TEXTO PARA DISCUSSÃO

No. 478

Inequality of outcomes and inequality of opportunities in Brazil

François Bourguignon
Francisco H.G. Ferreira
Marta Menéndez

DEPARTAMENTO DE ECONOMIA
www.econ.puc-rio.br

DEPARTAMENTO DE ECONOMIA
PUC-RIO

TEXTO PARA DISCUSSÃO
N^{o}. 478

INEQUALITY OF OUTCOMES AND INEQUALITY OF OPPORTUNITIES IN BRAZIL

FRANÇOIS BOURGUIGNON
FRANCISCO H.G. FERREIRA
MARTA MENÉNDEZ

OUTUBRO 2003

This draft: October 2003.
Comments welcome.

Inequality of Outcomes and Inequality of Opportunities in Brazil

François Bourguignon, Francisco H.G. Ferreira and Marta Menéndez*

Keywords: Inequality of Opportunities, Intergenerational Educational Mobility

JEL Codes: D31, D63, J62

Abstract: This paper departs from John Roemer's theory of equality of opportunities. We seek to determine what part of observed outcome inequality may be attributed to differences in observed 'circumstances', including family background, and what part is due to 'personal efforts'. We use a micro-econometric technique to simulate what the distribution of outcomes would look like if 'circumstances' were the same for everybody. This technique is applied to Brazilian data from the 1996 household survey, both for earnings and for household incomes. It is shown that observed circumstances are a major source of outcome inequality in Brazil, probably more so than in other countries for which information is available. Nevertheless, the level of inequality *after* observed circumstances are equalized remains very high in Brazil.

* Respectively: World Bank; World Bank and PUC-Rio; and Delta, Paris. Correspondence to Francisco H.G. Ferreira, The World Bank, 1818 H. St. NW, Washington, DC, 20433, USA. Fax: (202) 522-1151. E-mail: fferreira@worldbank.org. The views expressed in this paper are solely those of the authors, and should not be attributed to the World Bank or its Board of Directors. We would like to thank Tony Atkinson, Richard Blundell, Sam Bowles, Jishnu Das, Quy-Toan Do, Yoko Kijima, Costas Meghir, Thomas Piketty and Martin Ravallion for helpful comments. All remaining errors are ours.

Introduction

Inequality of "outcomes" and inequality of "opportunities" have long been associated with very different views on social justice, in the literature on economic inequality. The first of these concepts refers to the distribution of the joint product of the *efforts* of a person and the particular *circumstances* under which this effort is made. It is mostly concerned with income inequality. The second concept refers to the heterogeneity in those circumstances that lie beyond the control of the individual, but that nevertheless significantly affect the results of his efforts, and possibly the levels of those efforts themselves. This distinction, the formulation of which is borrowed from Roemer (1998), building on earlier work by John Rawls, Amartya Sen and others, is well illustrated by the standard opposition between *inequality* and *mobility*. The US is often presented as more unequal than European societies but at the same time more mobile from a generation to the next, a feature which is sometimes taken as the sign of a more equal distribution of chances or opportunities in the US. [1]

Despite the obvious importance of the concept of inequality of opportunities, limited empirical work has been done in this area, in comparison with the huge literature on the inequality of outcomes.[2] The main reasons for this scarcity are probably: (i) the conceptual difficulty of separating out ‘circumstances’ and ‘efforts’; and (ii) the limited availability of variables that could satisfactorily describe ‘circumstances’. Both of these problems are still more acute in developing countries. Yet, knowing what part of observed outcome inequality may be attributed to circumstances, and in particular to family background, is as important there as in richer nations. Such knowledge should help define the actual scope for redistribution policies and, in particular, it should inform the choice between redistributing current income or expanding the opportunity sets of the poor, for instance through making the accumulation of human capital among children less dependent on their parents.

[1] For comparisons of mobility between the US and European countries, see Burkhauser et al. (1998), or Checchi et al. (1999).
[2] By contrast, social mobility has always been a leading theme of the sociological literature. However, it is not clear whether that literature translates easily into standard economic inequality concepts.

These issues are particularly relevant in the case of Brazil, because of its high level of (outcome) inequality. To what extent is this high inequality due to very unequal opportunities that individuals inherit from their parents and to what extent is it the result of some heterogeneity in their efforts, or in the returns to these efforts? Focusing on human capital and education, a natural way of answering this question consists of studying the 'demand for schooling' or, in other words, how much parents invest in their children, conditionally on their own characteristics. That part of schooling inequality which is explained by the characteristics of parents is then taken to quantify the inequality of opportunities, whereas the remainder is attributed to heterogeneous individual efforts. This approach has been followed by Behrman, Birdsall and Székely (2000) for Latin American countries. Barros and Lam (1993), and Lam (1999) applied similar methods to Brazil. Given the nature of their data sets, however, these studies focused on expected *future* mobility or, in other words, the relation between the *potential* earnings of children when they become adults, and the earnings of their parents. The approach was not suitable for disentangling the *actual* share of the inequality of opportunities in today's levels of overall (outcome) inequality.

We follow a different strategy, based on direct information given by survey respondents about the education and occupational position of their parents, as available in the 1996 Brazilian household survey (PNAD). It turns out that this information permits measuring not only the extent of intergenerational educational mobility but also the way in which parents' characteristics, and some other circumstance variables, affect the earnings or income of their children, independently of the education of the children. By controlling for the year of birth, it is also possible to see how this influence of parents and social background has changed across cohorts and whether opportunities account for an increasing or decreasing proportion of total inequality. Although we focus on parental education as a central determinant of opportunity, there are other dimensions in the space of opportunities, which can be captured by some of the variables in our data set. These include parental occupation, race, and the region of origin.

The analysis reveals a sizeable inequality of opportunities in Brazil. First, parental education proves to be a powerful independent determinant of individual earnings or income, even when controlling for the individual's own schooling. In addition, parental education is a strong predictor of own schooling. It turns out that, for older cohorts, the distribution of schooling is almost fully reproduced – up to some increase in average schooling – across generations. At the same time, the analysis also shows that the inequality that would remain among people after controlling for the inequality of all observed circumstances - parents' education, occupation, race and region of origin - is still very high: higher than the total inequality of outcomes observed in many countries in the world. This may be because: (i) important circumstances that shape the economic opportunities people face are simply unobserved or unobservable; and/or (ii) the functioning of the Brazilian economy is generates high inequality in individual efforts and/or in the returns to those efforts.

The paper is organized as follows. Section 1 briefly discusses the theoretical background for the estimation work undertaken in the paper, in terms of the general relationship between inequality of outcomes and inequality of opportunities. Section 2 discusses estimates of the effect of circumstance variables on both individual schooling and earning outcomes. Section 3 then presents a range of measures of the influence of the inequality of observed opportunities on the inequality of individual earnings. Section 4 generalizes this analysis to the case of household per capita incomes. This allows us to account for the role that circumstances may play in welfare inequality through labor-force participation behavior, assortative mating and fertility. The concluding section places the preceding results in an international perspective and draws some simple implications of these results for our understanding of anti-inequality policy in Brazil.

1. Conceptual background.

Among the determinants of the earnings of an active individual at any point in time, one can find some personal characteristics which are independent of the individual's will. Following Roemer (1998), we call such characteristics "circumstances", and contrast them to characteristics that, on the contrary, reflect "efforts" made by the individual to increase his/her productivity and earnings. Let C denote the set of "circumstance" variables and E denote the set of "effort" variables. C typically includes fixed socio-demographic attributes like race, sex, region of origin, and the individual's family background. E corresponds essentially to the human capital accumulated by the individual once he is free to make decisions for himself. This may include the last part of formal schooling, but also on-the-job training, past decisions to change job or region of residence, or current effort at work.

One could then represent the most general form of this relationship between earnings, efforts and circumstances as $w_i = f(C_i, E_i)$. A first, possibly naïve, attempt to estimate such a relationship empirically would be to linearize the model as follows:

$$Ln(w_i) = C_i.\alpha + E_i.\beta + u_i \qquad (1)$$

where w_i denotes current hourly earnings, α and β are two vectors of coefficients and u_i is a residual term that accounts for unobserved circumstance – including sheer luck - and effort variables, measurement error, and transitory departures from the permanent level of earnings. All of these factors are assumed to be independent of the variables included in C and E. They are also assumed to have zero mean and to be identically and independently distributed across individuals.

If inequality were to be measured by the variance of the logarithm of earnings, and if one could reasonably assume that circumstances and efforts are mutually independent, then the following simple decomposition of total inequality would hold:

$$v(Ln\ w) = \alpha'.V(C).\alpha + \beta'V(E)\beta + v(u) \qquad (2)$$

where $v(\ )$ stands for variances and $V(\ )$ for variance-covariance matrices. In other words, total inequality could be explained simply as a weighted sum of the inequality of observed opportunities (first term on the RHS), the inequality of observed efforts (second term) and the inequality due to unobserved earning determinants.

But the formulation in (1) suffers from a number of obvious shortcomings. First, to assume complete additive separability between circumstances and efforts – and independence between all the unobservable and observable earning determinants - seems hopelessly inadequate. It seems natural, for instance, to allow for the possibility that efforts may themselves be partly determined by circumstances, so that $w_i = f(C_i, E_i(C_i))$. Formal schooling, for example, may be determined at least in part by family background. This effect of parental background on the educational outcomes of the next generation may occur because more educated parents provide more "home inputs" into an "education production function", such as books, vocabulary and quality time spent on homework, but it may also reflect individual learning about the returns to effort, which may themselves depend on the circumstances – and indeed on the previous mobility history – of the family.[3] Assuming, more modestly, that only unobserved effort determinants are orthogonal to observed circumstances, we can write a system of equations for efforts as:

$$E_i = C_i.b + v_i \qquad (3)$$

where b is a matrix of coefficients and v_i stands for a vector of unobserved effort determinants – one component for each component of the vector E_i. As usual the v_i's are supposed to be i.i.d. across individuals and with zero mean. Substituting (3) into (1) yields:

$$Ln(w_i) = C_i.(\alpha + \beta.b) + v_i\,\beta + u_i \qquad (4)$$

This reduced-form expression shows that, if the world were reasonably represented by a model consisting of equation (1) and the system (3), circumstances have a double effect on wages. They affect it directly, for given efforts, through the set of coefficients α. They also affect it indirectly through their influence on efforts, the size of this second effect being given by the scalar product $\beta.b$.[4] This restatement of the original model modifies the variance decomposition in equation (2) and, more generally, any decomposition of the distribution of individual wages into components associated with observed circumstances and efforts.

[3] Hanushek (1986) reviews the literature on the first, 'production function' view of the impact of family characteristics on individual schooling outcomes. Piketty (1995) models the second type of channel, where individual beliefs about opportunities and mobility are formed on the basis of the experience of their own lineages, and in turn rationally determine their own effort levels. Our empirical model allows for these indirect effects of circumstances (and past history) on welfare *through* efforts. It does not distinguish between the alternative channels, but it is consistent with both.

[4] For a clear presentation of this distinction between direct and indirect effects of circumstances on current earnings see Bowles and Gintis (2002).

If one were interested only on the total effect of circumstances on outcomes (measured by earnings), then an estimation of the reduced-form equation (4) would suffice. It would be impossible to identify the structural parameters α, $\boldsymbol{\beta}$ and $\boldsymbol{b}$ from the estimated coefficients, but it would be possible to treat those coefficients as capturing the combined effects of observed circumstances on outcomes. Indeed, the estimates of the complete effects of equalizing observed opportunities reported in section 3, and drawn in Figures 1-4, correspond to those which would be obtained from a direct OLS estimation of equation (4). But the distinction between direct and indirect effects of circumstances on welfare may matter, not only because they are of intrinsic interest, but also because they have sharply different implications for policy.

The intrinsic interest of identification of the structural model arises largely from the (closely related) literature on intergenerational mobility, which has long struggled to understand the *nature of the mechanisms* through which parental education affects the earnings of children. Particular interest has been attached to the relative magnitudes of the impact of the parents' education on the child's education, vis-à-vis the impact of the parents' education on the child's labor earnings *conditional* on own schooling. Given the relative importance of parental schooling within our set of circumstance variables, as will be shown below, our identification of the structural parameters in (1)-(3) has a bearing on that debate.

