The Measurement of Inequality of Opportunity:
Theory and an application to Latin America

Francisco H. G. Ferreira and Jérémie Gignoux[♦]
June 10th, 2008

Keywords: Inequality of opportunity; Latin America

JEL Codes: D31, D63, J62.

Abstract: What part of the inequality observed in a particular country is due to unequal opportunities, rather than to differences in individual efforts or luck? This paper estimates a lower bound for the opportunity share of inequality in labor earnings, household income per capita and household consumption per capita in six Latin American countries. Following John Roemer, we associate inequality of opportunity with outcome differences that can be accounted for by morally irrelevant pre-determined circumstances, such as race, gender, place of birth and family background. Thus defined, unequal opportunities account for between 24% and 50% of inequality in consumption expenditure in our sample. Brazil and Central America are more opportunity-unequal than Colombia, Ecuador, or Peru. "Opportunity profiles", which identify the social groups with the most limited opportunity sets, are shown to be distinct from poverty profiles: ethnic origin and the geography of birth are markedly more important as determinants of opportunity deprivation than of outcome poverty, particularly in Brazil, Guatemala and Peru.

[♦] Development Research Group, The World Bank. We are grateful to Caridad Araujo, Ricardo Paes de Barros, Marc Fleurbaey, Markus Goldstein, Peter Lanjouw, Marta Menéndez, Jose Molinas, John Roemer and Jaime Saavedra for helpful discussions and comments on an earlier draft. Helpful comments were also received in conferences or seminars at the World Bank, the IDB, the Catholic University of Milan, Universidad de los Andes in Bogotá, Colégio de México, and the Universities of Manchester and Lund. We also thank Carlos Becerra, Jofre Calderón, Leo Gasparini, and Luis Rubalcava for kindly providing us with access to data. The views expressed in the paper are those of the authors, and should not be attributed to the World Bank, their Executive Directors, or the countries they represent. Correspondence to fferreira@worldbank.org and jgignoux@worldbank.org.

1. Introduction

Economic inequality – usually measured in terms of income or consumption – is neither all bad nor all good. Most people view income gaps that arise from the application of different levels of effort as less objectionable than those that are due, say, to racial discrimination. Attitudinal surveys attest to this. When asked to place their views on a scale from 1 to 10, where 1 implied agreement with the statement that "Incomes should be made more equal," and 10 implied agreement with the statement that "We need larger income differences as incentives for individual effort", respondents in the 1999-2000 World Value Survey were evenly divided.[1] The median answer was 6. The two modes of the distribution, with approximately 20% of respondents each, were 1 and 10.

Attitudes to inequality vary for a number of reasons, but an important factor is whether inequalities are seen to be driven by differences in factors for which the individual can be held morally accountable (i.e. where he or she had a choice), or by factors that lie beyond the individual's responsibility. In an influential contribution, John Roemer (1998) calls the former "efforts", and the latter "circumstances". He describes "equality of opportunity" as a situation in which important outcomes – which he calls "advantages", and which would include measures of economic welfare such as earnings or household consumption – are distributed independently of circumstances. A situation, in other words, where the distribution of economic welfare *within* groups of people with identical circumstances would not vary *across* such groups.[2]

The distinction between inequality of opportunity and the more standard concept of inequality of outcomes is of interest to economists for at least three sets of reasons. First, if inequality of opportunity does affect attitudes to outcome inequality, then it may affect attitudes to redistribution and beliefs about social fairness. These beliefs may in turn affect the extent of redistribution actually implemented, and the level of investment and output generated. Alesina and Angeletos (2005) and Bénabou and Tirole (2006) are examples of models where such beliefs and attitudes play a key role in generating multiple equilibria with very different objective economic characteristics.

Second, there is a widespread normative view that inequality of opportunity matters for the design of public policy, since only differences due to opportunities should be the object of compensation by the state. This is the view in Arneson (1989), Roemer (1998) and Peragine (2004), to mention but a few. Third, it has also been suggested that inequality of opportunity might be a more relevant concept (than income inequality) for understanding whether aggregate economic performance is worse in more unequal societies (and if so, why). In addition to the role of beliefs and attitudes to redistribution, it is possible that the kinds of inequality that are detrimental to growth (such as inequality in access to good schools, or to financial markets) are more closely associated with the concept of opportunities, while other components of outcome inequality – such as those

---

[1] The World Value Survey is conducted by the Inter-university Consortium for Political and Social Research, based at the University of Michigan, and contains responses from representative samples in 69 countries.

