ISSN 1415-4765

TEXTO PARA DISCUSSÃO Nº 850

Reducing Schooling Inequality in Brazil: Demographic Opportunities and Inter-cohort Differentials

Carlos Eduardo Velez
Sergei Soares
Marcelo Medeiros

DEZEMBRO DE 2001

ISSN 1415-4765

TEXTO PARA DISCUSSÃO Nº 850

REDUCING SCHOOLING INEQUALITY IN BRAZIL: DEMOGRAPHIC OPPORTUNITIES AND INTER-COHORT DIFFERENTIALS

Carlos Eduardo Velez*
Sergei Soares**
Marcelo Medeiros**

Rio de Janeiro, dezembro de 2001

* Do Banco Mundial.

** Da Diretoria de Estudos Sociais do IPEA.

SUMMARY

RESUMO

O objetivo deste trabalho é a exploração das relações entre escolaridade e demografia no Brasil. Gostaríamos de apresentar uma investigação preliminar sobre quanto tempo melhorias no sistema educacional vão demorar para se refletir na distribuição educacional da população em idade ativa (PIA). Este tempo de resposta depende de dois fatores. O primeiro é a composição etária da população em idade ativa — o peso de cada coorte na população de 16 a 70 anos. O segundo fator é a distribuição da instrução formal dentro de cada coorte — a média e a desigualdade de anos de estudo completados com sucesso. Estes dois fatores — demografia e educação por coorte — definem tanto o nível educacional médio como a distribuição da educação dentro da PIA para um ano qualquer.

Neste texto, usamos uma projeção demográfica padrão e fazemos várias hipóteses sobre a evolução tanto da média como da desigualdade educacional. De acordo com essas hipóteses, é possível calcular quanto tempo melhorias no nível educacional de coortes sucessivas vão levar para se traduzir em melhorias significativas na distribuição da escolaridade da PIA.

Os resultados, infelizmente, são um tanto pessimistas. Calculamos que até melhorias fortes com relação à tendência observada vão demorar anos ou até décadas para se transformar em dotações educacionais significativamente maiores e melhor distribuídas para a PIA. Em outras palavras, a inércia demográfica é um fator forte impedindo transformações dramáticas na distribuição da educação em períodos menores que algumas poucas décadas.

ABSTRACT

The objective of this paper is to explore the interplay between schooling and demographics in Brazil. We would like to provide a preliminary answer to the question of how long will improvements in schooling of younger cohorts take o change the distribution of educational endowments of the total labor force.

This answer depends on two factors. The first is the demographic composition of the working age population — the weight each cohort has in the 16 to 70 year old population. The second is the distribution of schooling within each cohort — its average educational level and the inequality within each cohort. These two factors — demography and education by cohort — define the average educational level and the distribution of education for the working age population in any given year.

This paper takes a standard demographic projection and makes various hypotheses about the evolution of education — both the mean and inequality. According to these hypotheses, we will calculate how long improvements in the schooling of successive cohorts take to translate into significant improvements in the schooling of the working age population.

Our results are somewhat pessimistic. We calculate that even very strong departures from the observed trend will take many years or decades to translate into significantly different educational endowments for the working age population. In other words, we show that demographic inertia is a strong factor preventing changes in educational endowments in periods shorter than a few decades.

1 - INTRODUCTION

How long will improvements in schooling of younger cohorts take to change the distribution of educational endowments of the total labor force and, in turn, change the distribution of labor income in Brazil? When rates of return to schooling are significant, as they are in Brazil, the size and distribution of educational endowments determines to a large extent the distribution of labor income. However, improvements in the educational attainment of younger cohorts do not translate immediately into proportional improvements for all cohorts of the economy. Demographics might play an important role in that process.

This paper attempts to develop a demographic model linking the educational profiles of successive cohorts of individuals entering the labor force with the level and inequality of educational endowments of the whole labor force. We ask how the demographic transition might affect the impact of cohort educational profiles on the level and inequality of educational endowments of the entire labor force. For example, an aggressive education policy to improve high school completion should result in large differences between the educational profiles of younger versus older population cohorts in the labor force, with obvious effects on the distribution of the educational stock. However, the size and speed of that effect will depend on the pace of demographic transition in Brazil. Presumably, if demographic transition has not been completed and the fertility rates remain high, the effect will be larger and faster.

The profile of population growth for Brazil will show how large those effects will be and how long it will take to observe them. Moreover, it will show whether Brazil's position in the demographic transition provides an opportunity (or makes it more difficult) to reduce the inequality, or improve the level, of the educational stock of the labor force. Hopefully this model could be a device to show how much time is required to recover the full social benefits of sustained investments in education. That is, to understand the links and the lags between current policy actions and future outcomes. In other words, it would enhance the value of current policies in terms of the equity improvements for present and future generations.

2 - DEMOGRAPHIC BACKGROUND

According to the 2000 Census, the Brazilian population amounts to 170 million people, most of which live in the coastal urban area. Spatial differences are strong. Brazil is divided by geographers into five regions: South, Southeast, Center-west, North, and Northeast. The first two are the most developed and rich, the last one is the poorest. Population density can be considered high in the metropolitan areas of all regions, medium in non-metropolitan areas of the Southern and Southeastern regions and low in rural areas of the Northern and Center-western regions.

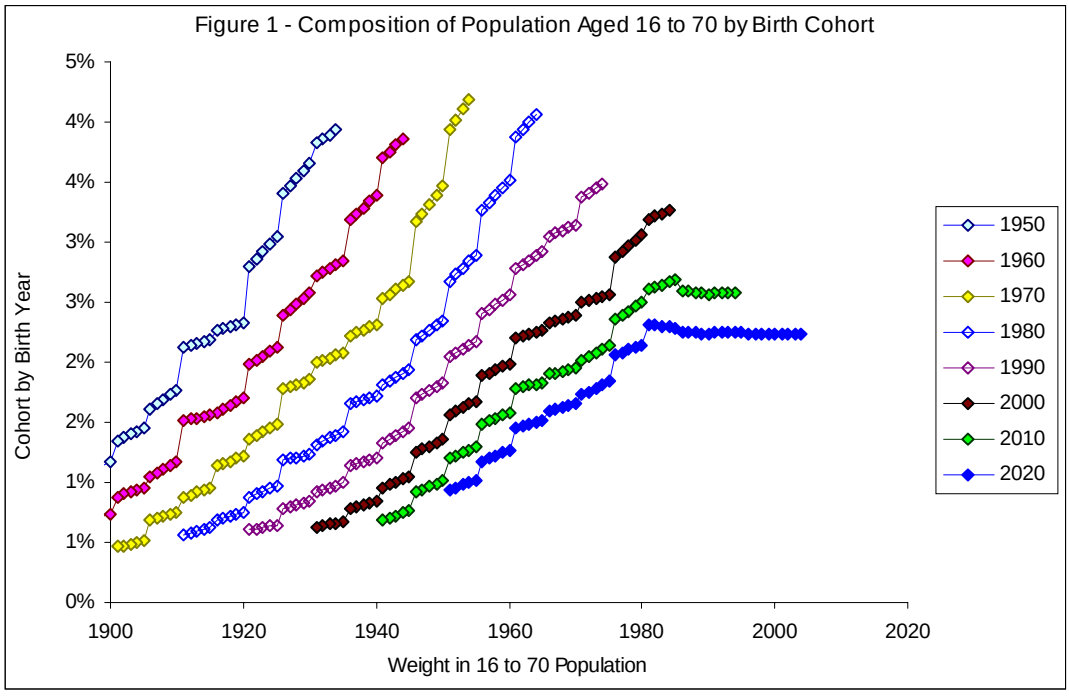
As in other countries, the Brazilian demographic history of the last hundred years can be divided into three periods. The first ranges from the early 1900's to the late

1930's, when birth and death rates were high. However, as mortality balances birthrate, a good part of population growth was due to international immigration. The second period begins after the 1930, when international migration is reduced and the falling mortality together with high fertility became the main reason for rapid population growth. Rates were at their peak, around 2.9% per year, during the decades of 1950 and 1960. The third period begins at the late 1960's, with a rapid fall in fertility rates and, therefore, of the population growth. Mortality keeps falling during the period but its level is not enough to undo the effects of the reduction in fertility. Population growth rates are estimated at 1.3% in the in the late 1990's.

