 3º grau. É um sistema gratuito moldado na tradição da história da educação bra sileira.

Por outro lado, deve ser lembrado que as origens do sistema nacional de pós-graduação permitem inferir que os investimentos no setor foram também justificados, dentro de uma visão das necessidades científicos e tecnológicos para o desenvolvimento econômico, como investimentos privativos do Estado, na medida em que visavam a melhoria da qualificação dos docentes das universidades por ele mantidas.

Mesmo no caso das instituições particulares, não é trivial o aporte de recursos públicos para manutenção de sua pós-graduação, em termos de bolsas de estudos, incluindo pagamento de taxas escolares, verbas para desenvolvimento de atividades acadêmicas, além de recursos para efetivo suporte da infra-estrutura de recursos humanos e materiais, como é o caso da Pontifícia Universidade Católica do Rio de janeiro (PUC/RJ).

A implantação, o desenvolvimento e o funcionamento do sistema brasileiro de pós-graduação são, pois, em sua quase totalidade, dependentes do aporte de recursos públicos. A pós-graduação, em termos de formulação de políticas e sua implementação, é considerada, para todos os efeitos, um investimento público, sendo absolutamente residual o volume de recursos captados junto a outros setores da sociedade.

Esta tradição de aporte de recursos públicos como elemento basilar na dinâmica da pós-graduação brasileira já estava presente nas origens do programa de formação de recursos humanos no exterior, desenvolvido pelo Brasil a partir dos anos 50, através da criação de agências governamentais como a CAPES e o CNPq, e impulsionado nas duas décadas seguintes. O conjunto de docentes formados no exterior, com bolsas do Governo brasileiro, constituiu, em boa medida, o núcleo básico central que permitiu a implantação de programas de mestrado e doutorado no País.

Até o início dos anos 80, significativa parcela, mais de 50%, dos docentes da pós-graduação tinham obtido seus títulos em universidades do exterior, em especial nos EUA, França e Inglaterra. Grande parcela dos restantes correspondem aos docentes da Universidade de São Paulo, que já havia desenvolvido programas próprios de titulação de seus docentes. A partir dos anos 80, como resultado do esforço nacional de pós-graduação, invertem-se as proporções, observando-se cerca de 50% dos docentes com titulação obtida no próprio País, majoritariamente na mesma instituição em que atuam. Embora esta espécie de "in breeding", cuja compensação hoje se busca realizar através de programas de pós-doutoramento no exterior, já esteja disseminada por uma larga parte do sistema nacional, ela com certeza se faz sentir com mais relevo nas instituições que congregam maior número de doutorados, isto é, as de São Paulo. De todo modo, verifica-se que um grande contingente dos mais de 20.000 docentes que atuam na pós-graduação (cerca de 44%) possuem títulos obtidos em outros países, o que significa uma saudável diversificação acadêmica e científica.

O conjunto dos cursos de pós-graduação parece, ao menos em termos quantitativos, bem atendido em termos de docentes. A relação média global aluno/docente situa-se em torno de 2,5, variando no intervalo de 1,1 nas áreas biológicas até 4 nas Ciências Humanas e Sociais. Este indicador situa-se em patamar bastante baixo, mesmo para países com longa tradição de pós-graduação.

O perfil deste corpo docente, contudo, ainda não atende integralmente às exigências de titulação mínima em nível de doutorado: cerca de 15% ainda não detêm este título, embora este contingente venha se reduzindo gradativamente. Em 1980, situava-se na casa dos 22%.

Quase 100% dos professores encontram-se vinculados às suas instituições em regime de dedicação exclusiva ou tempo integral. A pós-graduação brasileira reúne, assim, a quase totalidade da elite dos docentes do ensino superior brasileiro. Quase todos os doutores, que ainda constituem menos de 15% do corpo docente do sistema brasileiro de ensino superior parecem atuar na pós-graduação. Se, por um lado, este dado corresponde a uma situação esperada, cabe observar que, ao menos no caso brasileiro, existe uma nítida tendência a que, uma vez vinculado à pós-graduação, o docente reduza ou mesmo encerre suas atividades ligadas ao ensino de graduação. Menos de 30% dos docentes informam dedicar-se regularmente ao ensino de graduação. Deste modo, cria-se um círculo potencialmente perverso no qual, dentro do ensino superior, estabelecem-se ilhas de competência acadêmica e científica (a pós-graduação), em detrimento do segmento básico, que é o ensino de graduação. Este é, com efeito, um dos principais dilemas do ensino superior brasileiro sistematicamente apontado nos diversos planos nacionais de pós-graduação.

A produção da pós-graduação brasileira

O esforço de desenvolvimento da pós-graduação brasileira, sendo bastante recente, pode explicar algumas das deficiências no perfil de sua produção. Isto, contudo, não impede que se façam profundas reflexões sobre seu desempenho, em alguns itens bastante preocupante.

A titulação anual de mestres e doutores, embora tenha aumentado substancialmente em termos absolutos, nos últimos vinte anos, apresenta-se modesta em relação à dimensão do alunado em curso. Titulam-se, a cada ano, cerca de 4000 mestres e quase 1000 doutores no País, o que corresponde a cerca de 10% do corpo estudantil. Os dados são um pouco mais animadores quando se comparam as novas admissões com as titulações, situando-se estas últimas, ano a ano, em torno de 40% a 45% das primeiras. Se o fluxo anual parece equilibrado, quando se contrastam as admissões com as saídas, é preciso considerar que este equilíbrio é de certo modo perverso, pois entre os que saem, somam-se os que abandonam os cursos e os que trancam matrícula (evadidos em potencial). Em média, por ano, os que abandonam ou trancam matrícula, no mestrado, correspondem a quase o dobro dos que se titulam. No doutorado, a proporção é de um titulado para um evadido, real ou potencial.

Os dados levantados sugerem a existência de um elevado número de alunos que permanecem por longo tempo dentro dos programas. De fato, os tempos médios de titulação para mestrado estão bastante alongados, situando-se perto de 4 anos para mestrado e pouco mais de 5 anos para doutorado. Estes indicadores refletem uma realidade que importa modificar. Várias são as razões para este quadro : muitos estudantes procuram a pós-graduação mais como especialização que visando obter a titulação final, em especial no mestrado; um número não precisado de estudantes não se dedica aos estudos em tempo integral; excesso de exigências de inúmeros programas, em especial no que respeita aos créditos em disciplinas; deficiências na orientação de dissertações/teses. Tais questões foram, por sinal, abordadas com ênfase no III PNPG.

É verdade que observa-se uma significativa variação nos tempos médios de titulação, entre as diversas áreas do conhecimento. Assim, por exemplo, os programas de Ciências Agrárias, Engenharias e alguns ramos das Ciências Biológicas apresentam tempos perto dos 30 meses. Já nas Ciências Humanas e Sociais, encontram-se os tempos mais dilatados, por vezes superior a 60 meses, em nível de mestrado. Como os cursos desta última área são os relativamente preponderantes no conjunto, explica-se a elevação da média geral. De todo modo, há uma propensão ao excessivo prolongamento do mestrado que torna bastante demorada a trajetória de formação de novos cientistas e pesquisadores no País.

Este quadro sofre um agravamento quando se percebe que quase todos os programas de doutorado nacionais impõem como pré-requisito de admissão a titulação em nível de mestrado. São raros aqueles que admitem a passagem automática do mestrado para o doutorado, quando o desempenho do estudante assim o recomende. Esta prática, porém , apresenta-se contraditória quando se nota que um número importante de programas de pós-graduação, especialmente em São Paulo, apresenta estrutura curricular comum para os dois níveis. Prolonga-se, deste modo, para quase 10 anos o período médio para formação de pós-graduados plenos no País.

Adicionando o ingrediente da idade média dos candidatos à pós-graduação, forma-se um retrato preocupante. Dificilmente encontram-se estudantes de pós-graduação com idade inferior a 25 anos, situando-se a faixa média de idade para mestrandos perto dos 30 anos. A idade brasileira de formação de um doutor aproxima-se, pois, dos 40 anos, o que, sem dúvida, constitui uma fase tardia para obtenção desta titulação acadêmica.

É interessante comparar estes dados com o apoio recebido pelos estudantes de pós-graduação. Ao longo dos últimos 4 anos, a concessão agregada de bolsas de estudos, pela CAPES e CNPq, situou-se em torno de 20.000 bolsas anuais. Neste período, os índices de titulação de mestres e doutores variaram, anualmente, entre 3.500 a 4.000 formados. Há, portanto, no contraste destes números, uma sugestão de que o sistema não vem respondendo de forma ágil aos estímulos recebidos.