In addition, policymakers interested in reducing inequality of opportunities would benefit from knowing whether the effect of parental education on child's earnings operates predominantly through the direct effect (in which case, factors such as social network effects in access to employment and the importance of family wealth in establishing one's own business would be paramount) or largely through the effect on education (in which case the central policy challenge will lie in broadening access to continued schooling opportunities to children from disadvantaged backgrounds, through programs like cash transfers conditional on school attendance, or special after-class help for lagging students). Although a consideration of these specific policies lies outside the scope of this paper, the policy relevance of distinguishing direct and indirect effects of circumstances on outcomes lies behind our attempt to identify the structural model, rather than being satisfied with running an OLS regression of model (4).[5]

Such a decomposition of the impact of circumstances into direct and indirect effects would only make sense, however, if one could find unbiased estimates of the various sets of coefficients: α, β and b. Running OLS on equations (1) and (3) is unlikely to yield such estimates. In particular, the required assumption that u in equation (1) is orthogonal to C and E may be open to doubt. The problem is probably less serious for the circumstance variables. One may not be so much interested in the 'true' effect of the variables included in C but in their overall impact once their correlation with unobservable circumstances are taken into account. This overall impact will be truthfully described by the OLS estimate of α. In other words, if parental wealth or income is not observed, the estimate of α will account for the effect on children's earnings of both the variables in C and the effect of that part of parental wealth or income which is correlated with elements of C.

The real problem arises when unobservables (u) in the earnings equation cannot be assumed to be independent of the effort variables (E). Again, imagine that parental wealth has a direct impact on either the schooling or the current earnings of their children (or both), independently of the child's own education. This correlation between u and E, or u and v, introduces bias in the estimation of the (α,β) coefficients and therefore in the decomposition of the total inequality into circumstance and effort components.

One way out of this difficulty would be to observe instrumental variables, Z that would influence efforts but not earnings. Equation (3) would then be replaced by:

$$E_i = C_i.b + Z_i.d + v_i \qquad (5)$$

[5] Even the structural model (1)-(3) is still somewhat restrictive. Like any parametric model, it relies on certain maintained functional form assumptions. Additionally, we chose to be parsimonious in our specification, and have omitted various possible interaction terms between circumstance and effort variables, which might have allowed us to estimate different rates of returns to efforts for individuals in different circumstances.

with the vector Z_i being orthogonal to u_i. Then instrumenting the effort variables in (1) through (5) would yield an unbiased estimator of (α,β) and then an unbiased decomposition of total inequality into inequality of observed opportunities, or circumstances, and inequality of efforts. Models of this type have been extensively used in the literature on returns to education. In the standard Mincerian equation, for instance, instrumenting education by family background is standard practice to correct for the endogeneity of education. But if family background is an independent determinant of earnings in its own right, and indeed one is interested in separating out the impacts of the instrument and of the variable it is instrumenting for, as is presently the case, then some other instrument is required. Ability tests taken while attending school have sometimes been used. But that information is seldom available, particularly in developing countries.[6]

In the absence of an adequate set of instrumental variables Z, the only solution is to explore the likely effect of the potential bias in the estimation of β due to the correlation between u and v, and then to decide on that basis what is the most reasonable range of estimates. This is the approach we adopt in this paper, relying on the parametric bounds analysis which is described in detail in the appendix.

Once this method yields an interval for the unbiased estimates of the parameters α, β and b, it simply remains to define our measure of inequality of opportunities, and its contribution to the distribution of current earnings. An appealing way of measuring that contribution consists of evaluating (through the system of equations (1) and (3)) the counterfactual distribution of earnings which would attain if all the inequality due to the circumstance variables had been eliminated. This can be done in two steps. First, one can derive from OLS estimates of equation (1) above what would be the distribution of earnings if circumstances had been equalized for all individuals. The remaining inequality would thus be due essentially to differences in efforts and in the residual term, u_i. It is then possible to define hypothetical individual earnings:

$$Ln\tilde{w}_i = \overline{C}.\hat{\alpha} + E_i.\hat{\beta} + \hat{u}_i \qquad (6)$$

[6] Early contributions include Bowles (1972), Griliches (1972), Behrman and Taubman (1976). For a survey of all models of returns to schooling based on this kind of instrumentation see Card (2001).

where $\overline{C}$ stands for the mean circumstances in the population being studied and the notation ^ refers to OLS estimates of equation (1). Comparing that hypothetical distribution $\{ \tilde{w}_i , i = 1, 2, ..N \}$ with the actual distribution $\{w_i, i = 1, 2, ..N \}$ allows us to evaluate the distributional effect of the inequality of observed opportunities in *C*.

Of course, the preceding definition would be justified on normative grounds only if observed efforts in *E* and unobserved efforts in *u*, were solely the reflection of individual preferences and were unaffected by *C*. But, if efforts also depend on circumstances, the preceding evaluation of the effect of the inequality of opportunities is only partial. If all circumstances were included in *C*, then it is the residual, *v*, of equation (3) that would describe the effect of the heterogeneity of individual preferences on observed efforts. Thus, equalizing circumstances would lead to a partial equalizing of efforts that would in turn contribute to an additional equalizing of individual earnings. Formally, this would lead to a distribution of earnings $\{ \tilde{\tilde{w}}_i , i = 1, 2, ..N \}$ defined by :

$$Ln\tilde{\tilde{w}}_i = \overline{C}.(\hat{\alpha} + \hat{b}\hat{\beta}) + \hat{v}_i\hat{\beta} + \hat{u}_i \qquad (7)$$

where $\hat{b}$ and $\hat{v}_i$ are OLS estimates of equation (3).

The comparison of the actual distribution $\{w_i, i = 1, 2, ..N\}$ and the distribution $\{ \tilde{\tilde{w}}_i , i = 1, 2, ..N \}$ fits the definition that Roemer (1998) gives of the inequality of opportunities, even though in a very simple way.[7] It shows the *complete* effect of observed circumstances on the distribution of earnings. By contrast, the comparison of the actual distribution with the distribution $\{ \tilde{w}_i , i = 1, 2, ..N \}$ is only *partial*. Comparing the two approaches allows us to distinguish the part of the inequality of opportunities that goes through the direct effect of circumstances on outcomes, from

[7] Actually, interpreting Roemer's theory of the equality of opportunities literally would lead to measure the inequality of opportunities by comparing earnings for each percentile of the distribution of the residual term, v_i , and then aggregating across percentiles. In the present case, this would be equivalent to comparing the distribution obtained by equalizing the residual terms u_i and v_i on the one hand, with perfect equality, on the other. Equalizing circumstances as we do here is in some sense complementary of that approach and offers the advantage of using the actual distribution of earnings rather than equality as a benchmark.

the part that goes through the effect of circumstances on efforts. This is why we report results from both the partial and complete estimates below.

Before closing this conceptual section, we note that when circumstance variables include parental characteristics, it is natural to think of the preceding analysis as being related to intergenerational mobility. A direct measure of *income* mobility would be provided by the preceding model if parents' income was among the variables C. But other types of mobility may be behind equation (3). For instance, if parental education is included in C and individuals' schooling is in E, then part of system (3) actually describes intergenerational educational mobility. The coefficient b that relates the education level of children to that of parents is an (inverse) indicator of that mobility. For instance, if education is measured in number of years of schooling for both parents and children, then the extent to which b is less than unity would describe how fast differences in education tend to systematically lessen across generations.[8] It can be seen on equation (4) that the degree of intergenerational educational persistence, b, partly determines the share of current earnings inequality which is due to differences in circumstances or opportunities.

2. Earnings and educational mobility in Brazil

This method is now applied to Brazilian data. This section first describes the data and the nature of the variables being used. It then discusses the various estimates obtained both for the earnings equations and for the efforts equations. The latter can also be seen as describing intergenerational educational mobility in Brazil.

Data and variables

Data are from the 1996 wave of the Pesquisa Nacional por Amostragem a Domicilio (PNAD), the Brazilian Household Survey. For that year, information is available on the education and

[8] Behrman, Birdsall and Szekely (2000) define intergenerational educational mobility as the share of the residual in the variance of schooling in equation (3) rather than as the value of $1-b$. The two concepts are equivalent only when the distribution of schooling has the same variance for parents and children.

occupation of the parents of all surveyed household heads and spouses.[9] The analysis is restricted to urban areas because of the general imprecision of earnings and income measurement in rural areas.[10] It is also restricted to individuals 26 to 60 years old, in an effort to concentrate on individuals having finished schooling and potentially active in the labor market.

The analysis described in the preceding section is conducted on 5-year cohorts - from individuals born between 1936-40 up to those born between 1966-70. This permits not only measuring the role of the inequality of opportunities in shaping the inequality of observed earnings at a point in time, but also to study how this role may have changed over time. An important question is indeed whether the increase in the educational level of successive cohorts was accompanied by more or less inequality of opportunities or whether it corresponded to a uniform upward shift in schooling achievements, with constant inequality of opportunities. Comparing various cohorts observed at a single point in time allows us to shed some light on this issue.

The analysis first focuses on individual earnings, measured as "real hourly earnings from all occupations". In a second stage, the analysis will be conducted for welfare levels, measured by household income per capita. This will allow us to discuss the roles of labor supply behavior and fertility as additional channels through which the inequality of opportunities affects the inequality of outcomes.

The vector of circumstance variables, *C*, includes race dummy variables, parental education expressed as numbers of years of schooling [11], the occupational position of the father[12] - a 9-level categorical variable- and dummies for the regions of origin. The vector of effort variables, *E*, is

[9] The same information is available in both the 1982 and the 1988 surveys. Earnings in the 1982 survey were collected with respect to a different reference period (of three months) and are therefore not comparable with other years. The 1988 survey was used to check the robustness of some of the results reported in this paper.

[10] See Ferreira, Lanjouw and Neri (2003) for a discussion of the shortcomings of rural income data from the PNAD.

[11] Parental education is given in discrete levels. They were converted into years of schooling (here in brackets) using the following rule. No school or incomplete 1st grade (0); incomplete elementary (2); complete elementary, or complete 4th grade (4); incomplete 1st cycle of secondary or 5th to 7th grade (6); complete 1st cycle of secondary or complete 8th grade (8); incomplete 2nd cycle (9.5); complete 2nd cycle of secondary (11); incomplete superior (13); complete superior (15); master or doctorate (17).

[12] The 9-level occupational categories are: rural workers (1); domestic servants (2); traditional sector workers (3); service sector workers (4); modern industry workers (5); self-employed shopkeepers (6); technicians, artists and desk workers (7); employers (8); liberal professionals (9). This classification is borrowed from Brazilian sociological studies on occupational mobility (see Pero, 2001; Valle e Silva, 1978)

restricted to the individual schooling, measured in years [13], years of schooling squared, to capture possible non-linearities, and a migration dummy, defined as whether the observed municipality of residence is different from the one where the individual was born.[14] Descriptive statistics of the main variables are shown in Table 1.

Earnings equations

Earnings equations were estimated separately for men and women, and by cohort, using OLS for men and the standard two-stage selection bias correction procedure for women [15]. Results are shown in Table 2. Note that, unlike in the standard Mincerian specification, age or imputed experience do not appear among the regressors because we are treating cohorts as age-homogeneous by definition.