The story above happened in all regions, but not at exactly the same time. Except for migration, the demographic patterns of all regions followed, with some delay, what happened in the Southeast. Furthermore, in the last decades the demographic patterns of all regions became much more homogeneous than before, although we can still identify clear differences among them.

During the 1990's infant mortality has fallen significantly, but not enough to compensate reduced fertility. As a result, during this decade younger cohorts are smaller than their predecessors. Although the pressures for the supply of schooling caused by total population growth are reduced, other factors of pressure such as short distance migration and the increase of school enrollment are still in effect.

The net effect of all these changes is shown on Figure 1. This figure shows the composition of the potential working-age population, namely all those between 16



and 70 years of age in any given calendar year, in terms of cohorts. Up to the year 2000 population, younger cohorts are always more numerous than older ones. From year 2010 onwards, the functions are no longer monotonous and there is a point from which newer cohorts become less numerous.

Our hypothesis is that this inflexion may provide a window of opportunity for education that should not be missed.

3 - METHODOLOGY AND DATA

The methodology to be used in this paper will be the simplest possible capable of providing an answer to the questions on the interplay between the educational level and inequality of each cohort and the educational level and inequality of the population as a whole in a given calendar year. Let:

T index calendar years;

t index cohorts;

S_t be the final average educational level of cohort t;

I_t be a decomposable measure of final educational inequality of cohort t; and

T_t be the weight of cohort t in the population aged 16 to 70 in year T.

Then:

$$S_T = \Sigma_t \, T_t \, S_t \tag{1}$$

(the final educational level of the 16-70 population in year T is a weighted average of the final educational level of each cohort).

$$I_T = \Sigma_t \, f(T_t) \, I_t + \Sigma_t \, W(T_t) \, S_t \tag{2}$$

(the final educational inequality of the 16-70 population in year T is a weighted function of the final educational inequality of each cohort and of its final educational level)**.**

According to Bourguignon (1979), an inequality measure is decomposable if, for any partition[1] of the population, the total inequality can be expressed as a weighted sum of within and between group inequality. In equation (2), I_T represents total inequality in calendar year T, the partition is according to cohorts, and $f(T_t)$ and $W(T_t)$ represent the weights given to between-cohort and within-cohort inequality, respectively.

[1] A partition is a division of the population among groups where each individual belongs to one and no more than one group.

The decomposable inequality measure we decided to use is one of most common: one-half of the squared coefficient of variation. According to Shorrocks (1980), this measure corresponds to the member of the generalized entropy class with an inequality aversion parameter of 2. This inequality measure can be decomposed into within and between components by using the following decomposition weights:

$$f(T_t) = T_t \tag{3}$$

$$W = T_t (S_t / S_T)^2 \tag{4}$$

Henceforth, we will always refer to this inequality measure as I_2.

The methodology consists in simply estimating the final educational level and inequality of each cohort and then running a few simulations to show what impacts educational improvements for different cohorts will have on the 16-70 population for different calendar years.

The data we use are all from the *Pesquisa Nacional por Amostra de Domicílios* (PNAD) from 1977 to 1999. These are surveys covering the whole nation, except for the rural area of the Northern region, where the vast distances make a yearly survey too costly. The PNADs are multipurpose surveys with strong emphasis on labor market results and often carry a supplement covering some specific topic such as health, fertility, child labor, social mobility, among others. The sampling scheme has always been the same — stratified and clustered — but the strata change every time the Census Bureau Grid changes, which happens every 10 years with the national Census. The questionnaire has changed considerably over time, but schooling and age, the only variables important in this study, have been largely spared. That is, the questions regarding schooling and age have not changed much from 1977 to 1999.

The PNAD imposes two shortcomings upon our analysis. The first is that the same people are not followed over time. This means that we do not have real cohorts but pseudo-cohorts. In principle, this should not be a problem, if we believe in the PNAD sampling scheme.

The second problem is that the PNADs exist only from 1977 to 1999 — a mere 22 years. This means that any one cohort was followed only during a part of its evolution.

4 - THE EVOLUTION OF EDUCATION BETWEEN COHORTS

Figure 2 shows educational progress in Brazil. On the horizontal axis is the year of birth of each successive cohort from 1900 to 1983, on the vertical axis is the estimated final average educational level of the cohort. We will explain exactly how this estimate is made later on, but for now what is important is that average education is a monotonically increasing function of cohort date of birth but the

rate of increase is not fixed. Figure 2 shows that the education of each cohort increased at a more or less steady rate until about the 1940 cohort, accelerated for those born between 1940 and 1960, slowed its rate of growth for the 1960's cohorts and then accelerated again for the cohorts born after 1970.

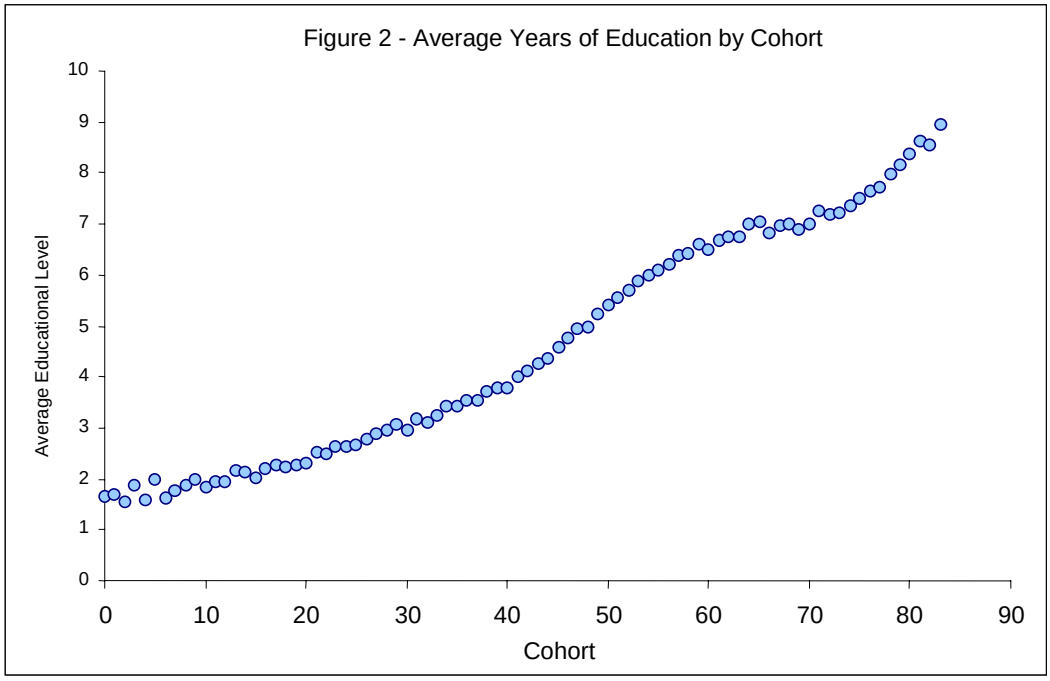
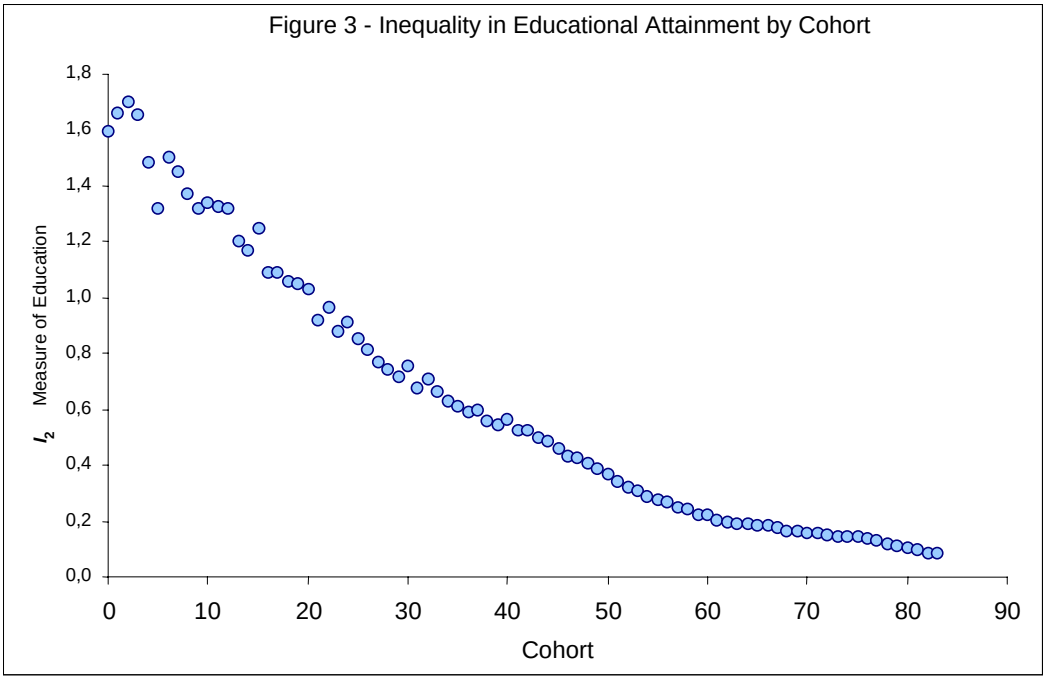