O aumento de eficiência na operação de um sistema complexo como o de pós-graduação constitui-se efetivamente num processo lento e sinuoso, que deve inclusive romper, no caso brasileiro, com largas tradições cristalizadas, como o mestrado-quase-doutorado.

Ao longo dos últimos anos observam-se alguns ganhos de produtividade no fluxo de operação dos cursos. Com efeito, o tempo médio de titulação no mestrado reduziu-se de 5 anos, em 1985, para 4 anos em 1989/90. O tempo do doutorado, contudo, manteve-se estável, em torno de 5 a 5 anos e meio.

A questão da produção científica/técnica e artística da pós graduação é outro ponto merecedor de discussão atenta. Os dados disponíveis sobre publicações do conjunto de docentes/pesquisadores ligados aos programas de mestrado e doutorado no Brasil são sugestivos.

Alguns dados resultantes do levantamento anual de acompanhamento da pós-graduação, realizado pela CAPES, vem apresentanto um perfil que merece consideração. Tomando-se uma lista de publicações que inclui desde livros publicados no País ou no exterior, artigos em periódicos nacionais ou estrangeiros até resumos publicados em anais de congressos, observa-se que, nos últimos anos, o volume da produção cresceu substancialmente. A taxa média de publicações **per capita** apresenta uma evolução bastante satisfatória, subindo de 0,6 em 1978 para 0,87 em 1982, 0,93 em 1985 e aproximando-se de 1,0 em 1989/90. O volume absoluto de trabalhos publicados passou de 8.000 em 1978 para cerca de 22.000 em 1989/90, o que representa um avanço considerável, num período de pouco mais de dez anos. Esta progressão situa o Brasil em posição privilegiada em relação aos demais países latino-americanos, argumento que se tornará mais claro ao ser examinada a composição desta produção científica.

Do total dos trabalhos publicados, cerca de 26% referem-se a artigos em periódicos nacionais, 10% a artigos em periódicos internacionais, 37% a trabalhos apresentados em congressos nacionais e 8% a trabalhos apresentados em eventos internacionais.

Tomando-se o conjunto de docentes, verifica-se que, já por longo tempo, cerca de um em cada três docentes/pesquisadores tem publicado anualmente um artigo em periódico nacional e que um em cada dez tem publicado um artigo em periódico estrangeiro. Os índices por grandes áreas do conhecimento são sugestivos. A relação total de artigos nacionais com total de docentes apresenta o seguinte perfil: 0,12 para Ciências Exatas, 0,36 para Ciências da Saúde, 0,44 para Ciências Biológicas e 0,49 para Ciências Humanas e Sociais. O com portamento da relação tomando-se o total de artigos em periódicos de circulação internacional é quase inverso: 0,05 para Ciências Humanas e Sociais, 0.07 para Ciências da Saúde, 0,32 para Ciências Exatas e 0,33 para Ciências Biológicas.

Estes dados sugerem um nítido perfil da produção das diferentes áreas, bem como o grau de internacionalização da ciência brasileira, que, por sua vez, reflete o seu nível de amadurecimento ou consolidação. Tomando-se como exemplo de ponta, observa-se que cerca de 70% da produção de artigos das Ciências Biológicas, notadamente as sub-áreas de Bioquímica, Fisiologia, Farmacologia, Imunologia e Biofísica, é publicada em periódicos de circulação internacional.

Ressalvada a maior contribuição proporcional de algumas áreas, cabe destacar que, desde 1978, o Brasil já se situava em primeira posição, dentre os países latino-americanos, em termos de artigos publicados em periódicos internacionais. Mais de 30% do total de tais artigos produzidos no continente são de autoria de pesquisadores brasileiros.

Esta marca atingida no período é alentadora. A observação, contudo, a um nível maior de detalhe, demonstra que, potencialmente, a produção poderia ser ainda maior. Para algumas subáreas do conhecimento foram trabalhadas algumas informações referentes à periodicidade de com que os docentes/pesquisadores publicam os resultados de suas pesquisas. Considerando basicamente as categorias de livros e artigos em periódicos publicados em ciclos de 4 anos, desde 1982, obtêm-se indicadores de razoável consistência e constância. Cerca de um terço à metade dos professores publicam ao menos um trabalho em cada ciclo; de 10% a 30% publicam pelo menos um trabalho em cada um de dois anos por ciclo; menos de 10% publicam pelo menos um trabalho em cada um de três anos por ciclo ; e menos de 5% publicam pelo menos um trabalho em cada um dos anos de cada ciclo.

Embora estes dados guardem aderência com a conhecida situação descrita pela "Lei de Lotka" e não seja recomendável propugnar uma epidemia de publicações, e tampouco haja necessária correlação entre quantidade e qualidade, causa espécie o considerável número de docentes que, por ciclo, não tem publicado ou o tem feito de modo bastante modesto. O aumento de produção científica pelo incremento da produtividade parece apresentar um potencial não desprezível. Por outro lado, cabe indagar das razões que estão determinando este comportamento da realidade e dos dados que a retratam, em particular aquelas relacionadas com as condições de trabalho e investigação.

Passando a considerar a questão dos trabalhos de dissertação/tese, as indicações disponíveis sugerem que um considerável número não se transformam em trabalhos publicados, fato mais verdadeiro entre as Ciências Humanas e Sociais e as da Saúde, do que entre as Ciências Biológicas, por exemplo. É também reduzida a incidência de trabalhos publicados em conjunto por docentes e discentes. É modesto o volume de participações discentes em congressos para apresentação de suas pesquisas de tese. Contam-se poucas instituições que exigem, para a defesa da dissertação ou tese, que um artigo dela resultante já tenha sido aceito para publicação em periódico de bom nível. São igualmente raras aquelas que ao menos solicitam a entrega, junto à dissertação, da forma para envio para publicação.

A falta do exercício de publicação dos trabalhos discentes, não submetendo-os ao crivo mais amplo do conjunto de docentes/pesquisadores, além de constituir permanente ameaça ao nível qualitativo dos trabalhos (embora aqui não se levante nenhuma suspeita apriorística quanto à competência dos orientadores e das bancas examinadoras), também contribui negativamente para um provável desperdício na investigação de temas, que podem estar sendo duplicados, sem troca de informações que poderia possibilitar o aprofundamento em conjunto da investigação de dado tópico. Este círculo é mais perverso obviamente nas áreas de pós-graduação e pesquisa menos consolidadas no País. No entanto, mesmo entre as mais consolidadas, são poucas aquelas em que a disseminação dos resultados das pesquisas discentes se faz com fluidez.

O nível geral de qualidade da pós-graduação

Desde 1975, a CAPES mantém um sistema de acompanhamento e avaliação dos cursos de mestrado e doutorado no País. Baseado no princípio de que a melhor metodologia é a de avaliação por pares, a CAPES, de início anualmente e agora a cada dois anos, vem reunindo Comissões de Consultores Científicos, formadas por docentes e pesquisadores nacionais por área/subárea do conhecimento, que, com base em relatórios anuais enviados pelos cursos à CAPES, relatórios de visitas de especialistas aos programas e outras informações complementares, realizam a avaliação periódica do desempenho da pós-graduação, por curso e por área.

Tal avaliação contempla o perfil do corpo docente de cada curso (dimensão, qualificação, dedicação, distribuição segundo as especialidades), atividades de ensino, atividades de pesquisa, produção científica, técnica e/ou artística docente e discente e alguns aspectos de processamento (fluxo de alunos, tempo médio de titulação, entre outros). Cada curso é avaliado dentro de sua própria evolução e comparado com seus congêneres de área/subárea. O resultado desta análise comparativa é expresso numa escala conceitual de cinco letras que se inicia pelo "A", atribuído aos programas de melhor nível, e finaliza pelo "E", conferido aos cursos de desempenho pouco aceitável.

A análise dos resultados desta avaliação sugere que um contingente significativo dos cursos atingiu níveis qualitativos bastante bons. Cerca de 65 a 70% dos programas de mestrado situam-se nas faixas conceituais "A" e "B". Agregando-se os localizados na faixa "C", alcança-se a proporção de 90%. Entre os doutorados, os percentuais são de 70% e 91%, respectivamente. O estudo mais detalhado destes dados indica a existência de uma correlação positiva entre conceito do curso, volume de sua produção, natureza da produção (especialmente publicações em periódicos e assemelhados) e produção docente **per capita**.

Assim, se resta ainda muito por fazer em termos da consolidação da pós-graduação brasileira, as informações disponíveis sugerem que os resultados dos esforços realizados podem ser considerados satisfatórios. Trata-se hoje de um momento de consolidação, que corre o risco de não se efetivar ou, ainda pior, tranformar-se num momento de desmantelamento e perdas, como se verá a seguir.