Circumstance variables have the expected effect on earnings. The coefficients of racial dummy variables are negative and significant for both blacks and ‘pardos’.[16] They are generally positive, but not always significant for people with an Asian origin. It is interesting that the racial gap against blacks and 'pardos' is less pronounced for women. It is not even always significant. Regional differences are important, too. With the South-East as a reference, being born in the North-East has a strong and significantly negative effect for both men and women. The effect of the other regions is generally also negative, but seldom significant. The estimated effect of mean parental education on individual earnings is always positive, significant, and relatively stable across cohorts. It is also sizable since it amounts to a 3 to 6 per cent increase in earnings, for each additional year of schooling of the parents. The difference between the education of the father and the mother is meant to detect a possible asymmetry in the role of the two parents. But no such asymmetry seems to be systematically present. Concerning the estimated effect of father’s

[13] The number of years of schooling directly provided in the PNAD is bounded at 15. For consistency with the scale used for parents' schooling, this variable was changed to 17 for individuals reporting a master or a doctorate degree.
[14] The decision to migrate might have been a decision of the individual him/herself when adult, or of his parents when he was still a child. While in principle it should be taken as a circumstance variable in the second case and as an effort variable only in the first case, we can not distinguish the two from the available data. We treat migration as an effort because results are more consistent with efforts than circumstances, as we will see below, but this interpretation should be treated with considerable caution.
[15] The Heckman correction was initially applied to men as well, but available instruments proved unsatisfactory.
[16] Race is self-reported in the PNAD: the respondent, rather than the interviewer, chooses his or her race. ‘Pardo’ is meant to refer to people of mixed-race, generally involving some Afro-Brazilian component.

occupation on earnings, it is generally positive - the reference category being rural workers - though seldom significant across cohorts, once we control for education, except for self-employed shopkeepers and employers. [17]

Turning now to the vector of "effort" variables, own education has the usual positive and significant effect on earnings for men. This effect is decreasing as one considers younger male cohorts. This is consistent with the negative coefficient generally found for the squared imputed experience term –i.e. age minus number of years of schooling minus first schooling age - in the standard Mincerian specification. This implies that returns to schooling increase with age, which is exactly what is found here.[18] The coefficient of schooling is sometimes insignificant, particularly for women. The reason is that the overall education effect is captured by the squared years of schooling term, which is positive and significant both for men and women. This means that the marginal return to education increases with the number of years of schooling.

The order of magnitude obtained for the return to schooling in the preceding equations is somewhat lower than previous estimates for Brazil. For instance, Ferreira and Paes de Barros (1999) found that the marginal return to a year of schooling lies in the range 12 to 15 percent for both men and women in 1999. In Table 2, marginal returns at 5 years of schooling range from 7 to 11 percent for men and from 11 to 13 percent at 10 years of schooling. This may be due in part to the specification being used here, which is not strictly comparable to the Mincerian model. The difference is also consistent, however, with the probable over-estimation of the returns to schooling in an earnings equation that does not include family background variables.[19]

---

[17] In fact, we observed that all father's occupational categories are highly significant determinants of individuals' years of schooling in our effort equations described further on. Only self-employed and employers seem to be significant determinants of earnings, independently from education.

[18] The conventional Mincerian specification is such that: $Lnw = a.S + b.Exp - c.Exp^2$ where $Exp = Age - S - 6$. Expanding the Exp term leads to: $Lnw = (a-b-12c).S + 2cAge.S - c.S^2$ + terms in Age or Age squared. If this equation is estimated within groups with constant age, one should indeed observe that the coefficient of S is higher in older cohorts. Note that the present specification also includes an independent S^2 term.

[19] The preceding intuition is fully confirmed by the data used in this study. In unreported regressions, though available from the authors upon request, we have re-estimated the preceding wage equation with years of schooling instrumented by parents' schooling achievements and the other exogenous variables of the model, and with parents' education excluded from the regressors. The coefficients of the number of years of schooling turn out to be substantially higher than in the previous case because they now partly account for the direct influence of parental education on individual earnings. Their order of magnitude is also comparable to what has been found in other earnings equations estimated for Brazil (see, for example, Ferreira and Paes de Barros, 1999).

Migration has a significant and positive effect on earnings, both for men and women. This sign would be consistent with a human capital interpretation of migration. Because the coefficient is rather large, amounting to a 10-18 per cent increase in earnings, it is tempting to consider this as an effort variable. But, it may also reflect the decision of parents to move to an area with better income opportunities when the surveyed individual was still a child, in which case this variable should be taken as indicative of circumstances. If this were true, however, the size of the estimated coefficients would suggest very much persistence in the earnings differential that might have motivated the migration of the parents.[20]

Effort equations and intergenerational educational mobility

Let us now turn to estimates of equation (3), and consider the impact of circumstances on efforts (i.e. schooling and migration). Because no significant model for migration was found, this variable is dropped from the analysis. Regression results for schooling are shown in Tables 3a and 3b separately for men and women. They call for several remarks.

The first set of remarks has to do with intergenerational educational mobility as measured, inversely, by the coefficient of parental education. The higher that coefficient, the stronger is parental education in determining the schooling of their children, and therefore the less mobility there is. Because education is measured for both parents and children in years of schooling, a unit value for that coefficient is a convenient reference. It would correspond to the perpetuation of differences in years of schooling across generations – this being consistent with an increase in mean schooling. Conversely, a coefficient less than unity means that educational differences tend to diminish across generations. From that point of view, a striking feature in tables 3a and 3b is that the coefficient of the mean schooling of parents is significantly below unity and has been decreasing continuously and significantly over time. Overall the gain is substantial. For people born in the early 1940s, a one-year difference in the schooling of their parents resulted in a difference of approximately 0.75 years in their own schooling. For those born in the late 1960s, the same initial difference in parental education resulted in approximately half a year of schooling.

[20] Note that we are considering migration across municipalities and not regions for which persistence of earning gap

Attributing this decline entirely to the general rise in mean education over time would be incorrect, because of the role of the intercept in this effort equation. If a majority of children are now going to school for 5 years whereas the majority was going to school only for 3 years 20 years ago, it may seem natural that the influence of parental education declined with time. This is not necessarily true, however. This 2-year addition to mean schooling achievement might very well hold for the whole population, whatever their family background. If this were the whole story, then only the intercept in the regressions reported in tables 3a and 3b would be increasing across successive cohorts, and the coefficients of all variables would remain constant. This is clearly true of race, for instance, for which no clear trend seems to be present. Black people have the same quantitative disadvantage in education in the 1960s - 1 to 2 years of schooling - as they had in the 1940s or the 1950s. Likewise, the disadvantage of being born in the North-East for men has remained approximately constant. In effect, only the coefficient reflecting the disadvantage of being born from parents with a low level of schooling seems to have been falling regularly over time. In other words, the conditional distribution of educational opportunities seems to have remained approximately constant over time, except with respect to educational family background.

An interesting feature of intergenerational educational mobility is that intra-household decision mechanisms seem to matter more for the education of women than for men. The transmission of education from parents to girls is higher, the greater the mother's schooling relative to the father's. This effect is significant and persistent across cohorts. For boys, however, although the mother's education still seems to weigh more than the father's, the difference is only significant for a single cohort. In both cases, it is difficult to find a trend in the evolution across cohorts.

A full understanding of the inequality of educational opportunities in Brazil would require a more detailed analysis. In particular, very much of the preceding discussion is based on measuring education in terms of the number of years of schooling. One might prefer a more general approach based on 'human capital' rather than years of schooling, where human capital might be measured by the cost of education, including forgone earnings, or possibly by the earnings that a

---

would be natural.

given schooling level actually commands. Also, the quality of schooling is totally ignored in the preceding description of intergenerational educational mobility. However, it cannot be ruled out that taking into account the quality of education, so as to get closer again to a concept of human capital, would modify the preceding conclusion of an increasing educational mobility in Brazil.[21] Unfortunately, there is no data source in Brazil which combines information on plausible measures of educational quality with information on the schooling of parents, for individuals currently active in the labor market.

3. Evaluating the inequality of opportunities and its influence on the inequality of earnings

The methodology presented in section 1 is now applied on the basis of the preceding earnings and effort equations. Making all variables in the analysis explicit, the two basic equations (1) and (3) now write:

$$Ln(w_i) = \alpha_0 + R_i\alpha_R + GR_i.\alpha_G + MPE_i.\alpha_P + DPE_i.\alpha_D + FO_i.\alpha_F + S_i.\beta_S + S_i^2.\beta_{S2} + M_i.\beta_M + u_i \qquad (8)$$

$$S_i = b_0 + R_i b_R + GR_i.b_G + MPE_i.b_P + DPE_i.b_D + FO_i.b_F + v_i \qquad (9)$$

where S is the number of years of schooling of the surveyed individual, S^2 the square of than number, and *M* his/her migrant status. *R, GR, MPE, DPE* and *FO* stand respectively for the race dummies, the regional dummies, mean parental schooling, the mother/father difference in schooling, and father occupation. α_R , α_G , b_R, and b_G are vectors of coefficients whereas other parameters (α_P , α_D , α_F , β_S , β_{S2} , β_M , b_P , b_D , b_F) are scalars. Estimates of all these coefficients are partly shown in Tables 2 and 3 and they have been discussed extensively in the preceding section. In principle, there should be an equation like (9) for all effort variables appearing in equation (8), that is, for *S*, S^2 and *M*. However, the equation for S^2 necessarily leads to conclusions similar as (9) whereas it was not possible to identify any significant effect of the circumstance variables on the migration status, *M*.

[21] See Albernaz, Ferreira and Franco (2002) for a discussion of the variance of educational quality in Brazil, as measured by standardized test scores, and of its determinants.

Identifying the role played by the inequality of opportunities in the distribution of current earnings consists of equalizing the circumstance variables either in equation (8) – for the “partial” effect - or in both equation (8) and (9) – for the “complete” effect. The distribution obtained in the first case accounts partially for the inequality of opportunities, because it ignores its role through determining the level of efforts, whereas the second distribution accounts also for this latter effect.

Simulation results are shown in Figures 1a-b and 2a-b. Although it would have been possible to show the whole distribution resulting from equalizing circumstances, only two summary inequality measures of these distributions are presented, to save on space. Figures 1a and 2a show the Gini coefficient for men and women respectively, whereas figures 1b and 2b show the Theil coefficient. In all figures, the top line represents the inequality actually observed for the various cohorts. The line below it shows the “partial” effect of equalizing circumstances, whereas the bottom line shows the “complete” effect. Finally, the dotted lines around the two bottom lines show upper and lower bounds for the preceding estimates, as explained in the appendix.

The complete effect of equalizing circumstances is to reduce the Gini coefficient by 8-10 percentage points, for all cohorts of both men and women. Ignoring the effect of circumstances on individual effort variables, in the 'partial effect' calculation, leads to a drop of approximately half of that effect. Again, this is true for practically all cohorts, for both men and women earners. The absolute and relative drops are more pronounced with the Theil coefficient than with the Gini. In any case, equalizing circumstance variables leads to a drop in the Theil of between 12 and 25 points for the complete effect – the 35 points drop observed for the 1941-1945 being clearly exceptional.

Naturally, the parametric bounds analysis used to address the endogeneity bias in the estimation of the two equations generates some imprecision in some of the estimates. Indeed, dotted curves in Figures 1a to 2b show that the unknown degree of endogeneity of efforts makes the estimated partial or direct effect of equalizing circumstances on the distribution of earnings rather imprecise. The distance between the extreme bounds is 2 to 4 percentage points for men in the case of the Gini coefficient and 4 to 6 points for women. By contrast, bounds on the complete effects are extremely close to the mean effects. In other words, there is very little ambiguity about the important role played by the inequality of opportunities in shaping the unequal distribution of earnings in Brazil. The statistical imprecision is only about the *way* in which this occurs, that is about how much of the effect of circumstances on earnings is mediated through the impact of circumstances on efforts.[22]

Several conclusions may be drawn from these experiments. The first is that, even after correcting for the inequality of observed opportunities, inequality in Brazilian earnings remains very high, with a Gini coefficient above 0.44 for all cohorts and sometimes even above 0.5. Two interpretations are possible. On the one hand, *observed* circumstance variables accounted for in the present analysis (i.e. parents' education, father's occupation, region of origin and race) are actually a limited subset of all circumstances. Among those potentially observable, parental income or wealth are not available in our data set. Yet, they are accounted for in part by observed variables with which they are correlated, such as parental education, occupation and region of origin.[23] It would thus be surprising if their inclusion drastically modified the preceding results. As for circumstance variables that are intrinsically unobservable, they are by their very nature analytically and practically irrelevant for understanding the sources of inequality in Brazil. On the other hand, it may simply be the case that observed non-opportunity related earnings inequality is very high in Brazil, and presumably higher than in other countries. Possible reasons for this include structural characteristics of the labor market that would tend to magnify differences across people with the same observed background, or possibly a large variance of persistent and non-persistent shocks in the earning career of individuals.