Figure 3 shows the same for the I_2 measure of education of the cohort. Once again, the most important fact is a monotonic relation — each successive cohort



has less internal inequality than the previous one. It is important to note that the I_2 measure does not bear a linear relation to the amount of income inequality explained by education, and this is due to the highly nonlinear returns to education in Brazil.

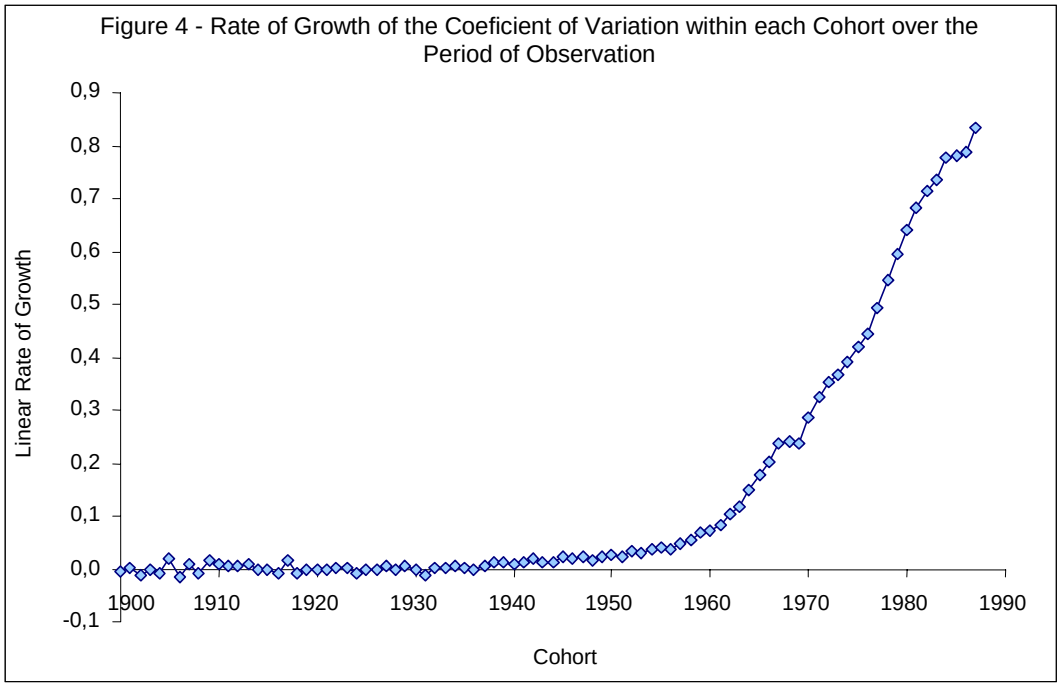
Before going into the interactions between the educational level and educational inequality of each cohort, its weight in the population and the education levels and inequalities of the whole population by year, it is important to understand better how education levels and inequality evolve within each cohort over time.

5 - THE EVOLUTION OF EDUCATION WITHIN EACH COHORT

In principle, all cohorts are born with zero average education, and over time this number increases up to the point at which there is no one in the cohort that still in school and then stabilizes. In Brazil, after 30 years of age, very few people are still in school. In 1999, only 2.9% of the population 30 or over were still in any kind of regular learning. The year of 1999 may be used as an upper bound, given that each successive cohort is completing more education than its predecessors.[2]

Since we observed cohorts from 1977 to 1999, this means that any cohort born previous to 1947 should no longer show any increases in education over the period of observation and even those born previous to 1952 (those 25 and older in 1977) should show very little. This is indeed what we observe, Figure 4 shows the

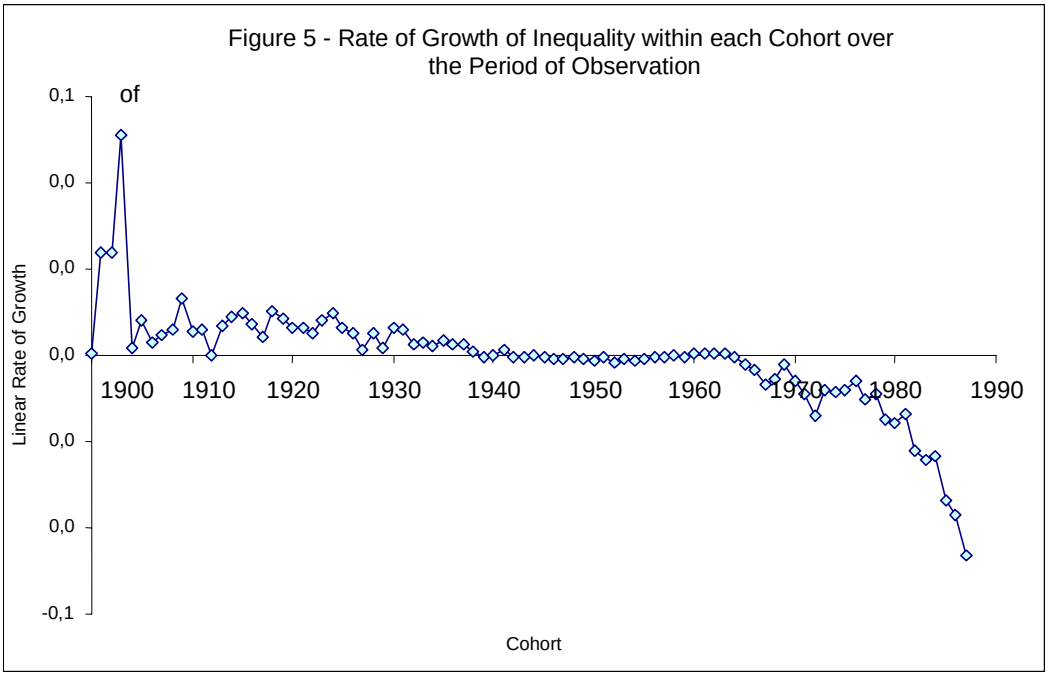
Figure 4 - Rate of Growth of the Coeficient of Variation within each Cohort over the Period of Observation



[2] For example, in 1997 only 2.1% of people 30 or older were involved in education; in 1995, only 1.6%; in 1988, only 1,1%; and in 1986, only 1.0%.

then increases strongly and monotonously up the last cohort observed — the one born in 1987.