EDUCATION AND GROWTH IN BRAZIL:

SOME CROSS-SECTIONAL EVIDENCE

Lawrence J. Lau, Dean T. Jamison, Shu-Cheng Liu and Steven Rivkin[1]

March 1991

Department of Economics
Stanford University
Stanford, California 94305-6072
U.S.A.

---

[1]This paper is prepared for presentation at the Workshop on Education, Growth and Inequality in Brazil, March 25-26, 1991, Rio de Janeiro, Brazil. The authors are, respectively, Professor of Economics, Stanford University, Professor of Education and Public Health, University of California at Los Angeles, Visiting Scholar, Stanford University and Ph. D. Candidate, University of California at Los Angeles. The authors wishes to thank the World Bank for financial support. They are grateful to Dr. Nancy Birdsall and Dr. Robert Kaplan for helpful discussions. The views expressed herein are the authors' and do not necessarily reflect those of the World Bank or the institutions with which they are affiliated.

1. Introduction

Physical capital, labor, human capital and technical progress (or equivalently, total factor productivity) are the four principal sources of economic growth of nations. The rate of growth of labor is generally constrained by the rate of growth of population. For industrialized countries, the rate of growth of the labor force is seldom higher than two percent per annum, even with international migration. For developing countries, the rate of growth of the labor force is rarely higher than five percent, despite a generally higher rate of growth of population. Consequently, the rate of growth of capital, both physical and human, and technical progress have been found to account for a major proportion of economic growth, especially for countries with high growth rates. For example, Jorgenson, Gollop and Fraumeni (1987) find that between 1948 and 1979, capital formation accounted for 46 percent of the economic growth of the United States, growth of labor input, including human capital, accounted for 31 percent, and technical progress accounted for 24 percent.[2]

The importance of the contributions of capital, both physical and human, and technical progress to the growth of output can be readily understood with the help of some simple arithmetic. Starting with an aggregate production function:

$$Y = F(K, L, ED, t), \tag{1.1}$$

where Y is real output, K, L, and ED are the quantities of physical capital, labor and human capital respectively, and t is an index of chronological time, the rate of growth of output can be expressed in the familiar equation of growth accounting:

$$\frac{d\ell nY}{dt} = \frac{\partial \ell nE}{\partial \ell nK} \cdot \frac{d\ell nK}{dt} + \frac{\partial \ell nF}{\partial \ell nL} \cdot \frac{d\ell nL}{dt} + \frac{\partial \ell nF}{\partial ED} \cdot \frac{dED}{dt} + \frac{\partial \ell nF}{\partial t} \ . \tag{1.2}$$

[2]Jorgenson, Gollop and Fraumeni (1987), p. 21.

The four terms on the right-hand side of equation (1.2) may be identified as the contribution of capital, labor, human capital and technical progress respectively to the growth in output.

The production elasticity of output with respect to measured labor input can typically be estimated as approximately 0.6 for industrialized countries and between 0.3 and 0.4 for developing countries.[3] Thus, given the rate of growth of measured labor force, which is typically no higher than 2 percent per annum in industrialized countries and 5 percent per annum in developing countries, the maximum rate of growth that can be accounted for by the growth in labor input is on the order of 1.2 percent for developed countries and 2.0 percent for developing countries. Any growth in output of in excess of 2.0 percent per annum in a developing country is attributable to the growth in capital inputs and to technical progress. For a developing country that grows at 4 percent per annum, at least 50 percent of the growth in output may be attributed to physical and human capital and technical progress. In the short and intermediate runs, physical capital is especially important for another reason--it is the only input that can be readily varied. Human capital and technical progress, because of the long gestation periods that they entail, can be changed only in the longer run.

Lau and Yotopoulos (1989) introduce a new approach to the estimation of the aggregate production function with cross-sectional data. In Section 2, the new approach for studying the relationship between output, inputs and technical progress, based on the concept of an aggregate meta-production function,[4] is presented. In Section 3 there is a brief discussion of the data and the statistical model used. In Section 4, the results obtained from applying the model to cross-sectional state data from

[3]See, e.g., Boskin and Lau (1990) for estimates of the production elasticity of labor for industrialized countries and Lau, Jamison and Louat (1990) for estimates of the production elasticity of labor for developing countries. One may note that our estimate of the production elasticity of labor for Brazil, based on cross-section state data, is also approximately 0.4. See Section 4 below.

[4]The term "meta-production function" is due to Hayami and Ruttan (1970, 1985). See also Lau and Yotopoulos (1989) and Boskin and Lau (1990).

Brazil for the years 1970 and 1980 are presented. In Section 5, these results are interpreted in the light of findings from other studies. Some concluding remarks are made in Section 6.

2. The Meta-Production Function Approach

The new approach to the estimation of aggregate production functions from pooled time-series and cross-section data is based on the Lau and Yotopoulos (1989) modification of the concept of the meta-production function, introduced by Hayami and Ruttan (1970, 1985), through the use of time-varying, state- and commodity-specific augmentation factors. The basic assumptions for this approach are:

(1) All states have access to the same technology, that is, they have the same underlying aggregate production function F(.), sometimes referred to as a meta-production function, but may operate on different parts of it. The production function, however, applies to standardized, or "efficiency-equivalent", quantities of outputs and inputs, that is:

$$Y^*_{it} = F(K^*_{it}, L^*_{it}, ED_{it}) \quad , \quad i = 1,\ldots,n \quad ; \tag{2.1}$$

where Y^*_{it}, K^*_{it} and L^*_{it} are the "efficiency-equivalent" quantities of output, capital and labor respectively of the ith state at time t, ED_{it} is a measure of human capital of the ith state at time t, proxied by the average number of years of education of the labor force, and n is the number of states. The assumption of a meta-production function implies that F(.) does not depend on i (but may depend on t).

(2) There are differences in the technical efficiencies of production and in the qualities and possibly definitions of measured inputs across states. However, the measured outputs and inputs of the different states may be converted into standardized, or "efficiency-equivalent", units of inputs by multiplicative state- and output- and input-specific time-varying augmentation factors. The "efficiency-

equivalent" quantities of output and inputs of each state are not directly observable. They are, however, assumed to be linked to the measured quantities of outputs, Y_{it}'s, and inputs, K_{it}'s and L_{it}'s, through possibly time-varying, state- and commodity-specific augmentation factors $A_{ij}(t)$'s, $i = 1,...,n$; $j = K,L$[5]:

(2.2) $$Y^*_{it} = A_{i0}(t)Y_{it} \ ;$$

(2.3) $$K^*_{it} = A_{iK}(t)K_{it} \ ;$$

(2.4) $$L^*_{it} = A_{iL}(t)L_{it} \ ; \ i = 1, ..., n.$$

These assumption require some explanation. Together they imply that the aggregate production function is the same everywhere in Brazil in terms of standardized, or "efficiency-equivalent", units of outputs and inputs. Moreover, measured inputs of any state may be converted into equivalent units of measured inputs of another state. For example, one unit of capital in state A may be equivalent to two units of capital in state B; and one unit of labor in state A may be equivalent to one-third of a unit of labor in state B. These conversion ratios may also change over time. In terms of the measured quantities of inputs, the production functions of any two states are <u>not</u> necessarily the same. However, in terms of "efficiency-equivalent" units, the production functions are identical across states.

We note that in terms of the <u>measured</u> quantities of outputs, the production function may be rewritten as:

(2.5) $$Y_{it} = A_{i0}(t)^{-1}F(K^*_{it},L^*_{it},ED_{it}) \ , \ i = 1,...,n \ ;$$

so that the reciprocal of the output-augmentation factor $A_{i0}(t)$ has the interpretation of the possibly time-varying level of the technical efficiency of production, also referred to as output efficiency, in the ith state at time t.

There are many reasons why these commodity augmentation factors are not likely to be identical

---

[5]One can, in principle, also allow the quality of education to vary across states as well. However, given only two observations per state, the differences in the quality of education, if any, cannot be easily identifiable. It is therefore assumed that the average quality of education is approximately the same across states.

across states. Differences in climate, topography and infrastructure; differences in definitions and measurements; differences in quality; differences in the composition of outputs; and differences in the technical efficiencies of production are some examples. The commodity augmentation factors are introduced precisely to capture these differences across states. The commodity augmentation factors are assumed to have the constant exponential form with respect to time. Thus:

(2.6) $$Y^{*}_{it} = A_{i0} \exp(c_{i0}\, t) Y_{it} \;;$$

(2.7) $$K^{*}_{it} = A_{iK} \exp(c_{iK}\, t) K_{it} \;: \text{and}$$

(2.8) $$L^{*}_{it} = A_{iL} \exp(c_{iL}\, t) L_{it} \;;\; i = 1,\ldots,n;$$

where the A_{i0}'s, A_{ij}'s, c_{i0}'s, and c_{ij}'s are constants. We shall refer to the A_{i0}'s and A_{ij}'s as augmentation level parameters and c_{i0}'s and c_{ij}'s as augmentation rate parameters. For at least one state, say the ith, the constants A_{i0} and A_{ij}'s can be set identically at unity (or some other arbitrary constants), reflecting the fact that "efficiency-equivalent" outputs and inputs can be measured only relative to some standard. Econometrically this means that the constants A_{i0}'s and A_{ij}'s cannot be uniquely identified without some normalization.