A second important conclusion of the analysis is that the proportion, or even the absolute value of inequality due to observed opportunities in actual inequality seems rather stable across cohorts, regardless of whether the Gini or Theil coefficients are considered. This finding is consistent with

[22] The fact that the bounds are close to each other for the complete effect and far apart for the partial effect is not really surprising. It may be seen from (7) that the estimation of the complete effect is in some sense equivalent to a reduced form model, where (log) earnings is a function of circumstances only. There would be no OLS bias in that case. This is in contrast with the partial effect model (6), where the coefficients may be strongly biased because of the endogeneity of efforts.

[23] To check how income is correlated with education, occupation and region of origin in the current generation – which is the best we can do - we ran a regression of income on these variables for the PNAD respondents with children. R-squares are in the range 0.31 - 0.42, depending on the cohort.

the idea that circumstances are essentially responsible for differences in the “starting point” of individual earnings. Later in the life-cycle, however, other factors are responsible for changes in inequality. The fact that both actual inequality and circumstance-corrected inequality increase in a parallel way with age is consistent with persistent circumstance-independent labor market shocks for active individuals.[24]

On the basis of the increased intergenerational educational mobility discussed above, one might have expected to see the proportion of the inequality accounted for by opportunities fall, when moving from older to younger generations. That this is not the case may have several explanations. Two of them are as follows. First, the change in intergenerational educational mobility found in the education equations above may be too limited to have a significant effect on the inequality of earnings. Second, returns to education tend to be lower for the youngest generations, which somewhat compensates the equalizing effect of more intergenerational educational mobility.

Finally, it is interesting to evaluate the role of particular circumstance variables in the preceding results. The complete effect of equalizing individual circumstance variables is shown for the Gini coefficient on figures 3a and 3b – results are qualitatively analogous for the Theil coefficient. It can be seen there that, of all circumstance variables, parental education plays the largest role in determining inequality. Father's occupation is the second most important variable but, as it is correlated with education, its independent contribution is smaller than suggested by figures 3. With respect to education, it is worth emphasizing that results are not very different when parental schooling is not equalized across the board, but instead a lower bound is imposed, as if schooling were compulsory until a certain age. This means that it is the inequality of education at the bottom of the distribution that really matters to explain the contribution of the inequality of opportunities to actual inequality in earnings. Interestingly enough, race alone seems to account for very little, when parental occupation and education are already controlled for. These results suggest that the most effective policies for reducing inequality of opportunities in Brazil would be those that reduced the effect of parental education on their child’s schooling and earnings. This conclusion is even stronger for women than for men.

[24] This law of increasing variance of earnings breaks down in Brazil only for the older cohorts of men and women.

4. The effects of the inequality of opportunities on the distribution of household income

The analysis so far has referred to the earnings of active individuals. The same type of analysis may also be conducted for economic welfare levels, whether the individuals are active or not. Why should one rule out women who are out of the labor force from the measurement of the inequality of opportunities? That she is outside the labor force and has many children may well be a channel through which circumstance variables affect the economic welfare of an individual. This section thus focuses on measuring the effect of the inequality of opportunities faced in the past by household heads and spouses on the distribution of economic welfare among them. The welfare measure used in this section is the monetary household income per capita. Thus each adult in the population – or more precisely: all household heads and spouses - is imputed a welfare level equal to income per capita in his/her household. With this new definition, the opportunities faced by household heads and spouses now affect not only their earnings, as before, but also their participation behavior, fertility, non-labor income and, of course, the matching of individuals within couples.

In an effort to capture these effects the previous model was re-estimated substituting household per capita income for earnings and considering the whole population of adults rather than earners only. In effect, three different models were estimated, with the objective of identifying various channels through which opportunities affect welfare: individual earnings, labor supply behavior, fertility, etc.. . In all cases, the welfare level of individual i is given by:

$$Y_i = \frac{y_{hm(i)} + y_{sm(i)} + y_{0m(i)}}{n_{m(i)}} \qquad (10)$$

where *m(i)* is the household of which *i* is the head or the spouse, $y_{jm(i)}$ is the earnings of member *j* (=*h* for household head ,*s* for spouse), $y_{0m(i)}$ is non-labor income – really, income on top of the earnings of *h* and *s* - and $n_{m(i)}$ is the number of persons in the household.

The first model generalizes directly the approach from the previous section. The earnings of each active member, *h* or *s*, is taken as a function of the circumstances and efforts of that person:

Evidence of this age-dependence of earnings inequality in other countries is analyzed in Deaton and Paxson (1994).

$$Log(y_i) = C_i\alpha + E_i\beta + u_i\,; \qquad E_i = C_i b + v_i \tag{I}$$

The estimation is run only on individuals with positive earnings. The simulations are carried out as before by equalizing circumstances and re-computing earnings for all active persons. Then (10) is applied to these simulated earnings, and the resulting income per capita is imputed to all individuals, active or not. This model therefore estimates only those effects of the inequality of opportunities on the distribution of welfare which are mediated through the channel of individual earnings, treating household composition, occupational decisions and non-labor incomes as given.

The second model aims at identifying effects on household size and composition. Besides individual earnings, family size is now allowed to depend on circumstance and effort variables of the two spouses, according to the familiar multinomial logit specification. This model can be described as follows :

$$Log(y_i) = C_i\alpha + E_i\beta + u_i\,; \qquad E_i = C_i b + v_i\,; \quad \Pr\{n_{m(i)} = k\} = \frac{e^{Z_{m(i)i}\lambda_k}}{\sum_p e^{Z_{m(i)}\lambda_p}} \tag{II}$$

where Z_i reasonably includes the circumstances and efforts of the two parents $\{C_{hm(i)}, C_{sm(i)}, E_{hm(i)}, E_{sm(i)},\}$ and $k = \{2, 3, .., 7$ *and more)* is family size.

The third model is completely reduced-form. We now simply regress the welfare level of individual *i* on circumstance and effort variables. But this income per capita naturally depends on the circumstances and efforts of both spouses since it incorporates the earnings of both. The model thus writes :

$$Log(Y_i) = C_{hm(i)}\alpha_h + C_{sm(i)}\alpha_s + E_{hm(i)}\beta_h + E_{sm(i)}\beta_s + \varepsilon_i \;\; ; \;\; E_i = C_i b + v_i \tag{III}$$

This estimation is run on two parent/adult households, whereas model (I) continues to apply to singles or single parent households.

Model (III) takes into account various channels through which the inequality of opportunities may affect the distribution of personal welfare. It does so in reduced-form, however, and it is thus impossible to disentangle them from just this model. Behind equation (III), is not only the way in which individual circumstances and efforts affect individual earnings, but also how they affect labor supply behavior, the presence of non-labor income, the matching of spouses, or fertility, since the household size is also present in the definition of the level of welfare.

Using these three models, the effect of equalizing the circumstances of household heads and spouses on the distribution of welfare may be simulated in various ways. Comparing these ways permits identifying some of the channels through which circumstances affect the distribution of welfare. Consider the three following simulations:

$$y_i^* = (y_{hm(i)}^I + y_{sm(i)}^I + y_{0m(i)}) / n_{m(i)}$$

$$y_i^{**} = (y_{hm(i)}^I + y_{sm(i)}^I + y_{0m(i)}) / n_{m(i)}^{II}$$

$$y_i^{***} = y_i^{III}$$

where the left-hand side variable is income per capita in the household which *i* belongs to after equalizing circumstance variables according to model (I), (II), and (III). Model (I) shows the effect of equalizing circumstances through the sole channel of the earnings of active people. Comparing $\{y_i^*\}$ and $\{y_i^{**}\}$ should inform on the role played by the household size channel. Finally, comparing the distribution $\{y_i^*\}$ and $\{y_i^{***}\}$ - should show by difference the role played together by the labor supply channel, fertility, matching and non-labor income. All of these simulations are performed with the specification for the 'complete effect' of equalizing circumstances – i.e. taking into account the indirect effect of circumstances, which operates through their impact on effort choices.

It may be seen in Table 4 that the drop in inequality obtained from equalizing the effect of circumstances on earnings is approximately the same when considering the distribution of household income per capita as when analyzing individual earnings (in the previous section). In both cases, the Gini coefficient falls by some 8 to 10 percentage points, whereas the fall in the Theil coefficient for the distribution of welfare roughly corresponds to the average fall observed for male and female earnings inequality – see figures 1b and 2b. To the extent that potential earnings and family background tends to be positively and rather strongly correlated within couples, such a result was to be expected.

The role of household composition (including fertility) may be identified by comparing the effects of models I and II in table 4. Unsurprisingly, one can check that the complete effect of equalizing circumstances is limited for oldest cohorts and increases continuously as age goes down. For younger cohorts, the effect that circumstances have on per capita income through the household size effect is rather sizable: more than 4 percentage points of the Gini and 8 points of the Theil coefficient. Yet, this result must be understood with care. The analysis behind the results shown in table 4 is purely mechanical. More or less children reduce or increase income per capita in a pure arithmetic sense. The induced effects of fertility on labor-force participation, or the fact that a large number of children may depend on the potential earnings of parents are not taken into account in this calculation. To do so, a structural model should be used.

Moving now to the more general, reduced-form model where circumstances are allowed to affect the distribution of welfare through any number of channels, contrasting results are obtained. The additional drop in inequality in model (III) with respect to model (I) is substantial for the oldest cohorts, amounting to 6 percentage points of the Gini for the oldest one and 3 points for the one that follows. But then the fall in inequality is only 2 percentage points on average for all younger cohorts. A comparable pattern is observed with the Theil coefficient. Together, these results suggest that life-cycle related phenomena may be at work.

Fertility is one of these phenomena and it has already been analyzed. Other channels through which equalizing circumstances may affect the distribution of welfare are labor supply, choice of partner, and non-labor income. Their effect may be gauged from comparing the 3rd and 4th block in table 4. These residual effects were largest for the oldest cohort. Indeed, they were responsible for a drop in the inequality of welfare among the oldest generations, but for an increase among the youngest. For the oldest cohorts, one may think that the incidence of retirement income may play an important role. It has been shown in previous work that pensions are so regressive in Brazil that they do much to explain the excess in the country's inequality relative, say, to the United States – see Bourguignon, Ferreira and Leite (2002). As these pensions are likely to go to individuals with favorable family circumstances, a large effect may indeed be expected through this channel for the oldest cohorts in the present sample. For the youngest cohorts the unequalizing effect of equalizing circumstances that goes through labor supply and non-labor income is more puzzling, even though it is limited in size.

A possible explanation would be the negative dependence of female participation on non-labor and male labor income. By equalizing circumstances, this income source is becoming smaller among the richest households and bigger among the poorest ones. If participation behavior depended only on current income and not independently on circumstances, it would follow that female participation would go up among the richest and down among the poorest. Thus equalizing circumstances would lead to more inequality on the participation account.

Even if this argument helps explain the results shown in Table 4, the labor supply channel for the effect of equalizing circumstances is likely to be far more complicated than implicitly assumed here. First, circumstances may have a direct effect on participation, independently of income. Second, potential earnings are also likely to be an important determinant of labor supply with the opposite effect of equalizing circumstances on inequality. Third, fertility behavior is also likely to affect labor supply, as mentioned above. For all these reasons, it would seem necessary to rely on a structural model that would include all these dimensions of behavior to fully understand the origin of the preceding effects. Doing so might also show that relying on pure monetary income to define welfare, as done here without any correction for labor supply, may not be satisfactory.

In summary, the role of the inequality of opportunities in shaping the distribution of household welfare is larger than in the case of individual earnings. If observed circumstances could be equalized, welfare inequality would go down by much more than earnings inequality. The drop in the Gini coefficient is around 12 percentage points for incomes and around 8 points for earnings. As the arithmetic effect that goes through individual earnings in the case on welfare inequality is comparable in size to what was found for individual earnings, additional effects must be at work in the case of household income per capita. The simple simulations undertaken in this section suggest that household size and composition, labor force participation and non-labor income are important channels through which opportunities can affect the distribution of household income. A precise identification of the contribution of each channel remains for future work. At this stage, only the household size channel has been shown to be potentially very important, particularly among younger households.