A related pattern can be seen in the evolution of inequality. Figure 5 shows that the I_2 measure does not change significantly until the 1965 cohort and then there is a strong downward trend for all successive cohorts. In other words, cohorts aged 12 have already achieved their final inequality, as measured by one half of the coefficient of variation squared. This is not as intuitive as the effect on average education.

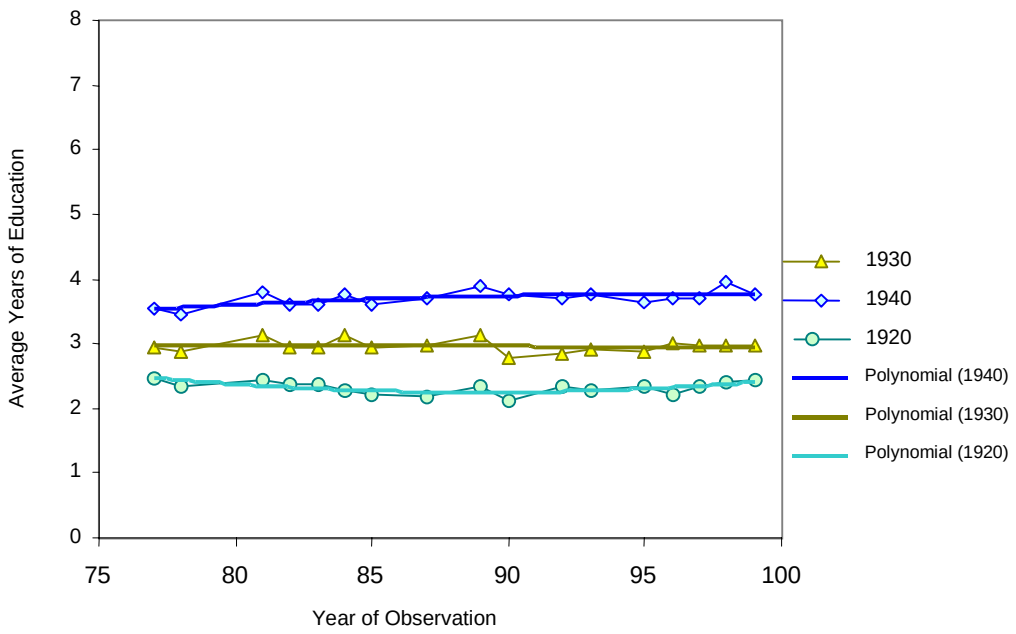


The I_2 measure is one-half of the variance divided by the square of the mean. When each cohort comes into the world none of its members has any education, the mean is zero, and I_2 is not even defined. Once at least one child finished one year of schooling, I_2 becomes defined and then increases very quickly as a part of the cohort acquires some education, yielding a positive denominator, but the numerator, average years of education, remains very low. I_2 then falls mostly because this denominator is increasing. Since we only observe each cohort after it is 10 years old, we do not observe the increasing part of the curve, only the downward part. We will also see later on that it becomes stable within each cohort before average educational level does. This is why we observe changes in the I_2 measure over the 1977-1999 period only with the 1965 cohort while the average changes over the same period for all cohorts after the one born in 1945.

Another way to observe the evolution within cohorts and over time is to look at the average education of each cohort from 1977 to 1999. Figures 6 through 8 show this for 10 cohorts born from 1920 to 1975. Figure 6 shows that the cohorts born in 1920, 1930, and 1940 show no increase at all in average education over the period; Figure 7 shows very slight increases for cohorts born in 1945, 1950, and 1955; and finally Figure 8 shows the large increases in the education of cohorts born from 1960 onwards, whose members were still overwhelmingly in

school during the observation period. Finally, each successive cohort attains a final educational level superior to that of its predecessors.

Figure 6 - Evolution of Cohorts Born before 1940



Figures 9 through 11 show the I_2 measure for the same cohorts as Figures 6 through 8. The message is again clear: the inequality of education is stable from 1977 to 1999 for cohorts born until 1965 but falls over the observation period for those born after 1965. Of course, the final value of I_2 for each cohort is lower than for its predecessor.

Finally, if we shift the curves on Graphs 8 and 11, we can see how the level and inequality in education vary as different cohorts age. This is what is shown on Figures 12 and 13, which may be the most important figures thus far. Let us start with Figure 12.

Figure 12 shows age on the horizontal axis and educational level on the vertical one. Each line represents the evolution of the average years of education of a given cohort. More recent cohorts have curves that lie above those of older cohorts, showing progress in the final educational level achieved. What is not clear is whether the form of these curves changes. Educational improvement can occur in two different ways: either new cohorts can have greater education at any given age previous to leveling off and then stabilize at the same age as previous cohorts or new cohorts can have the same education at each age but take longer to level off.