In this study, the aggregate meta-production function is specified to have the Cobb-Douglas functional form. For a Cobb-Douglas production function with three inputs, physical capital, labor, and human capital (average education), the production function, in terms of "efficiency-equivalent" inputs, takes the form:

(2.9) $$\ln Y_{it} = -\ln A_{i0}(t) + \ln Y_0 + a_K \ln K^{*}_{it} + a_L \ln L^{*}_{it} + a_E\, ED_{it} \,.$$

By substituting equations (2.6) through (2.8) into equation (2.9), we obtain equation (2.10), which is written entirely in terms of observable variables:

(2.10) $$\ln Y_{it} = \ln Y_0 - \ln A_{i0} + a_K \ln A_{iK} + a_L \ln A_{iL} + a_K \ln K_{it} + a_L \ln L_{it} + a_E ED_{it} + (-c_{i0} + a_K\, c_{iK} + a_L\, c_{iL})\, t \,,$$

which simplifies into:

(2.11) $$\ell n\ Y_{it} = \ell n\ Y_0 + \ell n\ A^*_{i0} + a_K\ \ell n\ K_{it} + a_L\ \ell n\ L_{it} + c^*_{i0}t\ ,$$

where A^*_{i0} and c^*_{i0} are state-specific constants. Equation (2.11) is the most general specification possible under our maintained hypotheses of a single Cobb-Douglas meta-production function and constant exponential commodity-augmentation representation of technical progress.

If there were a minimum of three observations per state, then it is possible to estimate the parameters of equation (2.11). Taking first differences of equation (2.11), we obtain:

(2.12) $$\ell n\ Y_{it} - \ell n\ Y_{it-1} = c^*_{i0} + a_K\ (\ell n\ K_{it} - \ell n K_{it-1}) + a_L(\ell n\ L_{it} - \ell n\ L_{it-1}) + a_E(ED_{it} - ED_{it-1})$$

By taking first differences of equation (2.12), we obtain:

(2.13) $$(\ell n\ Y_{it} - \ell n\ Y_{it-1}) - (\ell n\ Y_{it-1} - \ell n\ Y_{it-2}) = a_K\ [(\ell n\ K_{it} - \ell n\ K_{it-1}) - (\ell n\ K_{it-1} - \ell n\ K_{it-2})]$$
$$+ a_L[\ell n\ L_{it} - \ell n\ L_{it-1}) - (\ell n\ L_{it-1} - \ell n\ L_{it-2})] + a_E\ [(ED_{it} - ED_{it-1}) - (ED_{it-1} - ED_{it-2})],$$

which can be used for the estimation of a_K , a_L , and a_E .

However, since only two observations per state are available, equation (2.13) cannot be implemented. Instead, we fall back on equation (2.12) and make the assumption that $c^*_{i0} = c^*_0$, that is, the rate of technical progress is the same across the states:

(2.14) $$\ell n\ Y_{it} - \ell n\ Y_{it-1} = c^*_0 + a_K\ (\ell n\ K_{it} - \ell n\ K_{it-1}) + a_L(\ell n L_{it} - \ell n\ L_{it-1}) + a_E\ (ED_{it} - ED_{it-1})\ .$$

Equation (2.14) is the actual estimating equation used in this study.

Our new approach is applied to Brazilian state data for 1970 and 1980. By pooling data across states, it is hoped that the separate effects of economies of scale and technical progress, usually confounded by the simultaneous expansion of scale with time in the data of a single state, can be more readily identified. (At any given point in time, production at different scales is observed. The same scale of production may be observed at different points in time.) Moreover, inter-state data typically have greater variability in the relative quantities of inputs than intra-state data, thus facilitating the identification and estimation of the aggregate production function.

3. The Data and the Statistical Model

Data

We use data from individual Brazilian states for the years 1970 and 1980. Data for 1975 are not used because comparable labor force data are not available for that year. A number of states are also omitted from the sample because of incomplete data. A list of the individual Brazilian states and their code numbers is presented in Appendix 1. What follows is a brief description of the variables.[6]

(1) Real Output (Y)

The aggregate real output of each state is measured as the real Gross Domestic Product (GDP) in 1970 prices.

(2) Capital (K)

Capital stock data are not available for the individual Brazilian states. As a proxy, the annual quantity of industrial consumption of electricity in each state in 1970 and 1980 is used instead.

(3) Labor (L)

Labor is measured as the number of persons in the economically active population.

(4) Human Capital (ED)

Human capital is measured as the average number of years of formal education per person of the labor force (also referred to as the economically active population). The data are presented in Table 3.1.

(5) Time (t)

Time is measured in terms of the number of years chronologically with the year 1970 being set equal to zero.

---

[6]The authors are grateful to Dr. Robert Kaplan for providing the basic data used in this analysis.

The average annual rates of growth of real output, physical capital, labor and average education are presented in Table 3.2. The decade of the 1970s, especially the latter half of it, is a period of extraordinarily rapid growth. The average annual rate of growth of real GOP for Brazil as a whole was approximately 10 percent. It is apparent from Table 3.2 that State 4, Amapa, must have been an outlier. It turns out that the empirical results are affected significantly depending on whether State 4 is included or not. The results presented here are based on a sample without State 4.

The Statistical Model

We introduce stochastic disturbance terms $\in_{it}$'s into the <u>first-differenced</u> form of the natural logarithm of the aggregate production function (equation 2.14). We assume:

(3.1) $E(\in_{it}) = 0$;

and

(3.2) $V(\in_{it}) = \sigma^2$; $\forall i, t$;

and the stochastic disturbance terms are uncorrelated across states. In the first-differenced form, our stochastic assumptions amount to saying that the influence of the stochastic disturbance terms is permanent--they raise or lower the production function permanently until further changes caused by future stochastic disturbance terms. In fact, given only two observations per state, unless the A^*_{i0}'s are assumed to be identical across states, a model without first-differencing cannot be implemented.

4. Empirical Results

As mentioned earlier, the decade of the 1970s was a period of extremely rapid growth in Brazil. In most of the states, capital, labor, and average education grew rapidly. This led to an identification problem--it is difficult to disentangle the effects due to each of the inputs.

Tests of Hypotheses

First, the hypothesis of no educational effect is tested. This hypothesis is rejected when no trend term (c^*_0) is included. When a trend term is included, we can also reject the hypothesis that education has no effect. Next, we test the hypothesis of constant returns to scale in the physical inputs, capital and labor. This hypothesis is rejected at the 5 percent level of significance. These results are presented in Tables 4.1 and 4.2.

The estimation results from the first-differenced equation is presented in Table 4.3. The estimated effect of education ranges from 0.21 with the trend term and 0.48 without the trend term. In an attempt to separate the effect of technical progress, represented by the trend term, and the effect of education, we subtract an assumed rate of technical progress of between 1 and 5 percent per annum from the rate of growth of real output before running the regression in equation (2.14) without the trend term.

The results are presented in Table 4.4. The effect of education remains positive and statistically significant. What it says is that even with a rate of technical progress of 5 percent per annum, an increase in average education of the labor force of one year increases output by at least 17.8 percent. This is a very large effect. The effects estimated by Lau, Jamison and Louat (1990) are much smaller.

A 5 percent per annum rate of growth of total factor productivity or technical progress is an extremely high rate. Traditional estimates of the rate of growth of technical progress are on the order of a couple percent per annum. The Boskin and Lau (1990) estimates of the rates of technical progress for France, W. Germany, Japan, the United Kingdom and the United States in the postwar period do not exceed 2.0 percent per annum if constant returns to scale are maintained and do not exceed 4.0 percent otherwise. The Lau (1989) estimates of the rate of technical progress for China, Hong Kong, Singapore and Taiwan between 1952 and 1984 do not exceed 4.3 percent. It is reasonable to suppose that the rate of technical progress in Brazil during the 1970s does not exceed 5 percent per annum.