Nevertheless, the conclusion that inequality in Brazil remains very high even after equalizing observed circumstances holds in the case of the distribution of welfare, just as it did for individual earnings. When all effects are taken into account as in the bottom rows of Table 4, equalizing circumstances still leaves a Gini coefficient of approximately 0.5. This is undoubtedly high by all international standards.

5. Summary and conclusion

This paper tried to quantify the role of the inequality of opportunities – associated with people's race, region of origin, the education and the occupation of their parents – in generating inequality in current earnings and incomes in Brazil. We estimated the impact of opportunities (or circumstances) both directly on earnings and incomes, and indirectly on the level of efforts – such as schooling - undertaken by individuals. We took account of the biases arising from the lack of adequate instruments for correcting for the endogeneity of some of the income determinants, by showing upper and lower bounds for the unbiased estimates, for plausible values of the bias.

Altogether, the inequality of observed opportunities is responsible for a very substantial proportion of total outcome inequality in Brazil. It accounts for approximately 8-10 percentage points of the Gini coefficient for individual earnings. Fifty-five to 75 percent of this share can be attributed to parental schooling alone, and 70 to 80 percent when the father's occupation is added. The effect of opportunities is even higher for household income per capita, amounting to some 12 percentage points of the Gini (and even more for some cohorts). The reason for this difference with individual earnings is that opportunities affect welfare levels both through earnings and through additional channels. This is true in particular of household size and composition and, probably to a lesser extent, of labor-force participation, non-labor income and assortative mating.

Although international comparisons of the intergenerational transmission of inequality are always difficult - because of differences in definition, methodology and data type - the preceding figures nevertheless look high by international standards. Table 5 shows that the share of the variance of (log) wages explained by the (log) wage of the parents lies between 1 and 35 percent in most existing studies for developed countries. The circumstance variables used in this study – which exclude parental income - already explain between 25 and 30 per cent of the variance of the (log) earnings rate, and they would probably explain still more if the parental income or wealth was observed. Likewise, the share of the variance of (log) family income is between 2 and 42 per cent in existing studies in other countries. It is between 32 and 44 per cent for Brazil. Of course both comparisons must be related to intergenerational educational mobility figures. The singularity is obvious here too. Parents' schooling explain at most 20 per cent of children's schooling in existing studies. This proportion is between 35 and 47 per cent in Brazil. In effect, the inequality of opportunities in Brazil compares only with the maximum estimate obtained in the literature for (log) family consumption in the US.

Our finding that the degree of intergenerational educational persistence in Brazil is high by international standards is consistent with the results of two other recent studies which use the same PNAD data set we used: Dunn (2003) and Ferreira and Veloso (2003). If anything, these studies find shares of variance in earnings (Dunn) and schooling (Ferreira and Veloso) accounted for by parental variables which are even higher than ours although, in both cases, this may be due to the estimation techniques employed.[25]

The role played by inequality of observed opportunities in shaping the inequality of current earnings and incomes in Brazil is clearly important. The effects also appear to be larger than in

[25] Dunn (2003) instruments for fathers' earnings using father's education. As he acknowledges: "If fathers' educations are independently positively correlated with sons' earnings, then the IV elasticity estimate will be upwards-inconsistent."(p.5). He correctly treats his estimate as an upper-bound. Ferreira and Veloso (2003) use a simple OLS estimator, and their coefficient on parents' education will be upwardly biased if unobserved determinants of children's schooling – such as parental wealth and ability – are positively correlated with parental education.

other countries for which data is available. Yet, the inequality that they leave unexplained remains very substantial. In effect, correcting for observed disparities in opportunities would leave the Gini coefficient at 0.45 and above for individual earnings and 0.48 and above for income per capita, levels which are higher than total inequality in many countries, including those listed in Table 5. Thus, although the inequality of opportunities in Brazil is important, the country would still rank high in an international comparison of inequality even after eliminating that particular source of inequality.

Figures 1 and 2 also indicated that both the direct and the indirect effects of circumstances on earnings are important, each in its own right. This suggests that policies aimed at equalizing opportunity may be warranted *both* inside and beyond the classroom. Parental education – and, to a lesser extent, occupation – do affect the length of children's school careers. Efforts to reduce this dependence, through conditional cash-based assistance to students and their families, or through after-school programs for students who may be falling behind, might well deserve consideration. But family background clearly also impacts on earnings directly, even after conditioning on own schooling. This is consistent with hypotheses that both employment and career advancement opportunities may be allocated in part through socially-based networks.[26] In this paper, we have not presented any evidence on the existence of such networks, or on their welfare properties, but further investigation of their operation in the Brazilian case is warranted, and may be relevant for the ongoing debate on affirmative action in Brazil.

[26] Evidence from elsewhere suggests that socially-based networks can be effective in matching workers from certain families to coveted jobs. One example is the traditional segmentation of blue-collar occupations in Bombay by *jati* (or sub-caste) groups. While such labor-market effects are likely to impact on educational decisions (as explored by Munshi and Rosenzweig, 2003), they also constitute a direct impact of family background on earning opportunities, conditional on the child's schooling attainment.

Appendix: Taking into account the endogeneity of efforts and estimating bounds for the inequality of opportunities

The measure of the inequality of opportunities used in this paper depends on the estimates of the coefficients of the equation that explain earnings as a function of circumstances and efforts –i.e. equation (1). Because of the endogeneity of efforts and the absence of adequate exogenous instrument to deal with it, it is necessary to discuss the implications of the bias that this may imply for the estimates of the coefficients of the earning equation and therefore for the estimation of the inequality of opportunities. This appendix describes the method used to obtain 'bounds' on these estimates, which was inspired by the 'bounds analysis' developed by Manski and Pepper (2000), in a different context.

Consider the model :

$$Lnw_i = X_i \beta + u_i \qquad \text{(i)}$$

where it is assumed that the error term u_i is not necessarily orthogonal to all explanatory variables in *X*. Assume without loss of generality that all the variables have zero mean and define the following covariance matrices:

$$\Sigma = \begin{bmatrix} X'X & X'u \\ u'X & u'u \end{bmatrix} \quad \text{and} \quad S = X'X$$

The bias of OLS estimates of equation (i) is given by B in (ii):

$$E(\hat{\beta}) = \beta + B \;\; \text{with} \;\; B = S^{-1}X'u = S^{-1}(\rho_{Xu} \otimes \sigma_Y)\sigma_u \qquad \text{(ii)}$$

where ρ_{Xu} stands for the correlation coefficients between the components of *Y* and the residual term, *u*, and σ_V is the standard error of variable *V*. Evaluating the bias vector *B* thus requires knowing σ_u and ρ_{Xu} . An unbiased estimator of σ_u is readily obtained for any set of correlation coefficients ρ_{Xu} . Indeed, it can be shown that :

$$\sigma_u^2 = \hat{\sigma}_u^2 + B'SB$$

where $\hat{\sigma}_u^2$ is the variance of the OLS residuals. Substituting the value of the bias given in (ii) in that expression yields:

$$\sigma_u^2 = \hat{\sigma}_u^2 /(1-K) \qquad \text{(iii)}$$

with K given by:

$$K = (\rho_{Xu} \otimes \sigma_Y)' S^{-1} (\rho_{Xu} \otimes \sigma_Y). \qquad \text{(iv)}$$

For any set of correlation coefficients ρ_{Xu}, equations (ii)-(iv) thus permits computing the bias vector B and thus obtaining unbiased estimates of the coefficients of the model, β, and of the variance of the error term. As these correlation coefficients are not known, the idea is to use Monte-Carlo methods: we draw randomly a large number of values for them and derive bounds for the estimates of coefficients β, as well as for the results of the micro-simulation exercises undertaken in this paper, on the basis of these coefficients.

An important point is that not all vectors of correlation coefficients ρ_{Xu} are possible, which contributes to reducing the size of the bounds. In practical terms, some components of ρ_{Xu} are constrained to be zero – in this paper we have assumed that the residual term *u* is orthogonal to circumstances in equation (1) - while the others are drawn independently from a uniform distribution defined on (-1, 1). The correlation coefficient vector ρ_{Xu} must satisfy the condition that the covariance matrix Σ be positive. All drawings such that this condition is not satisfied have been deleted.

We used a set of 300 valid drawings for each simulation leading to an estimation of the inequality of opportunities. Figures 1-3 in the text report the mean and extreme values in the set of 300 results obtained for each calculation. To the extent that these extreme values do not depend on any specific assumptions on the correlation between the explanatory variables and the residual term, they appear as 'natural bounds' for the analysis of the inequality of opportunities. As the distance between them turned out to be rather limited, it did not seem necessary to impose further arbitrary restrictions on the correlation coefficient vector, ρ_{Xu}, as in bounds analysis (see, for instance, Manski and Pepper, 2000).

References :

Albernaz, A., F.H.G. Ferreira and C. Franco (2002): “Qualidade e eqüidade no ensino fundamental brasileiro”, *Pesquisa e Planejamento Econômico*, **32** (3).

Barros, R. P. and D. Lam (1993): “Desigualdade de renda, desigualdade em educação e escolaridade das crianças no Brasil”, *Pesquisa e Planejamento Econômico*, **23** (2), pp.191-218.

Behrman, J., N. Birdsall and M. Szekely (2000), “Intergenerational mobility in Latin America: deeper markets and better schools make a difference”, in N. Birdsall and C. Graham (eds), *New markets, new opportunities? Economic and social mobility in a changing world*, Washington, D. C. : Brookings Institution, p. 135-67

Bourguignon, F., F.H.G. Ferreira and P. Leite (2002), “Why are income distributions different? A comparison of Brazil and the United States”, Mimeo, The World Bank, Washington, D. C.

Bowles, S. (1972), “Schooling and inequality from generation to generation”, *Journal of Political Economy*, **80** (3), pp.S219-S51

Bowles S. and Gintis , H. (2002), “The inheritance of inequality”, *Journal of Economic Perspectives*, **16** (3), pp. 3-30

Burkhauser R., D. Holtz-Eakin, and S. Rhody (1998), “Mobility and Inequality in the 1980s: A Cross-National Comparison of the United States and Germany”, in S. Jenkins, A. Kapteyn and B. M. S van Praag (eds.) *The distribution of welfare and household production: International perspectives*, Cambridge University Press, 111-75.

Card, D. (20021), “Estimating the returns to schooling: progress on some persistent econometric problems”, *Econometrica*, **69** (5), pp. 1127-60

Checchi, D., A. Ichino, A. Rustichini (1999), “More Equal but Less Mobile? Education Financing and Intergenerational Mobility in Italy and in the US”, *Journal of Public Economics* **74** (3), pp. 351-93

Deaton, A. and C. Paxson (1994), “Intertemporal choice and inequality”, *Journal of Political Economy*, **102** (3), pp. 437-67

Dunn, Christopher (2003): “Intergenerational Earnings Mobility in Brazil and Its Determinants”, University of Michigan, unpublished.

Ferreira, F.H.G.; P. Lanjouw and M. Neri (2003): “A Robust Poverty Profile for Brazil Using Multiple Data Sources”, *Revista Brasileira de Economia*, **57** (1), pp.59-92.

Ferreira, F.H.G. and R. Paes de Barros (1999): "The Slippery Slope: Explaining the Increase in Extreme Poverty in Urban Brazil, 1976-1996", *Brazilian Review of Econometrics*, **19** (2), pp.211-296.

Ferreira, S. G. and F. A. Veloso (2003): “Mobilidade Intergeracional de Educação no Brasil”, Banco Nacional de Desenvolvimento Econômico e Social (BNDES), Rio de Janeiro, unpublished

Goux, D. and E. Maurin (2001): "La Mobilité Sociale et son évolution: le rôle des anticipations réexaminé", *Annales d'Economie et de Statistique*, **62**, forthcoming.