Figure 7 - Evolution of Cohorts Born before 1955

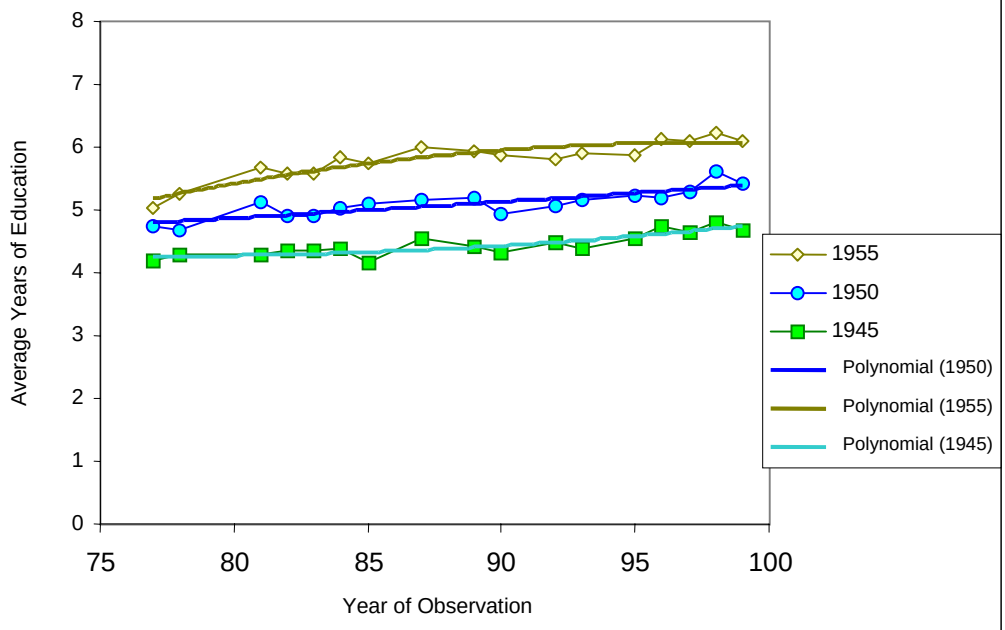


Figure 8 - Evolution of Cohorts Born before 1960

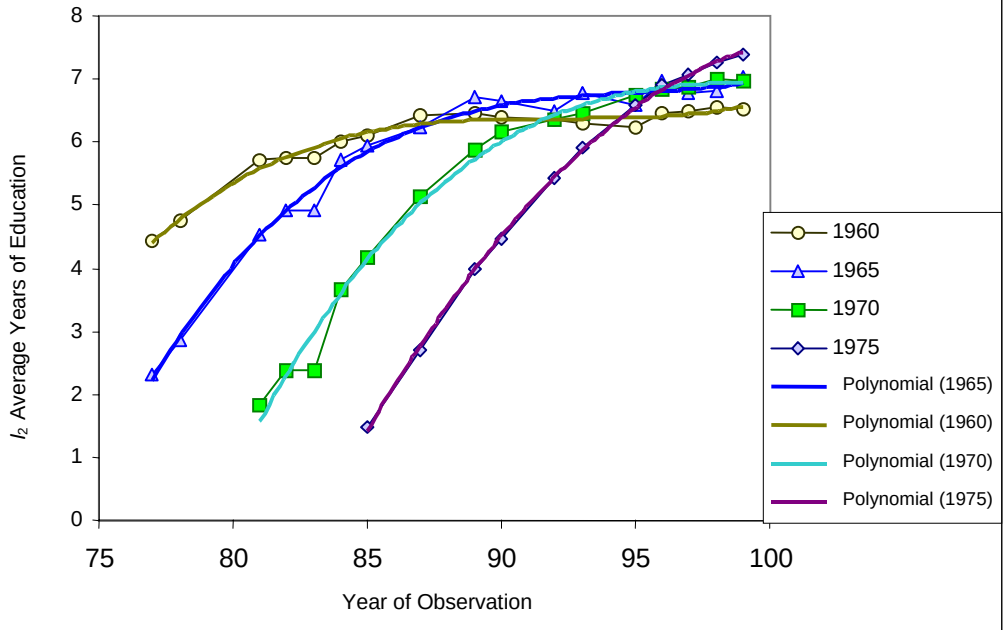


Figure 9 - Evolution of Cohorts Born Before 1940

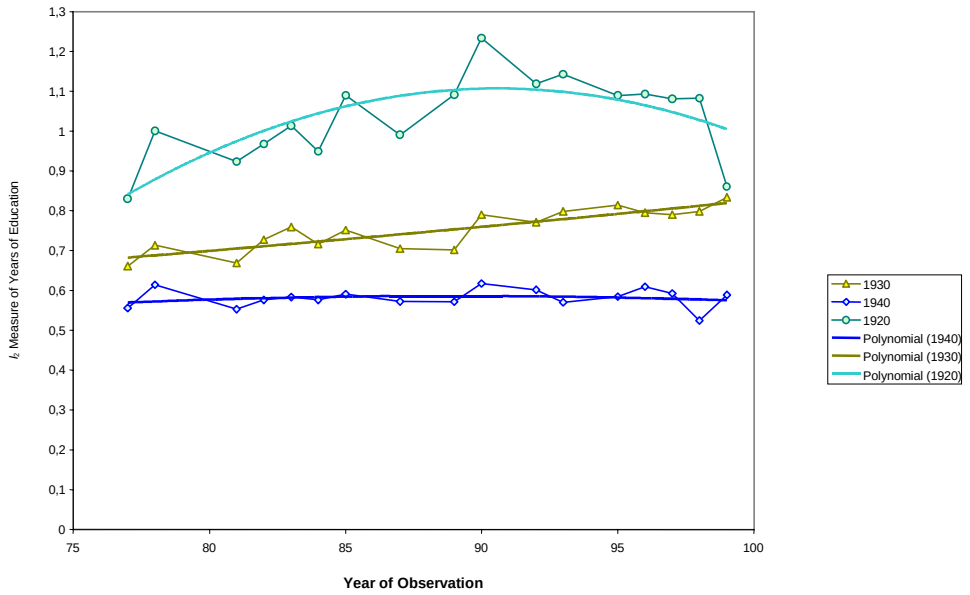


Figure 10 - Evolution of Cohorts Born before 1955

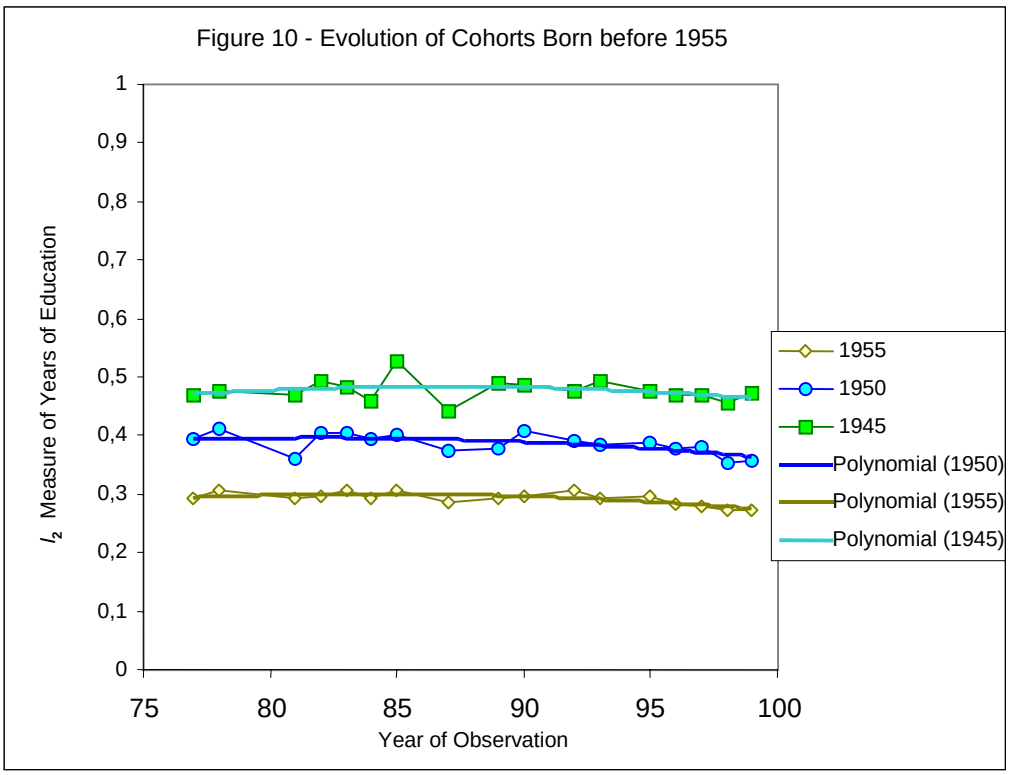


Figure 11 - Evolution of Cohorts Born before 1960

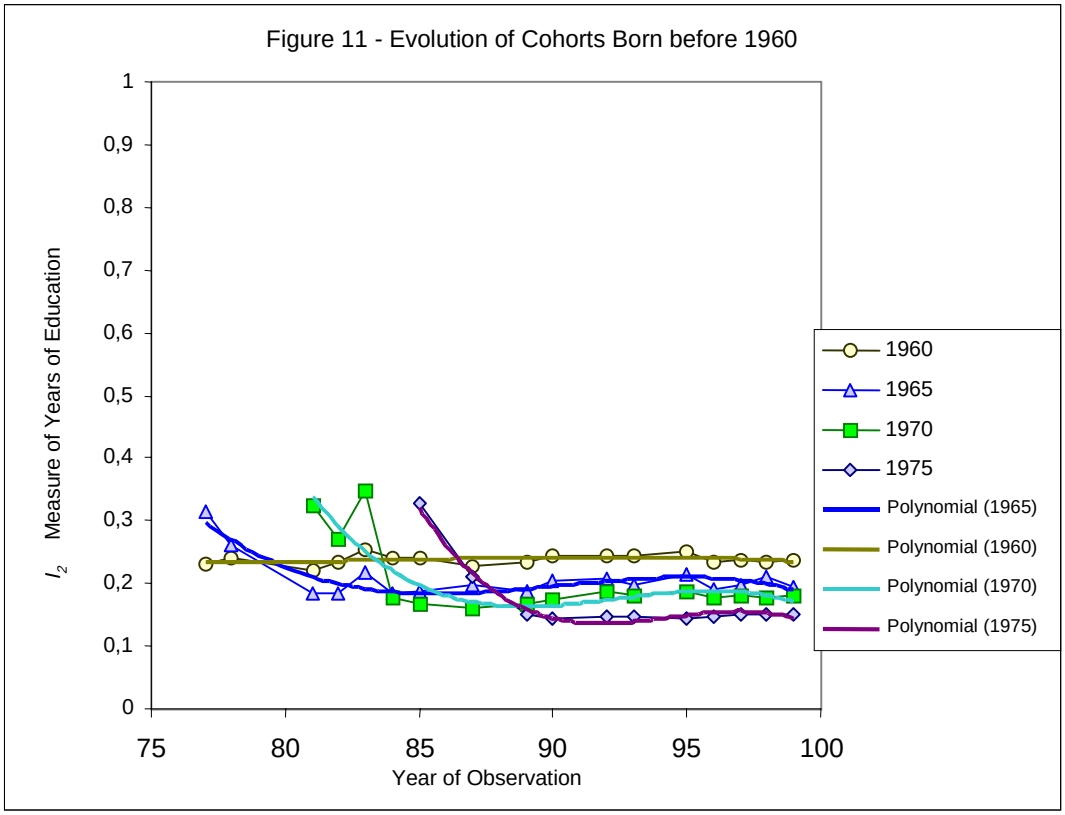


Figure 12 - Evolution of Education Levels of Cohorts Born from 1955 to 1975

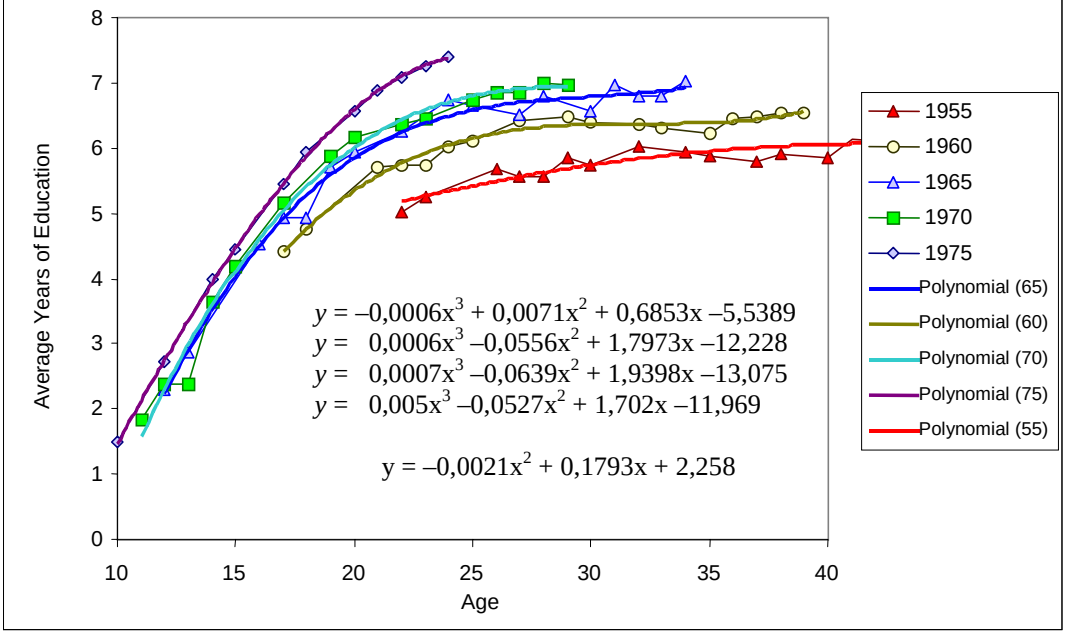


Figure 12 appears to show that successive cohorts, at least those born from 1960 to 1975, have higher education levels at any given age and appear to level off at more or less the same age. This is equivalent to saying that most educational improvement involves advancing further in the educational ladder in the same time rather than staying longer in school. In other words, kids are doing better because they are repeating less. This is coherent with most analysis in the education literature in recent years.