Estimates of the Combined Contribution of Human Capital and Technical Progress

Equation (1.2) may be used to obtain an alternative estimate of the growth in output due to human capital and technical progress. Rewriting equation (1.2) as:

$$(4.1) \qquad \frac{\partial \ell nF}{\partial E}\frac{dE}{dt} + \frac{\partial \ell nF}{\partial t} = \frac{d\ell nY}{dt} - \frac{\partial \ell nF}{\partial \ell nK}\frac{d\ell nK}{dt} - \frac{\partial \ell nF}{\partial \ell nL}\frac{d\ell nL}{dt},$$

the combined contribution of human capital and technical progress may be estimated once values of the production elasticities with respect to physical capital and labor are specified, the rates of growth of physical capital and labor being known, observed quantities. Such an exercise has been carried out with the capital elasticity assumed to be 0.6 and then 0.4 and the labor elasticity assumed to be 0.4. The results are presented in the last column of Table 4.5. The estimated contributions are unfortunately small and mostly negative for a production elasticity of capital of 0.6, with an (unweighted) average of -4 percent per annum. They are higher and more reasonable with a production elasticity of capital of 0.4. This finding of a negative and/or small residual is primarily due to the rapid rate of growth in the capital input.

The estimated combined contributions of human capital and technical progress are plotted against the rates of growth of average education. A positive correlation is clearly discernible.

5. Comparison with Other Studies

6. Concluding Remarks

We have implemented a new method of analyzing productivity and technical progress, based on the concept of an aggregate meta-production function, using two cross-sections of data from individual

Brazilian states.

We have not made explicit adjustments for the quality of capital or labor, as were done by Denison (1962, 1967, 1979, 1985) and Jorgenson, Gollop and Fraumeni (1987). Instead, we introduce human capital, measured as the average number of years of formal education per person of the labor force, as an explicit variable in the aggregate production function. We also introduce a time trend to capture the effect of technical progress. Any improvement in inputs not captured by the human capital variable should be reflected in the technical progress term.

References

Boskin, M.J. and L.J. Lau (1990), "Post-War Economic Growth of the Group-of-Five Countries: A New Analysis," Technical Paper No. 217, Stanford, CA: Center for Economic Policy Research, Stanford University (mimeographed).

Bowman, M.J. (1980), "Education and Economic Growth: An Overview," in T. King, ed., Education and Income: A Background Study for World Development Report, 1980, World Bank Staff Working Paper No. 402, Washington, D.C., 1-71.

Bowman, M.J. (1987), "The Relevance of Education," in G. Psacharopoulos, ed., Economics of Education: Research and Studies, Oxford: Pergamon Press, 305-307.

Denison, E.F. (1962), "United States Economic Growth," Journal of Business, 35: 109-121.

Denison, E.F. (1967), Why Growth Rates Differ: Post-War Experience in Nine Western Countries, Washington, D.C.: Brookings Institution.

Denison, E.F. (1979), Accounting for Slower Economic Growth: The United States in the 1970s, Washington, D.C.: Brookings Institution.

Denison, E.F. (1985), Trends in American Economic Growth, 1929-1982, Washington, D.C.: Brookings Institution.

Denison, E.F. and W. Chung (1976), How Japan's Economy Grew So Fast, Washington, D.C.: Brookings Institution.

Denison, E.F. and R.P. Parker (1980), "The National Income and Product Accounts of the United States: An Introduction to the Revised Estimates for 1929-80," Survey of Current Business, 60 (12): 1-26.

Griliches, Z. and D.W. Jorgenson (1966), "Sources of Measured Productivity Change: Capital Input," American Economic Review, 56: 50-61.

Hayami, Y. and V.W. Ruttan (1970), "Agricultural Productivity Differences Among Countries," American Economic Review, 60: 895-911.

Hayami, Y. and V.W. Ruttan (1985), Agricultural Development: An International Perspective, revised and expanded ed., Baltimore: Johns Hopkins University Press.

Hicks, N.L. (1980), "Is There a Trade-Off between Growth and Basic Needs?" Finance and Development, 17: 17-20.

Hicks, N.L. (1987), "Education and Economic Growth," in G. Psacharopoulos, ed., Economics of Education: Research and Studies, Oxford: Pergamon Press, 101-107.

Jamison, D.T. and L.J. Lau (1982), Farmer Education and Farm Efficiency, Baltimore: The Johns Hopkins University Press.

Jorgenson, D.W., F.M. Gollop and B.M. Fraumeni (1987), Productivity and U.S. Economic Growth, Cambridge, MA: Harvard University Press.

Jorgenson, D.W. and Z. Griliches (1967), "The Explanation of Productivity Change," Review of Economic Studies, 34: 249-283.

Jorgenson, D. W. and Z. Griliches (1972), "Issues in Growth Accounting: A Reply to Edward F. Denison," Survey of Current Business, 52 (5), Part II: 65-94.

Kawagoe, T., Y. Hayami and V.W. Ruttan (1985), "The Intercountry Agricultural Production Function and Productivity Differences Among Countries," Journal of Development Economics, 19: 113-132.

Lau, L.J. (1980), "On the Uniqueness of the Representation of Commodity-Augmenting Technical Change," in L.R. Klein, M. Nerlove and S.C. Tsiang, eds., Quantitative Economics and Development: Essays in Memory of Ta-Chung Liu, New York: Academic Press, 281-290.

Lau, L.J. (1989), "A Comparative Analysis of Economic Development Experiences in Chinese Societies," in Y.C. Jao, V. Mok and L.S. Ho, eds., Economic Development in Chinese Societies: Models and Experiences, Hong Kong: Hong Kong University Press, 9-23.

Lau, L.J., D.T. Jamison and F.F. Louat (1990), "Education and Productivity in Developing Countries: An Aggregate Production Function Approach," Working Paper, Washington, D.C.: The World Bank (mimeographed).

Lau, L.J. and P.A. Yotopoulos (1989), "The Meta-Production Function Approach to Technological Change in World Agriculture," Journal of Development Economics, 31: 241-269.

Lau, L.J. and P.A. Yotopoulos (1990), "Intercountry Differences in Agricultural Productivity: An Application of the Meta-Production Function," Working Paper, Stanford, CA: Department of Economics, Stanford University (mimeographed).

Lindbeck, A. (1983), "The Recent Slowdown of Productivity Growth," Economic Journal, 93: 13-34.

Lockheed, M.E., D.T. Jamison and L.J. Lau (1980), "Farmer Education and Farm Efficiency: A Survey," Economic Development and Cultural Change, 29: 37-76.

McMahon, W.W. (1987), "Expected Rates of Returns to Education," in G. Psacharopoulos, ed., Economics of Education: Research and Studies, Oxford: Pergamon Press, 187-196.

McMahon, W.W. (1984), "The Relation of Education and R & D to Productivity Growth," Economics of Education Review, 3: 299-313.

Mincer, J. (1974), Schooling, Experience and Earnings, New York: Columbia University Press.

Mincer, J. (1979), "Human Capital and Earnings," in D. M. Windham, ed., Economic Dimensions of Education, Washington, D.C.: National Academy of Education, 1-31.

Psacharopoulos, G. (1985), "Returns to Education: A Further International Update and Implications," The Journal of Human Resources, 20: 584-604.

Psacharopoulos, G. (1987), "Earnings Functions," in G. Psacharopoulos, ed., Economics of Education: Research and Studies, Oxford: Pergamon Press, 218-223.

Psacharopoulos, G. and A.M. Arriagada (1986), The Educational Attainment of the Labor Force: An International Comparison, Discussion Paper No. 38, Education and Training Series, Washington, D.C.: The World Bank.

Schultz, T.W. (1961), "Investment in Human Capital," American Economic Review, 51: 1-17.

Schultz, T.W. (1975), "The Value of the Ability to Deal with Disequilibria," Journal of Economic Literature, 13: 872-876.

Solow, R.M. (1956), "A Contribution to the Theory of Economic Growth," Quarterly Journal of Economics, 70: 65-94.

Solow, R.M. (1957), "Technical Change and the Aggregate Production Function," Review of Economics and Statistics, 39: 312-320.