Griliches, Z. and W. Mason (1972), “Education, Income and Ability”, *Journal of Political Economy*, **80** (3), pp.S74-S103

Hanushek, E. (1986): “The Economics of Schooling: Production and Efficiency in Public Schools”, *Journal of Economic Literature*, **24**, pp.1141-1177.

Lam, D. (1999), “Generating extreme inequality: schooling, earnings, and intergenerational transmission of human capital in South Africa and Brazil”, Population Studies Center, University of Michigan, Report 99-439

Manski, C. and J. Pepper (2000), “Monotone instrumental variables: with an application to the returns to schooling”, *Econometrica*, **68** (4), pp. 997-1010

Mulligan, C. (1999), “Galton versus the human capital approach to inheritance”, *The Journal of Political Economy*, **107** (6), pp. 184-224.

Munshi, K. and M. Rosenzweig (2003), “Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy”, Harvard University, unpublished.

Pero, V. (2001) "Et, à Rio, plus ça reste le même... Tendências da mobilidade social intergeracional no Rio de Janeiro", ANPEC, Salvador.

Piketty, T. (1995): “Social Mobility and Redistributive Politics”, *Quarterly Journal of Economics*, **CX** (3), pp.551-584.

Roemer, J. E. (1998): *Equality of Opportunity*, (Cambridge, MA: Harvard University Press)

Solon, G. (1999), “Intergenerational mobility in the labor market”, in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Amsterdam, North-Holland, Vol. 3A, pp. 1761-1800

Taubman, P. (1976), “The determinants of earnings: genetics, family and other environment; a study of white male twins”, *American Economic Review*, **66** (5), pp. 858-70

Valle Silva, N. (1978), *Posição social das ocupações*. Rio de Janeiro: IBGE.

Table 1. Descriptive statistics.

Cohort	b1936_40	b1941_45	b1946_50	b1951_55	b1956_60	b1961_65	b1966_70
Mean monthly earnings (Reais, all jobs)	494.9	680.7	743.7	749.4	683.8	607.1	490.1
Mean number of years of schooling	4.4	5.5	6.2	7.0	7.4	7.7	7.6
Mean father's number of years of schooling	2.2	2.4	2.7	2.9	3.1	3.5	3.5
Mean mother's number of years of schooling	1.6	1.9	2.2	2.4	2.8	3.2	3.2
Race (Percents)							
Branca (Whites)	59.9	59.1	59.0	59.6	58.5	59.5	57.5
Preta (Blacks)	6.6	6.7	6.3	6.2	5.8	5.3	5.6
Amarela (Asians)	0.4	0.7	0.5	0.5	0.4	0.4	0.3
Parda (MR)	33.1	33.5	34.2	33.8	35.3	34.9	36.7
Regions (Percents)							
North	4.5	4.6	5.3	5.7	5.7	5.5	5.6
North East	33.7	34.6	33.3	29.7	29.2	29.4	32.0
South East	39.1	37.3	36.9	37.5	36.7	34.9	30.7
South	18.9	19.1	19.8	22.0	22.3	23.2	22.8
Center-West	3.9	4.5	4.7	5.2	6.1	7.1	9.0
Migrants (Percents)	70.4	69.1	68.8	66.4	63.2	59.4	57.6
Number of individuals	2939	4254	6585	8807	10052	10304	8288

Table 2.a: Wage equations by cohort, men. a), b)

	b1936_40		b1941_45		b1946_50		b1951_55		b1956_60		b1961_65		b1966_70	
Race														
Branca (Whites, omitted)														
Preta (Black)	-0.437	**	-0.257	**	-0.309	**	-0.287	**	-0.261	**	-0.178	**	-0.238	**
	(0.08)		(0.07)		(0.06)		(0.05)		(0.05)		(0.04)		(0.05)	
Amarela (Asians)	-0.013		0.293		0.503	**	0.070		0.305	*	0.143		-0.355	
	(0.32)		(0.16)		(0.16)		(0.16)		(0.15)		(0.15)		(0.22)	
Parda (MR)	-0.229	**	-0.300	**	-0.271	**	-0.228	**	-0.191	**	-0.174	**	-0.217	**
	(0.05)		(0.04)		(0.03)		(0.03)		(0.03)		(0.02)		(0.02)	
Parental schooling														
Mean parental schooling (years)	0.040	**	0.040	**	0.050	**	0.034	**	0.035	**	0.031	**	0.031	**
	(0.01)		(0.01)		(0.01)		(0.01)		(0.00)		(0.00)		(0.00)	
Mother/father difference (years)	0.006		0.015		0.003		0.008		-0.001		-0.001		0.008	
	(0.01)		(0.01)		(0.01)		(0.01)		(0.00)		(0.00)		(0.00)	
Region dummies														
South East (omitted)														
North	-0.217		-0.059		-0.125		-0.076		-0.192	*	-0.122	*	-0.166	*
	(0.12)		(0.11)		(0.08)		(0.07)		(0.06)		(0.05)		(0.06)	
North East	-0.207	**	-0.149	**	-0.108	**	-0.210	**	-0.162	**	-0.248	**	-0.212	**
	(0.05)		(0.04)		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)	
South	-0.180	**	-0.157	**	-0.118	**	-0.046		-0.043		-0.092	**	-0.096	**
	(0.06)		(0.05)		(0.04)		(0.03)		(0.03)		(0.03)		(0.03)	
Center-West	-0.166		-0.121		-0.002		-0.228	**	-0.062		-0.142	**	-0.046	
	(0.13)		(0.09)		(0.07)		(0.06)		(0.05)		(0.05)		(0.05)	
Years of schooling	0.089	**	0.072	**	0.080	**	0.075	**	0.061	**	0.031	**	0.029	**
	(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)	
Years of schooling-squared	0.002	*	0.003	**	0.002	**	0.003	**	0.003	**	0.005	**	0.004	**
	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
Migrant dummy	0.048		0.177	**	0.168	**	0.111	**	0.115	**	0.128	**	0.157	**
	(0.04)		(0.04)		(0.03)		(0.02)		(0.02)		(0.02)		(0.02)	
Sample size	1882		2605		3892		5045		5634		5812		4714	
Adjusted R-squared	0.400		0.418		0.450		0.421		0.409		0.421		0.362	

a) Dependent variable is the log of hourly wage rate. Regressions also include dummy variables for father's occupation. Coefficients are not reported because of space constraint. b) OLS estimates, standard errors in brackets; *=significant at the 5% prob. Level; **=significant at the 1% prob. level.

Table 2.b: Wage equations by cohort, women. [a), b)]

	b1936_40	b1941_45	b1946_50	b1951_55	b1956_60	b1961_65	b1966_70
Race							
Branca (Whites, omitted)							
Preta (Black)	0.152	-0.086	-0.125	-0.144 **	-0.208 **	-0.120	-0.047
	(0.12)	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)	(0.07)
Amarela (Asians)	0.513	-0.276	-0.060	0.184	0.453 *	0.256	-0.001
	(0.37)	(0.27)	(0.22)	(0.20)	(0.23)	(0.22)	(0.27)
Parda (MR)	-0.153 *	-0.169 **	-0.163 **	-0.225 **	-0.106 **	-0.095 **	-0.136 **
	(0.07)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Parental schooling							
Mean parental schooling (years)	0.060 **	0.049 **	0.053 **	0.026 **	0.038 **	0.041 **	0.041 **
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
Mother/father difference (years)	0.015	-0.004	0.001	-0.004	0.009	-0.002	0.003
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Region dummies							
South East (omitted)							
North	-0.197	-0.101	-0.063	0.016	-0.077	-0.091	-0.100
	(0.17)	(0.09)	(0.08)	(0.06)	(0.06)	(0.06)	(0.07)
North East	-0.107	-0.259 **	-0.244 **	-0.213 **	-0.304 **	-0.283 **	-0.271 **
	(0.07)	(0.05)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
South	-0.154 *	-0.091	-0.028	-0.079 *	-0.059	-0.011	-0.028
	(0.08)	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
Center-West	-0.331 **	-0.195	-0.162 *	-0.146 *	-0.077	-0.069	-0.117 *
	(0.14)	(0.11)	(0.08)	(0.07)	(0.06)	(0.05)	(0.05)
Years of schooling	0.051	0.069 **	0.026	0.006	0.016	-0.020	-0.017
	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Years of schooling-squared	0.003	0.003 **	0.006 **	0.008 **	0.007 **	0.009 **	0.009 **
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Migrant dummy	0.134 *	0.128 **	0.122 **	0.109 **	0.092 **	0.098 **	0.137 **
	(0.06)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)
Self-selection correction term	0.304	0.155	0.234	0.232	0.909 **	0.804 **	0.617 **
	(0.28)	(0.23)	(0.22)	(0.22)	(0.22)	(0.15)	(0.19)
Standard error of residual	0.811	0.740	0.785	0.773	0.909	0.858	0.818
Number of obs	1057	1648	2692	3760		4490	3573
Censored obs	249	316	426	505	4418	648	677

a) Dependent variable is the log of hourly wage rate. Regressions also include dummy variables for father's occupation. Coefficients are not reported because of space constraint. b) Two-stage Heckman estimates, standard errors in brackets; *=significant at the 5% prob. Level; **=significant at the 1% prob. level.

Table 3.a: Schooling determinants, men. [a), b)]

	b1936_40		b1941_45		b1946_50		b1951_55		b1956_60		b1961_65		b1966_70	
Race														
Branca (Whites, omitted)														
Preta (Black)	-1.148	**	-1.256	**	-0.909	**	-1.285	**	-1.560	**	-1.449	**	-0.987	**
	(0.33)		(0.29)		(0.25)		(0.22)		(0.21)		(0.20)		(0.21)	
Amarela (Asians)	5.947	**	2.874	**	2.151	**	2.786	**	1.726	**	2.471	**	-1.307	
	(1.29)		(0.66)		(0.69)		(0.70)		(0.67)		(0.69)		(0.98)	
Parda (MR)	-0.772	**	-1.224	**	-1.118	**	-1.101	**	-1.114	**	-1.213	**	-0.755	**
	(0.19)		(0.16)		(0.14)		(0.12)		(0.11)		(0.10)		(0.11)	
Parental schooling														
Mean parental schooling (years)	0.725	**	0.746	**	0.691	**	0.627	**	0.585	**	0.559	**	0.494	**
	(0.04)		(0.03)		(0.03)		(0.02)		(0.02)		(0.02)		(0.02)	
Mother/father difference (years)	0.072		0.004		0.048	*	0.009		-0.009		0.016		0.004	
	(0.04)		(0.04)		(0.03)		(0.02)		(0.02)		(0.02)		(0.02)	
Region dummies														
South East (omitted)														
North	-1.363	**	-0.721		-0.688	*	-0.431		-0.329		-0.446		-0.454	
	(0.48)		(0.45)		(0.34)		(0.29)		(0.25)		(0.25)		(0.26)	
North East	-0.993	**	-0.869	**	-0.662	**	-1.031	**	-1.095	**	-0.757	**	-1.006	**
	(0.19)		(0.16)		(0.14)		(0.13)		(0.12)		(0.11)		(0.12)	
South	0.031		-0.246		-0.628	**	-0.479	**	-0.323	**	-0.115		-0.355	**
	(0.23)		(0.21)		(0.17)		(0.14)		(0.12)		(0.12)		(0.13)	
Center-West	0.171		0.425		-0.422		0.355		-0.232		0.230		-0.276	
	(0.53)		(0.39)		(0.33)		(0.27)		(0.23)		(0.21)		(0.20)	
Sample size	1882		2605		3892		5045		5634		5812		4714	
Adj R-squared	0.432		0.437		0.436		0.409		0.437		0.422		0.349	

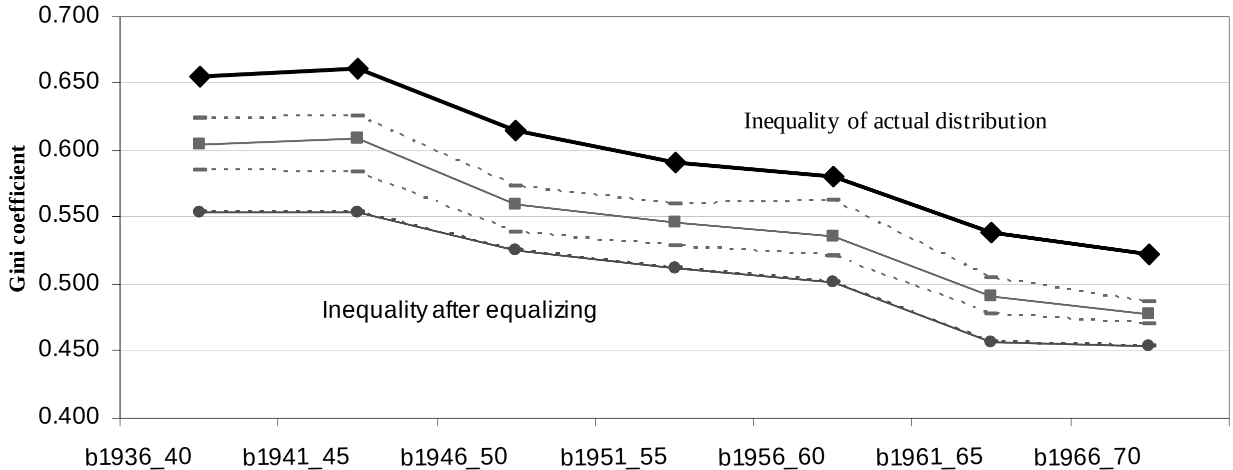
a) Dependent variable is years of schooling. Regressions also include dummy variables for father's occupation. Coefficients are not reported because of space constraint. b) OLS estimates standard errors in brackets; *=significant at the 5% prob. Level; **=significant at the 1% prob. level.