The practical impact of this upon our analysis is on how we will model final educational level of the cohorts born from 75 to 83, whose years of schooling had not yet reached its final value in 1999. What we do is take the 1974 cohort as a baseline and see how much more education a given younger cohort has at each observed age and then attribute to these newer cohorts the 1974 cohort final value multiplied by the average percentage difference between the two over the years of observation.

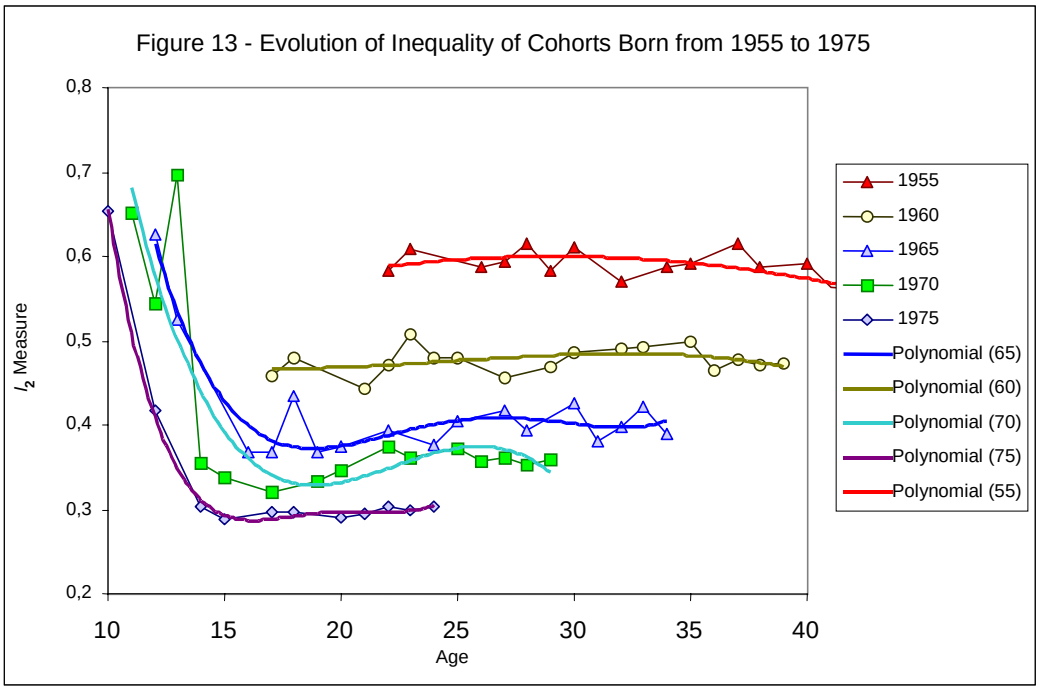


Figure 13 is easier to interpret. By age 20, educational inequality, as measured by the I_2 measure, levels off. This means that additional increases in the variance of education are matched by equal increases in the square of average education, leaving this inequality measure unchanged.

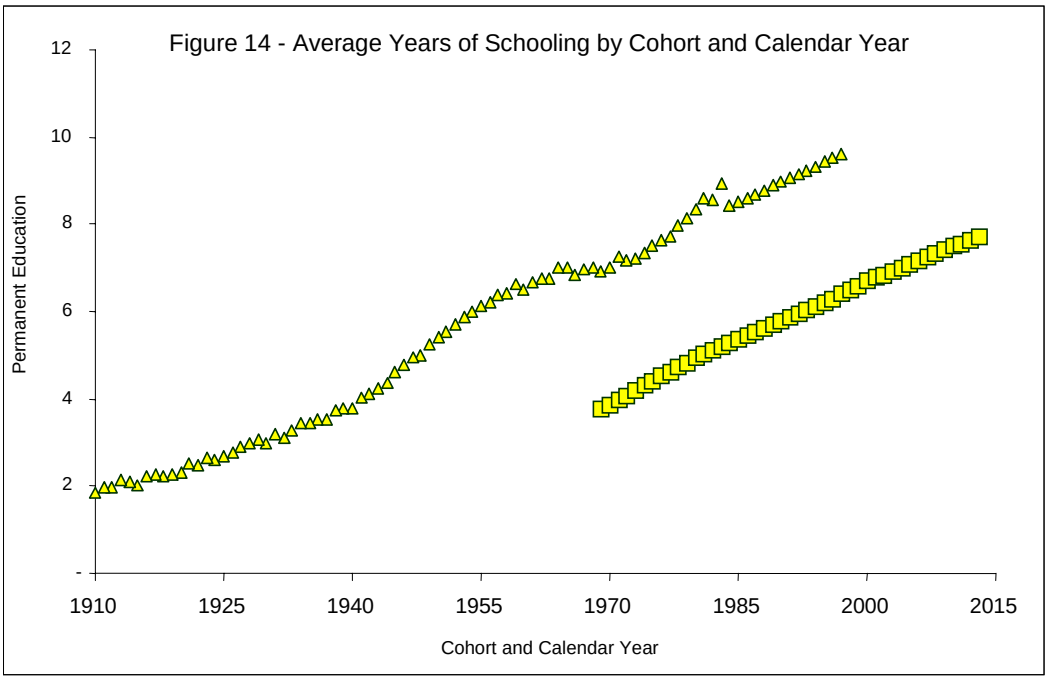
6 - SIMULATION

Figures 14 and 15 show how educational levels and inequality have evolved from one calendar year to the next from 1969 to 1999, in addition to a projection to