Appendix 1

State Codes

Rondonia
Roraima
Amapa
Acre
Amazonas
Para
Maranhao
Piaui
Ceara
Riogrande do Norte
Paraiba
Pernambuco
Alagoas
Sergipe
Bahia
Minas Gerais
Espirito Santo
Rio De Janeiro
Soa Paulo
Parana
Santa Catarina
Rio Grande Do Sul
Mato Grosso Do Sul } combined as 25
Mato Grosso
Goias
Distrito Federal

Table 3.1

Average Years of Education of Economically Active Population

State	Male 1970	Male 1980	Female 1970	Female 1980	All 1970	All 1980
2	2.09	2.75	4.31	4.68	2.32	3.06
3	2.62	3.48	4.77	5.91	2.94	3.98
4	2.73	3.90	4.41	5.93	3.02	4.38
6	2.01	3.24	4.09	5.19	2.32	3.71
7	2.22	3.13	3.45	4.95	2.44	3.51
8	1.00	1.71	1.41	2.61	1.08	1.93
9	0.92	1.66	2.01	3.55	1.10	2.08
10	1.20	2.05	2.68	3.96	1.45	2.55
11	1.36	2.36	3.31	4.64	1.64	2.93
12	1.13	2.01	2.56	4.09	1.36	2.50
13	1.80	2.82	2.88	4.16	2.03	3.20
14	1.04	1.87	2.03	3.23	1.23	2.22
15	1.25	2.37	1.92	3.71	1.41	2.75
16	1.40	2.23	2.19	3.59	1.56	2.57
17	2.47	3.81	4.07	5.51	2.77	4.24
18	2.47	4.04	4.40	5.79	2.79	4.46
19	4.98	5.93	5.49	6.66	5.11	6.16
20	4.04	5.22	4.81	6.06	4.23	5.48
21	2.37	3.92	3.56	5.14	2.57	4.22
22	3.18	4.66	4.19	5.71	3.38	4.94
23	3.85	4.87	4.68	6.01	4.06	5.22
24		3.59		5.16		3.93
25	2.05	3.11	3.70	5.44	2.25	3.52
26	1.92	3.33	3.73	5.35	2.17	3.77
27	5.09	6.59	5.94	7.21	5.32	6.81
Average	2.78	3.99	4.00	5.39	3.03	4.37

Table 3.2

Average Annual Growth Rates of Brazilian States (1970-1980)

State	GDP (Y)	Capital (K)	Labor (L)	Average Education (Year)
2	17.73%	52.08%	16.22%	0.0739
3	12.27%	27.10%	8.51%	0.1031
4	4.88%	24.54%	5.25%	0.1360
6	14.07%	22.42%	5.03%	0.1380
7	12.77%	26.50%	5.04%	0.1060
8	9.71%	23.29%	2.96%	0.0850
9	9.77%	23.05%	3.09%	0.0984
10	10.14%	14.14%	3.12%	0.1092
11	10.77%	21.16%	3.72%	0.1298
12	8.57%	14.75%	2.22%	0.1139
13	8.08%	10.51%	3.03%	0.1175
14	9.23%	18.45%	2.38%	0.0982
15	8.76%	13.46%	2.86%	0.1335
16	10.79%	23.73%	2.76%	0.1002
17	10.60%	13.41%	3.14%	0.1463
18	11.56%	23.48%	4.40%	0.1675
19	7.85%	10.60%	3.92%	0.1049
20	8.95%	10.53%	4.91%	0.1247
21	10.16%	13.85%	2.29%	0.1650
22	11.20%	16.02%	4.30%	0.1558
23	8.71%	13.41%	3.45%	0.1170
25	13.79%	15.65%	5.85%	0.1507
26	11.43%	32.26%	4.09%	0.1597
27	13.97%	40.60%	9.76%	0.1493
Average	10.66%	21.04%	4.68%	0.1243

Table 4.1 Tests of Hypothesis of No Education Effect

Cases	Test-Statistics	Critical Value at 5% level of significance
With trend	4.849	4.38
Without trend	108.437	4.35

Table 4.2 Tests of Hypothesis of Constant Returns

Cases	Test-Statistics	Critical Value at 5% level of significance
With trend	22.606	4.38
Without trend	13.230	4.35

Table 4.3 Estimates from First-Differenced Equation

Variables				
Constant	0.433 (3.1)	0.719 (11.5)		
ln Capital	0.100 (2.4)	0.095 (2.1)	0.139 (2.9)	0.427 (4.5)
ln Labor	0.400 (2.9)	0.373 (2.5)	0.410 (2.5)	0.146 (0.4)
Average Education	0.212 (2.2)		0.481 (10.4)	
Adjusted R^2	0.757	0.710	0.719	0.662

Note: Numbers in parentheses are t-ratios.

Table 4.4 Estimates under Alternative Assumptions of the Rate of Technical Progress

Variables	Rate of Technical Progress (percent per annum)				
	Variable Returns				
	1%	2%	3%	4%	5%
ln Capital	0.130 (3.0)	0.121 (2.9)	0.112 (2.8)	0.104 (2.7)	0.095 (2.4)
ln Labor	0.408 (2.7)	0.407 (2.9)	0.403 (3.0)	0.401 (3.0)	0.398 (3.0)
Average Education	0.419 (9.7)	0.358 (8.8)	0.297 (7.7)	0.238 (6.3)	0.178 (4.7)
Adjusted R^2	0.737	0.751	0.763	0.769	0.768

Table 4.5

Estimates of the Combined Contribution of Human Capital & Technical Progress

State	$\dot{Y}/Y-\dot{L}/L$	Value of Capital Elasticity (Sk)	
		0.6	0.4
2	1.51%	-20.01%	-9.60%
3	3.76%	-7.39%	-1.97%
4	-0.38%	-11.95%	-7.04%
6	9.04%	-1.39%	3.09%
7	7.73%	-5.14%	0.16%
8	6.76%	-5.45%	-0.79%
9	6.69%	-5.29%	-0.68%
10	7.02%	0.41%	3.24%
11	7.05%	-3.42%	0.82%
12	6.36%	-1.16%	1.79%
13	5.06%	0.56%	2.67%
14	6.86%	-2.79%	0.90%
15	5.90%	-0.46%	2.24%
16	8.02%	-4.55%	0.19%
17	7.46%	1.30%	3.98%
18	7.16%	-4.29%	0.41%
19	3.92%	-0.08%	2.04%
20	4.04%	0.67%	2.78%
21	7.87%	0.94%	3.71%
22	6.90%	-0.13%	3.07%
23	5.25%	-0.72%	1.96%
25	7.94%	2.06%	5.19%
26	7.34%	-9.57%	-3.11%
27	4.21%	-14.29%	-6.17%
Average	5.98%	-3.84%	0.37%

Note: The combined contribution is computed as:

$\dot{Y}/Y - Sk\,\dot{K}/K - 0.4\,\dot{L}/L$

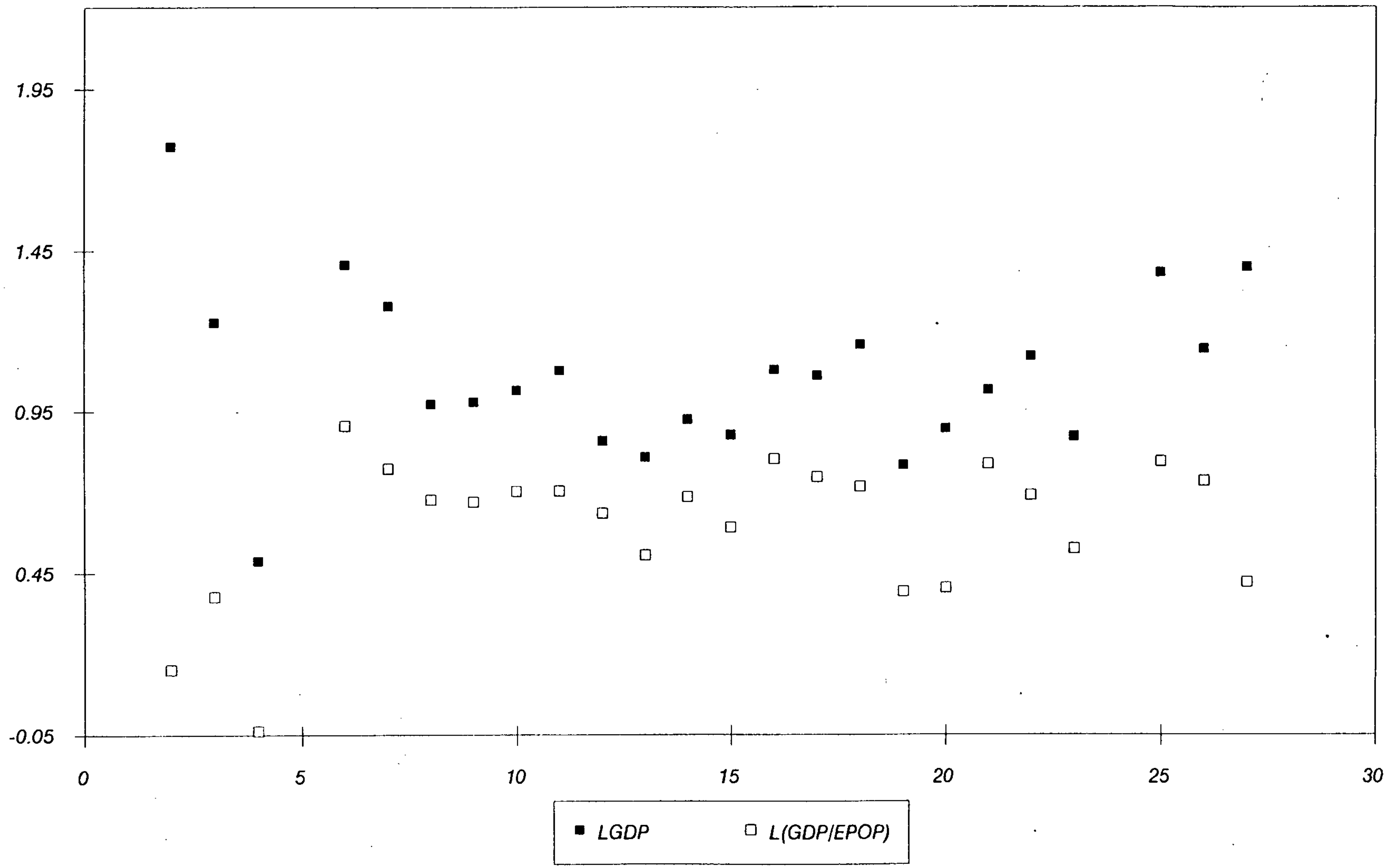


Figure 3.1

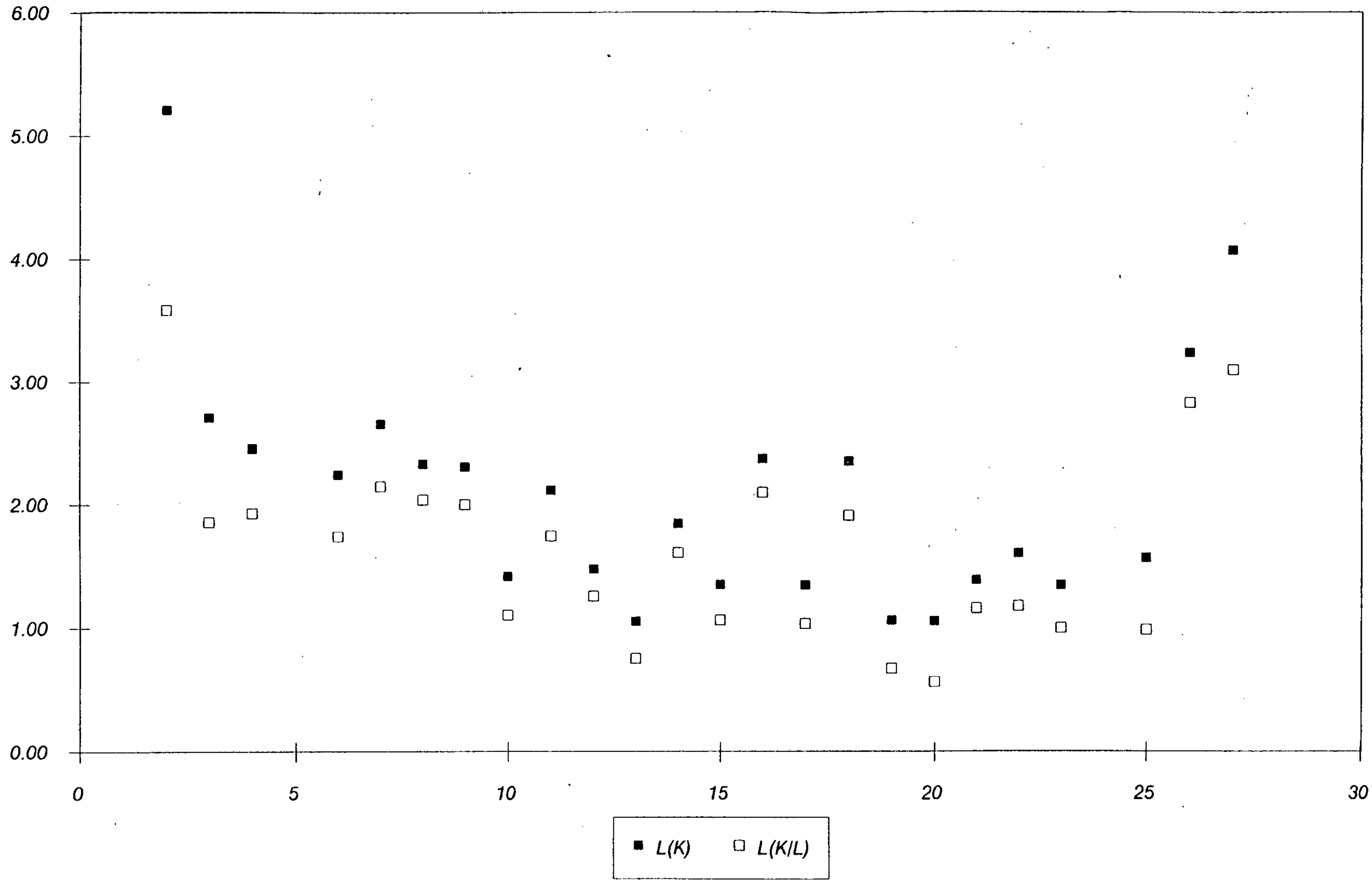


Figure 3.2

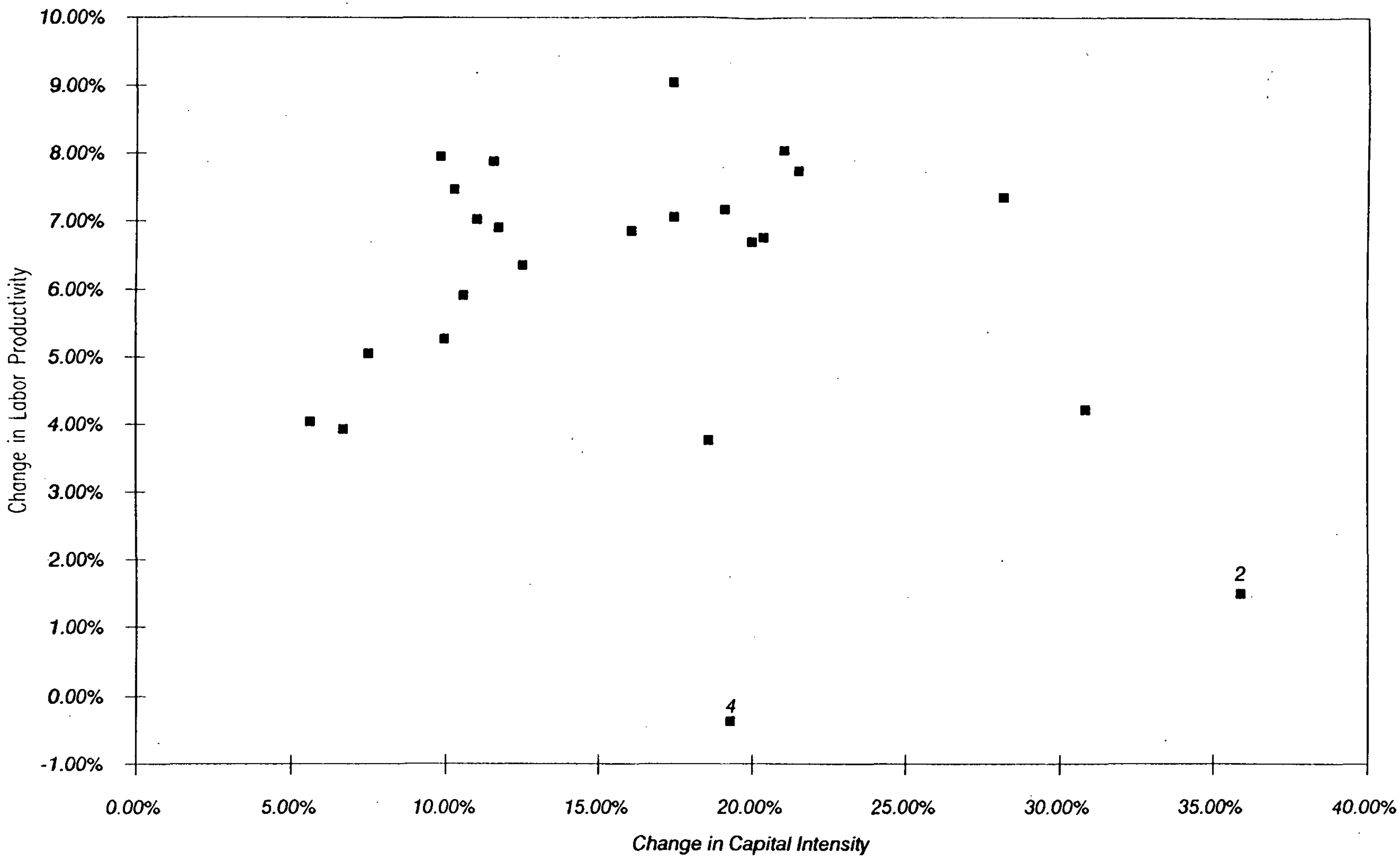

10.00%
9.00%
8.00%
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
-1.00%
Change in Labor Productivity
2
4
0.00%
5.00%
10.00%
15.00%
20.00%
25.00%
30.00%
35.00%
40.00%
Change in Capital Intensity