Table 3.b: Schooling determinants, women. [a), b)]

	b1936_40		b1941_45		b1946_50		b1951_55		b1956_60		b1961_65		b1966_70	
Race														
Branca (Whites, omitted)														
Preta (Black)	0.142	**	-1.154	**	-1.735	**	-0.950	**	-1.671	**	-1.588	**	-1.430	**
	(0.40)		(0.39)		(0.30)		(0.27)		(0.24)		(0.25)		(0.26)	
Amarela (Asians)	2.637	**	1.213		3.497	**	1.168		2.764	**	2.287	**	1.680	
	(0.99)		(1.02)		(0.82)		(0.84)		(0.80)		(0.80)		(0.93)	
Parda (MR)	-0.589	**	-0.898	**	-1.410	**	-1.343	**	-0.675	**	-1.097	**	-1.070	**
	(0.23)		(0.24)		(0.17)		(0.15)		(0.13)		(0.13)		(0.13)	
Parental schooling														
Mean parental schooling (years)	0.751	**	0.724	**	0.749	**	0.633	**	0.622	**	0.534	**	0.486	**
	(0.05)		(0.05)		(0.03)		(0.03)		(0.02)		(0.02)		(0.02)	
Mother/father difference (years)	0.088		-0.035		0.071	*	0.073	**	0.097	**	0.049	*	0.085	**
	(0.05)		(0.05)		(0.03)		(0.03)		(0.02)		(0.02)		(0.02)	
Region dummies														
South East (omitted)														
North	0.195		-0.164		0.039		0.417		0.098		-0.110		-0.029	
	(0.60)		(0.54)		(0.40)		(0.32)		(0.30)		(0.28)		(0.31)	
North East	-0.486	*	-0.285		-0.141		-0.141		-0.728	**	-0.325	*	-0.584	**
	(0.23)		(0.24)		(0.18)		(0.16)		(0.14)		(0.14)		(0.14)	
South	-0.194		-0.216		-0.492	*	-0.483	**	-0.502	**	-0.522	**	-0.730	**
	(0.28)		(0.27)		(0.20)		(0.17)		(0.15)		(0.14)		(0.15)	
Center-West	0.202		-0.108		-0.456		0.619		0.107		0.255		-0.143	
	(0.52)		(0.56)		(0.39)		(0.32)		(0.27)		(0.24)		(0.23)	
Sample size	1057		1648		2692		3760		4418		4490		3573	
Adj R-squared	0.469		0.364		0.444		0.384		0.396		0.400		0.362	

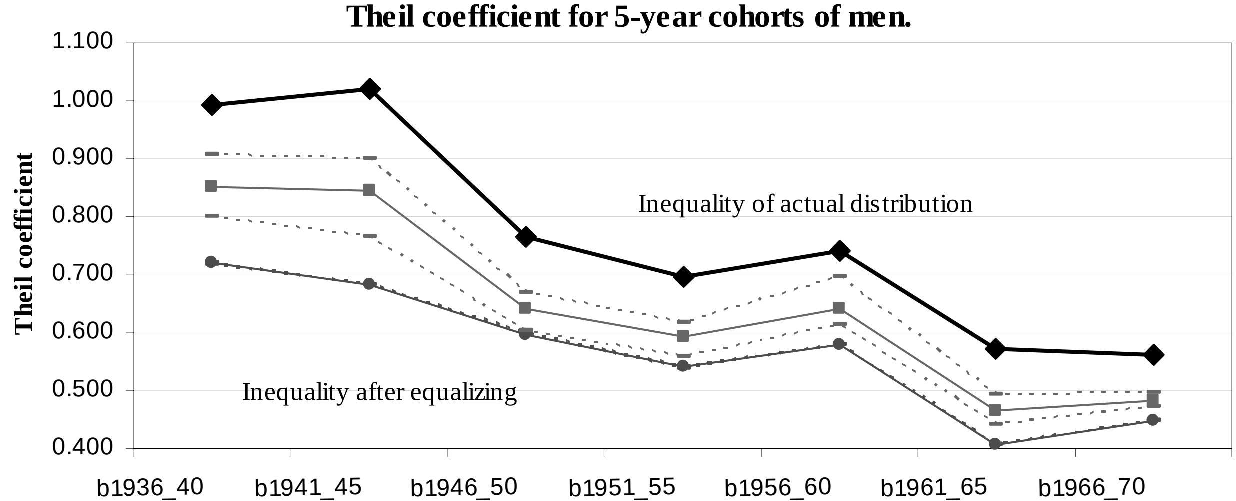
a) Dependent variable is years of schooling. Regressions also include dummy variables for father's occupation. Coefficients are not reported because of space constraint. b) OLS estimates standard errors in brackets; *=significant at the 5% prob. Level; **=significant at the 1% prob. level.

Figure 1a. Effects of equalizing circumstances on inequality (partial and complete effects).[a)]
Gini coefficient for 5-year cohorts of men.



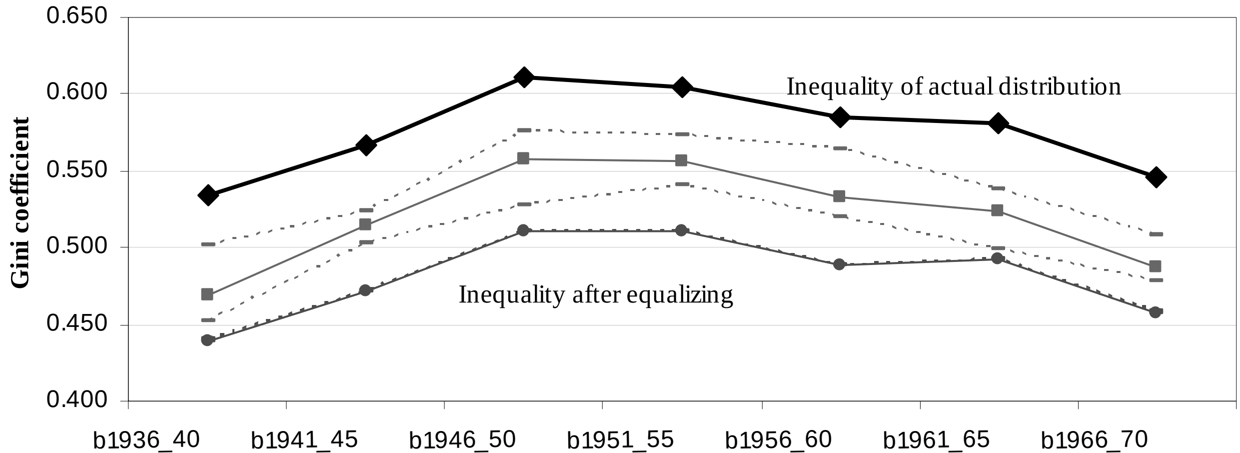
a) Partial and complete effect shown respectively on intermediate and bottom curves. Dotted lines correspond to upper and lower bounds.

Figure 1b. Effects of equalizing circumstances on inequality (partial and complete effects).[a)]
Theil coefficient for 5-year cohorts of men.



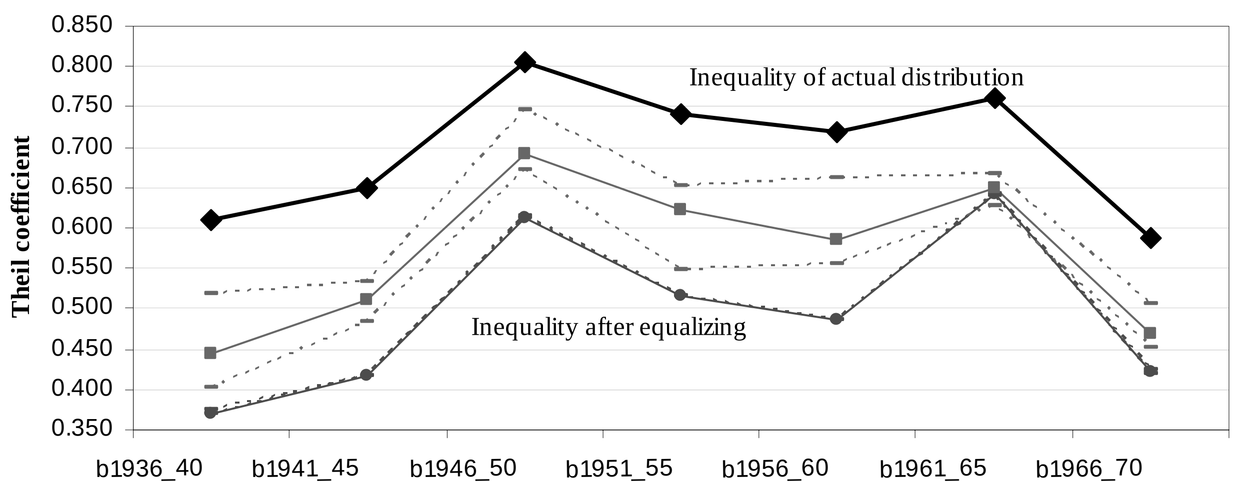
a) Partial and complete effect shown respectively on intermediate and bottom curves. Dotted lines correspond to upper and lower bounds.

Figure 2a. Effects of equalizing circumstances on inequality (partial and complete effects).[a)]
Gini coefficient for 5-year cohorts of women.



a) Partial and complete effect shown respectively on intermediate and bottom curves. Dotted lines correspond to upper and lower bounds.

Figure 2b. Effects of equalizing circumstances on inequality (partial and complete effects).[a)]
Theil coefficient for 5-year cohorts of women.



a) Partial and complete effect shown respectively on intermediate and bottom curves. Dotted lines correspond to upper and lower bounds.

Figure 3a: Complete effect of equalizing individual circumstance variables on inequality. Gini coefficient for 5-year cohorts of men.