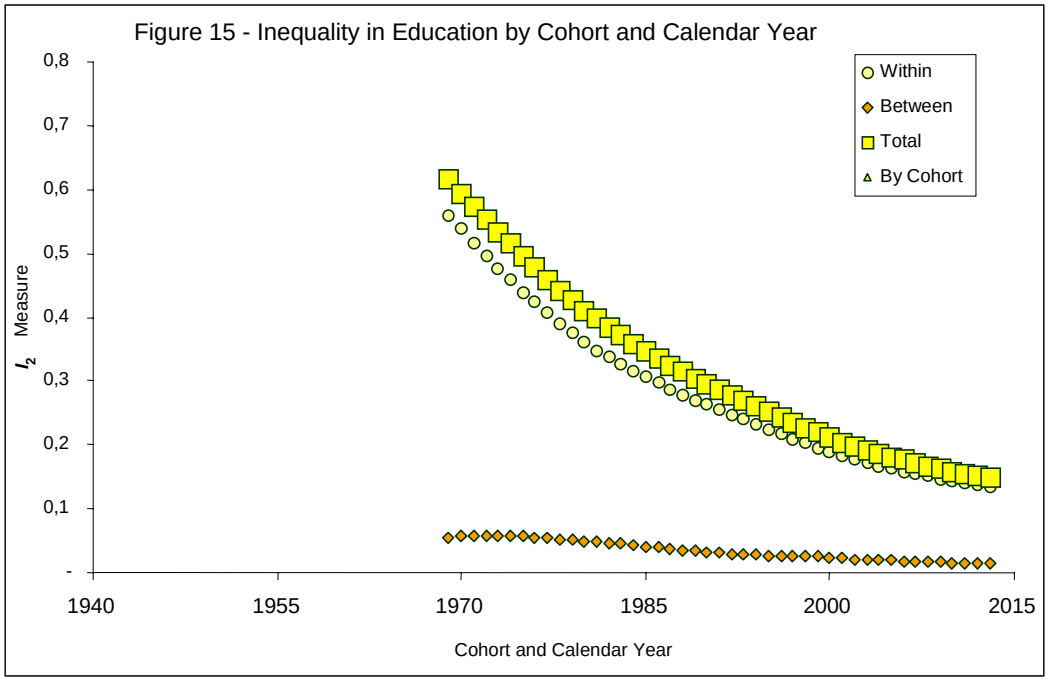
2013.[3] What is shown does not correspond exactly to what is observed in more recent years because those cohorts still increasing their education are imputed their final educational levels and inequality as explained above. For comparison, the education of each cohort is also shown on the same graph. These figures can be thought of as depicting the "permanent education levels and inequality" of people aged 16 to 70 in each year, even if they have not yet achieved these levels. Some interesting things are apparent from these graphs.

The first is that educational levels appear to be increasing in a very monotonous, slow, and linear fashion. This is not really surprising, given the monotonous and slow increase in educational levels of each cohort.

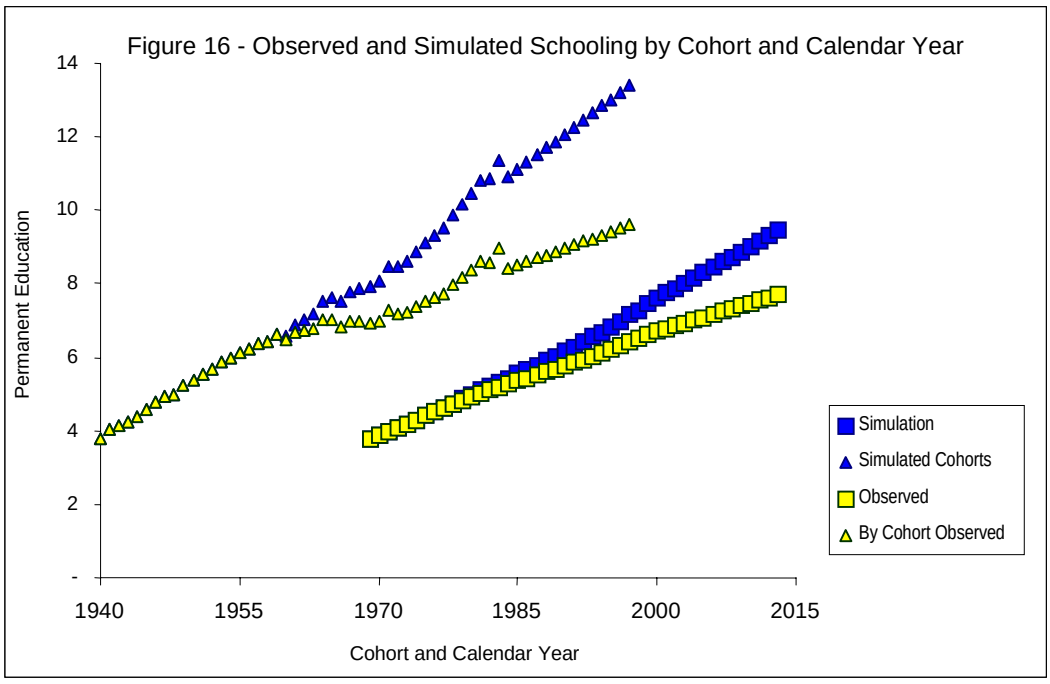
The second interesting fact is that inequality, a measured by the I_2 measure, is falling continuously. This fall is mostly due to within cohort inequality, as between cohort inequality is much smaller.[4] This is surprising, as it is not evident that within cohort inequality dominates total inequality. It is important to note that this fall does not necessarily mean that income inequalities due to education are falling — given Brazil's highly convex returns to education, the two may well go in different directions.

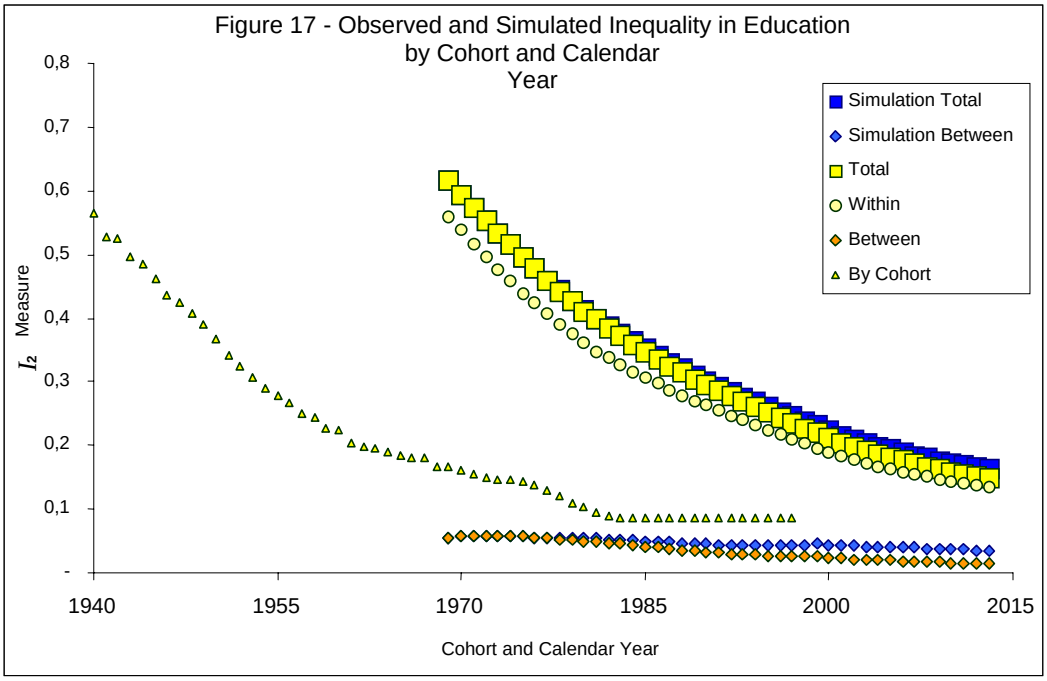


[3] The projection is very simple: the education of cohorts born after 1984 is linearly projected based on those born from 1961 to 1983.
[4] We have assumed constant within cohort inequality for cohorts born after 1987.