Figure 3.3

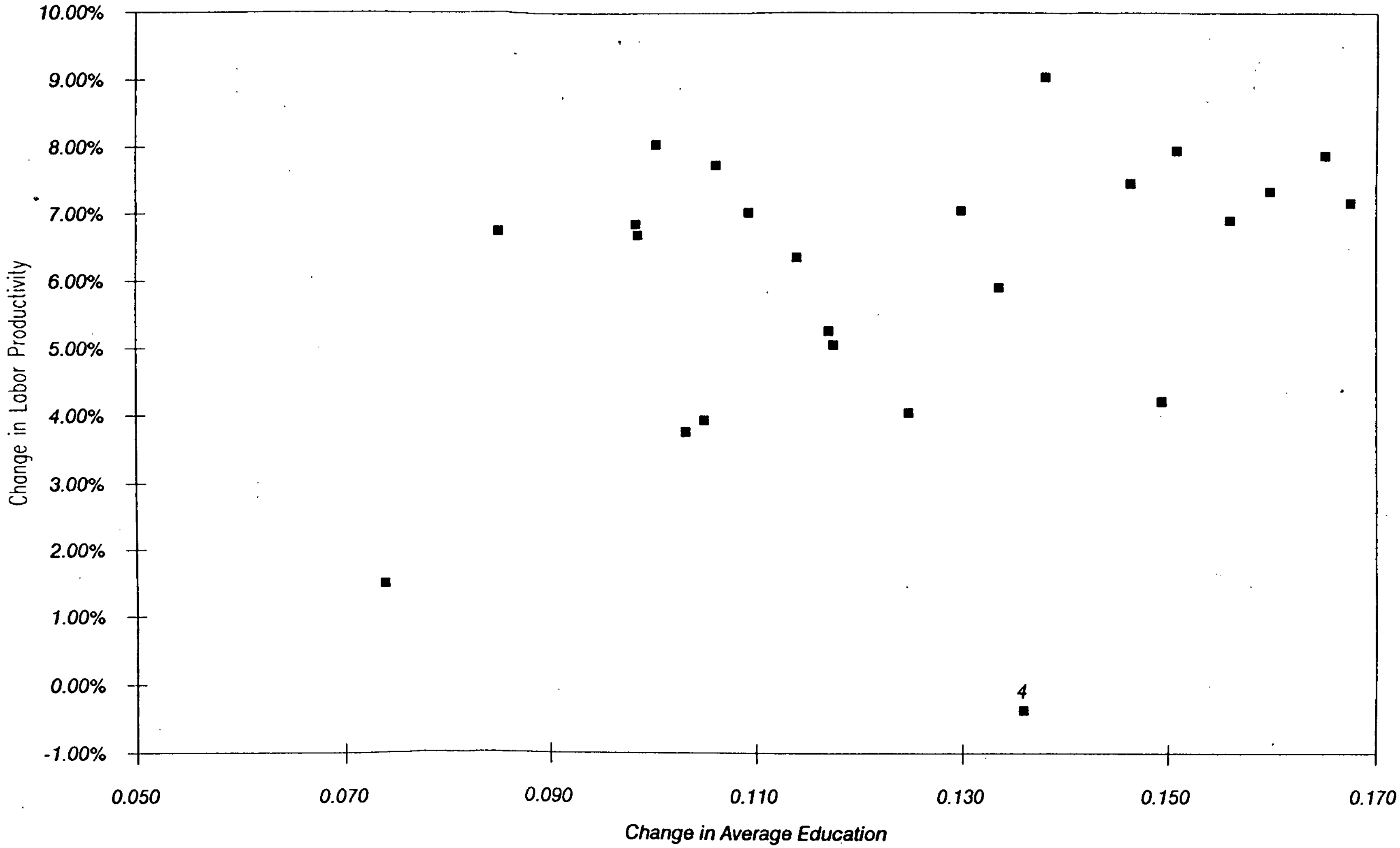
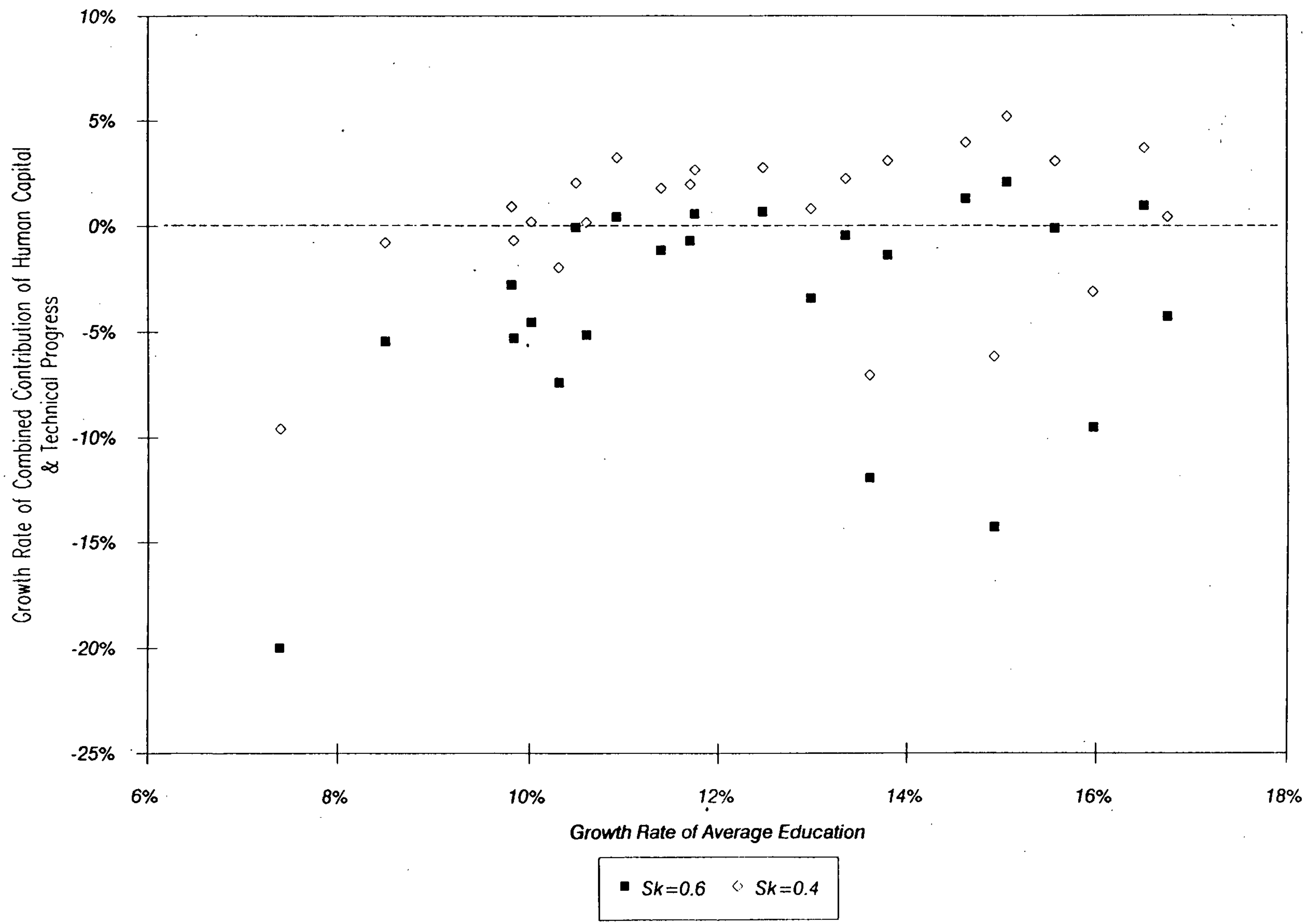

10.00%
9.00%
8.00%
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
-1.00%
Change in Labor Productivity
4
0.050
0.070
0.090
0.110
0.130
0.150
0.170
Change in Average Education

Figure 3.4


Growth Rate of Combined Contribution of Human Capital & Technical Progress
10%
5%
0%
-5%
-10%
-15%
-20%
-25%
6%
8%
10%
12%
14%
16%
18%
Growth Rate of Average Education
Sk=0.6
Sk=0.4