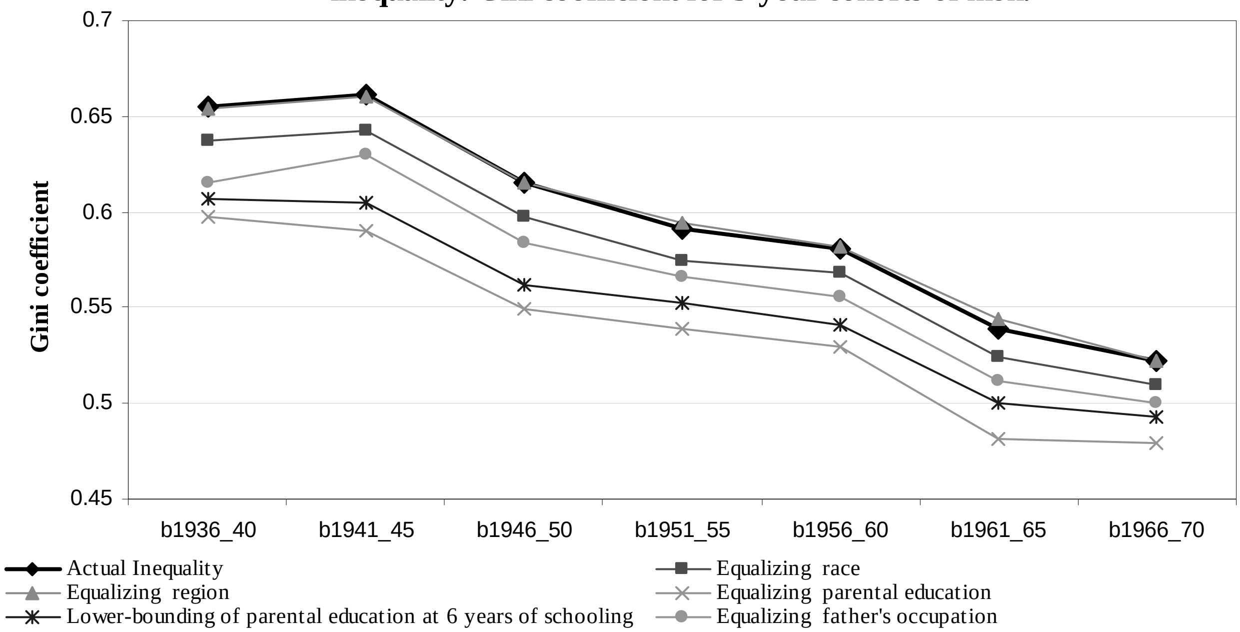


Figure 3b: Complete effect of equalizing individual circumstance variables on inequality. Gini coefficient for 5-year cohorts of men.

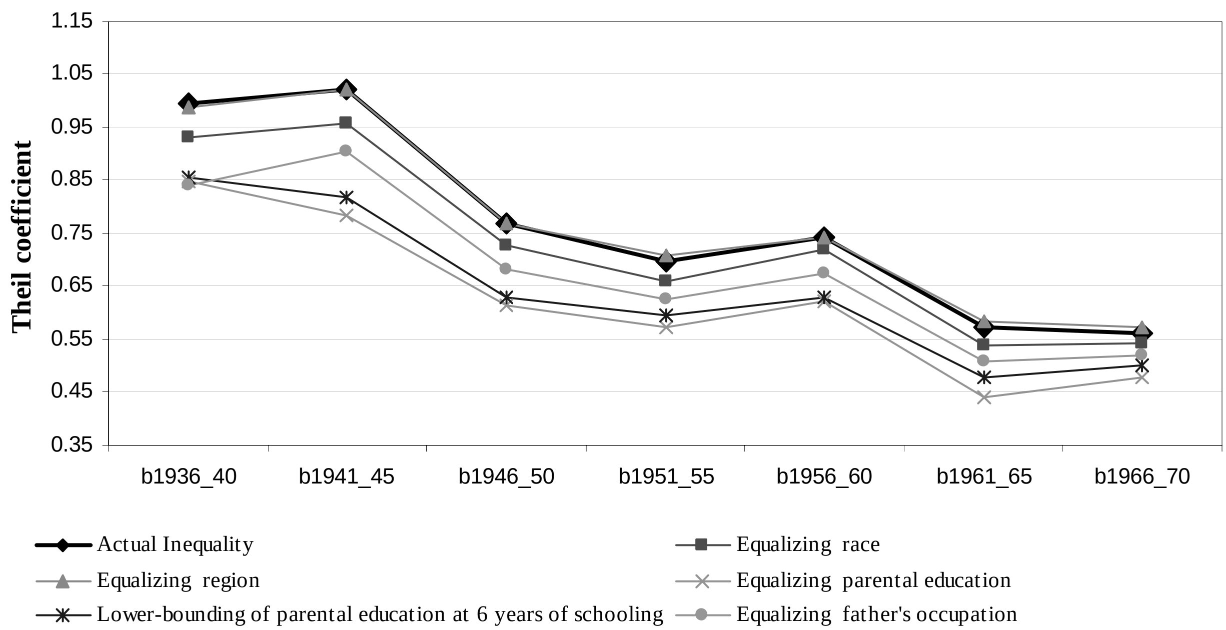


Figure 4a: Complete effect of equalizing individual circumstance variables on inequality. Gini coefficient for 5-year cohorts of women.

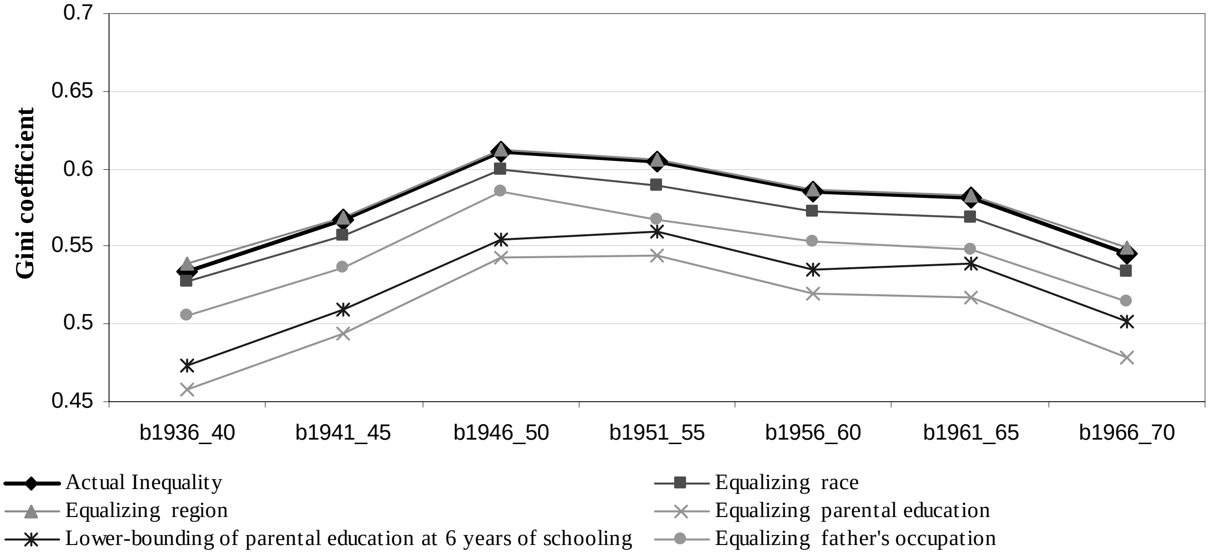


Figure 4b: Complete effect of equalizing individual circumstance variables on inequality. Theil coefficient for 5-year cohorts of women.

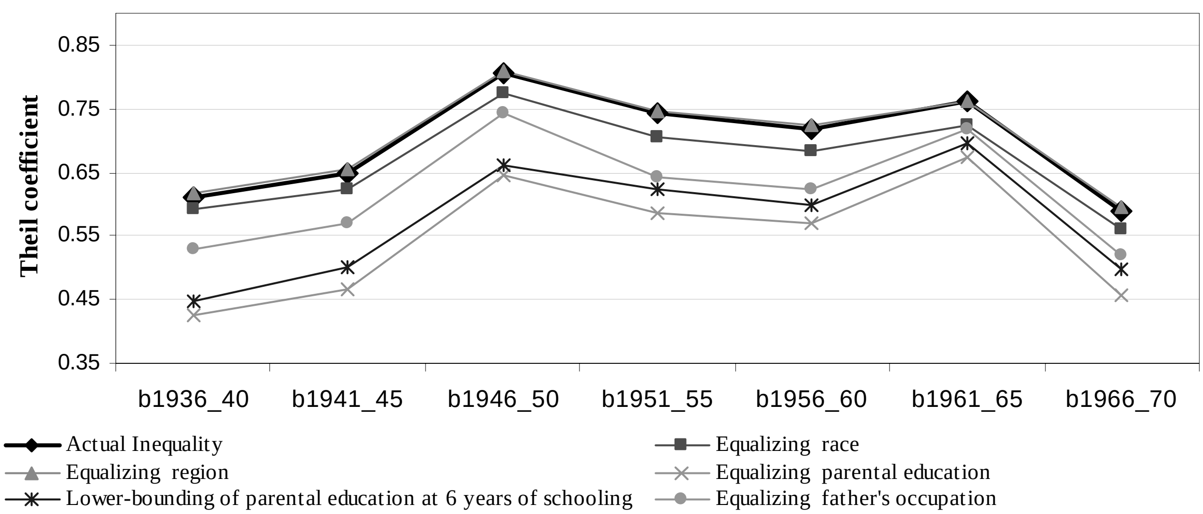


Table 4. Complete effects of equalizing circumstances on household income inequality : Gini and Theil coefficient for 5-year cohorts (adult men and women).

		b1936_40	b1941_45	b1946_50	b1951_55	b1956_60	b1961_65	b1966_70
Total Inequality	Gini	0.605	0.602	0.588	0.591	0.597	0.594	0.573
	Theil	0.75	0.736	0.682	0.72	0.709	0.691	0.635
Model simulated								
Model I: earnings effect (non-labor income, labor supply and household size kept constant).	Gini	0.534	0.521	0.505	0.504	0.509	0.5	499
	Theil	0.571	0.514	0.496	0.493	0.511	0.472	0.467
Model II : Earnings and household composition effects.	Gini	0.52	0.506	0.481	0.461	0.468	0.454	0.451
	Theil	0.516	0.473	0.43	0.402	0.419	0.388	0.382
Model III: all effects, reduced form.	Gini	0.474	0.492	0.481	0.489	0.49	0.478	0.482
	Theil	0.434	0.475	0.455	0.468	0.465	0.429	0.437

Table 5. International comparison of inequality of opportunities : share of the variance of economic status explained by parents characteristics. [a)]

Source	Definition of study and economic status	Number of studies	Range of R^2
Solon(1999)	Log wages , US studies based on (PSID)	11	.02-.28
	Log wages, US studies based on NLS	4	.02-.29
	Non-US studies	9	.01-.32
Mulligan (1999)	Years of schooling (various countries)	8	.02-.20
	Log earnings or wages (various countries)	16	.01-.35
	Log family income (various countries)	10	.02-.42
	Log family wealth (various countries)	9	.07-.58
	Log family consumption (US)	2	.35-.59
Dunn (2003)	Log earnings (Brazil)	1 (OLS, 2SLS: men only)	.28-.48
Ferreira & Veloso (2003)	Years of schooling (Brazil)	1 (OLS, different controls: men only)	.46-.66
This study			
	Years of schooling (Brazil)	6 cohorts/men-women	.35-.47
	Log earning rate (Brazil)[b)]	6 cohorts/men-women	.25-.30
	Log family income per capita (Brazil)[b)]	6 cohorts	.32-.44

a) Solon, Mulligan, Dunn and Ferreira & Veloso originally report the value of the intergenerational elasticity: the β coefficient in the Galtonian regression: economic status = $\alpha+\beta$*economic status of parents + residual. This estimated coefficient is close to the correlation coefficient between the economic status of parents and children, under the assumption that the variance of economic status is approximately constant across generations. For comparability with our results, the table reports the square of their estimated coefficient, which is directly comparable to R^2 statistics in the regressions undertaken in this paper.

b) R^2 statistics are those associated with the reduced form model where log earnings or log income per capita are regressed on circumstance variables only. These regressions are not shown here.

Departamento de Economia PUC-Rio
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marques de Sâo Vicente 225 - Rio de Janeiro 22453-900, RJ
Tel.(21) 31141078 Fax (21) 31141084
www.econ.puc-rio.br
flavia@econ.puc-rio.br