Figures 16 and 17 are identical to Figures 14 and 15, except that they show a simulation as well. In this simulation, we increase the education of all cohorts born after 1959 by (t - 1959)/10 years, where t is the cohort's year of birth. The final impact on average education of this very large increase in education of each is an increase of 1.74 years in 2013. On the other hand, the increase in inequality is very small — in 2013, the I_2 measure would be 0.162, and not 0.148, due to the smaller rate of reduction in between cohort inequality.

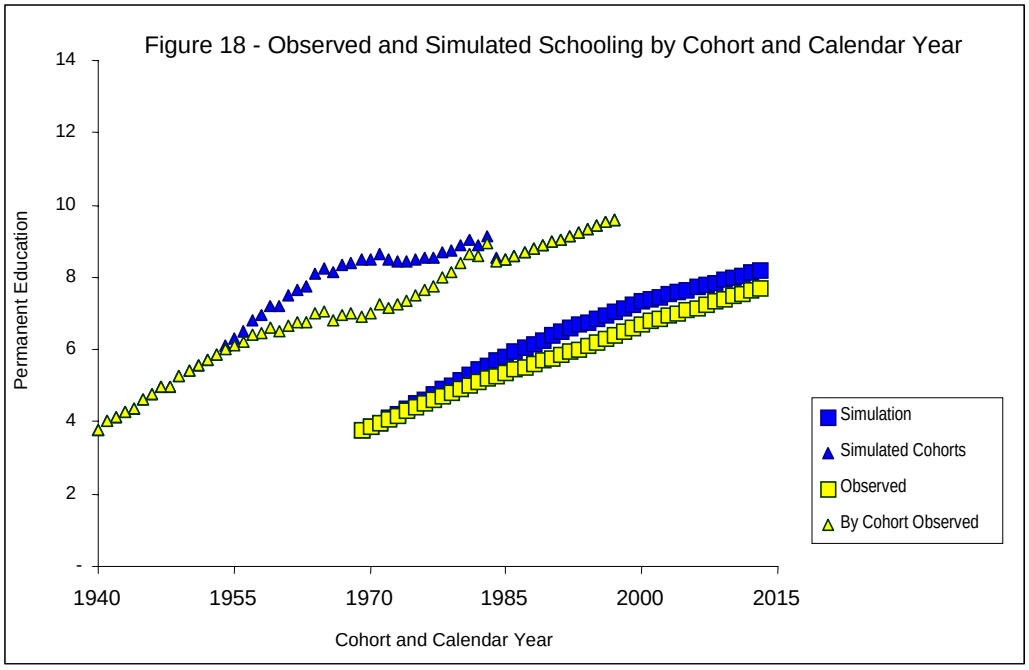


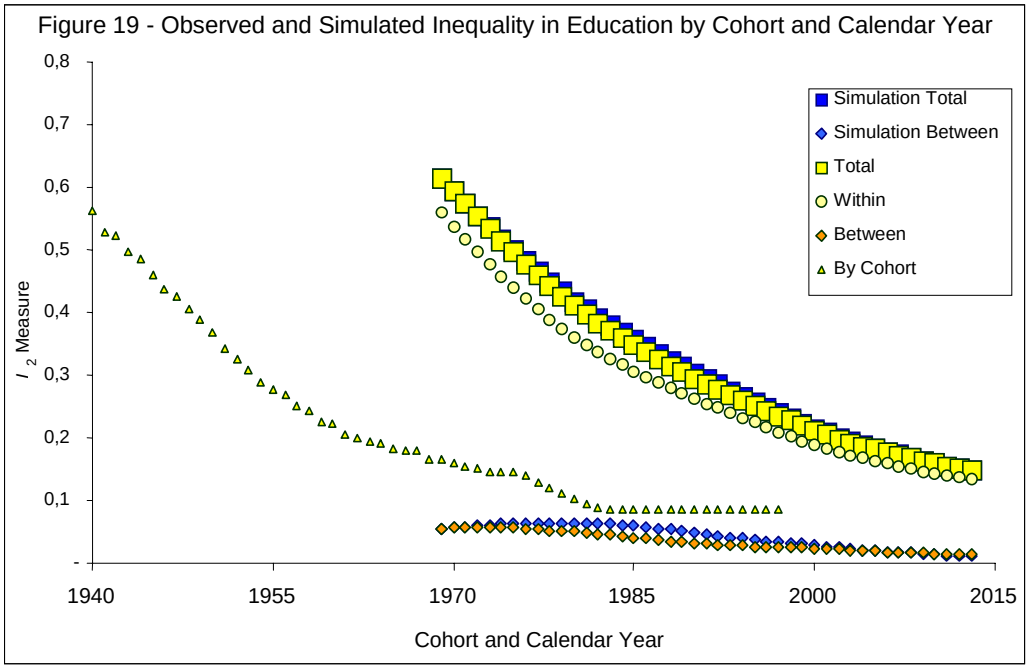


The fact that even a very large intervention, operated on more recent cohorts, does not significantly affect educational inequality is due to the dominant effect of within cohort inequality. Within inequality, as measured by I_2, is not independent from between cohort inequality because average cohort education composes the weights, but this effect is almost insignificant. On the other hand, average income increases, but relatively slowly, given the dramatic nature of the simulation at the cohort level. Both of these effects suggest the existence of strong demographic inertia.

The simulation above supposes very strong and incremental improvements in the educational system. Another simulation we ran is if the changes that improved the educational system in the 1990s had happened in the one or two decades before and are shown in Figures 18 and 19. The intervention was to increase education of cohorts from 1954 to 1983 in an inverted U form so that the 1983 cohort and those after follow the path they would follow anyway.

Since the transformations are much less dramatic than those in the first simulation, the impact on average education is also much less dramatic. The educational level of the 16 to 70 population in the final calendar year we look at, 2013, rises from 7.7 years to 8.1. The effects on inequality are also not very impressive: there is a small increase as the cohorts whose education was increased come into adult age, but it wears out by the year 2000.





7 - CONCLUSIONS

The main conclusion is that demographic inertia is strong in Brazil and improvements in the school system today will take long to translate into more education for the population as a whole. Between cohort inequality will always increase because increasing the educational level of younger cohorts is equivalent to giving more education to those cohorts that already have the most. On the

other hand, total inequality would relatively unaffected because within cohort inequality of older cohorts dominates total educational inequality within the population.

In response to the question raised in the title, there appears to be a window of opportunity, some of which has already been lost. However, given the slow nature of demographic transformations, this window is really very wide and there is still time to take advantage of it now.

BIBLIOGRAPHY

BOURGUIGNON, F. Decomposable income inequality measures. *Econometrica,* v. 47, n. 4, Jul. 1979.

DEATON, A., PAXSON, C. The effects of economic and population growth on national saving and inequality. *Demography*, v. 34, n. 1, Feb. 1997.

SHORROCKS, A. F. The class of additively decomposable inequality measures. *Econometrica*, v. 48, n. 3, Apr. 1980.