[2] Roemer (1998) was not, of course, the first economist or philosopher to argue that the space of opportunities was ethically the right one to focus on. Arneson (1989), Cohen (1989) and, to some extent, Sen (1985) had made broadly similar points. By providing a simple, yet powerful, formalization of the definition of equal opportunities, however, Roemer contributed to an increase in interest in the concept from applied economists.

arising from differential returns to effort – may actually have a positive effect on growth (World Bank, 2006; Bourguignon et al. 2007). Perhaps one of the reasons why the cross-country empirical literature on inequality and growth is so inconclusive is that it conflates the two kinds of inequality.[3]

But in order to make any empirical use of the concept of inequality of opportunity, whether in the design of taxation and public expenditures or in the study of the determinants of cross-country growth differences, it is first necessary to measure it. Some progress in that direction has been made. Bourguignon et al. (2003, 2007) parametrically estimate inequality of opportunity for various cohorts in Brazil, in 1996. Checchi and Peragine (2005) apply a non-parametric decomposition to measure inequality of opportunity for both income and educational achievement in Southern and Northern Italy.[4] Lefranc et al. (2006) use stochastic dominance rankings to compare the degree of inequality of opportunity among a set of OECD countries.[5] Barros et al. (2008) associate inequality of opportunity for children with unequal access to a set of basic services, and compute indices for a set of countries in Latin America. Cogneau et al. (2006) apply a variant of the Bourguignon et al. (2007) approach to a set of African countries.

Nevertheless, the empirical study of inequality of opportunity remains a nascent – though increasingly vibrant – field. This paper aims to make three contributions, the first two of which are methodological. First, we provide a simple conceptual framework which derives a class of indices of inequality of opportunity directly from Roemer's theory. The parametric measure proposed by Bourguignon et al. (2007) and the non-parametric indices in Checchi and Peragine (2005) are shown to be members of this class, which can therefore be seen as a unifying concept in the measurement of inequality of opportunity. Indices within the class differ along two dimensions: *decomposition path* and *estimation procedure*. Drawing on the earlier literature on path dependence in inequality decomposition (Foster and Shneyerov, 2000), we show that there exists a unique inequality index (the mean log deviation) for which our measure of inequality of opportunity is path-independent.[6] For that index, this class of measures collapses to a parametric and a non-parametric alternative along the estimation procedure dimension. We show that the two methods provide a narrow range of lower-bound estimates for inequality of opportunity in a set of six Latin American countries.

Second, we introduce the concept of an *opportunity-deprivation profile*: a vector of characteristics of the groups with the most limited opportunity sets in a given society (a precise definition follows). Following Roemer's (2006) suggestion that "the rate of economic development should be taken to be the rate at which the mean advantage level of the worst-off types grows over time." (p.243), we compare these profiles across our sample of six Latin American countries. We also compare these profiles to the analogous poverty profiles, and suggest an interpretation of the differences.

[3] See Forbes (2000) and Banerjee and Duflo (2003).
[4] See also Cogneau and Gignoux (2007), on earnings in Brazil.
[5] See also Hild and Voorhoeve (2004) on the philosophical implications of using stochastic dominance criteria for evaluating the extent of inequality of opportunity.
[6] Strictly, this uniqueness is within the set of inequality measures that satisfy the transfer principle, and use the arithmetic mean as the representative income (see Foster and Shneyerov, 2000).

The third contribution is substantive. We apply these two methodological innovations to a rich set of household data for six countries in Latin America: Brazil, Colombia, Ecuador, Guatemala, Panama, and Peru. In each case, we observe information on six "circumstance" variables, namely gender, race or ethnicity, place of birth, mother's education, father's education and father's occupation. We are not aware of a comparable data set being previously used for the comparative analysis of inequality of opportunity anywhere in the world.

The paper presents parametric and non-parametric estimates of our path-independent measure of inequality of opportunity for three distinct indicators of economic advantage – labor earnings, household per capita income, and household per capita consumption – and discusses the significant differences among them. A number of interesting cross-country patterns appear, both with respect to overall levels of inequality of opportunity, and to the relative importance of individual circumstance variables. Brazil, Guatemala and Panama are found to be systematically more opportunity-unequal than Colombia, Ecuador and Peru. Ethnic inequalities are also stronger in Brazil and in the two Central American countries, whereas geographic inequalities are greater in the two Central American countries and in Peru.

At the lower bound, inequality of opportunity is found to account for a substantial share of observed economic inequality in Latin America. For inequality in household consumption expenditures per capita, for instance, the (parametrically estimated) opportunity share ranges from 24% to 50%, depending on the country. The results are different for earnings and for household incomes, reflecting differences both in the economic mechanisms through which circumstances affect outcomes, and in questionnaire design and likely measurement error. The opportunity profiles also differ substantially among countries, with ethnicity being fundamental in Brazil but much less important in Colombia, for instance. Opportunity profiles also differ from poverty profiles, reflecting the fact that circumstances matter, but are not destiny: effort and luck enable some of those born in opportunity-disadvantaged groups to escape poverty, while others - born to more advantaged groups - fall into it.

The remainder of the paper is structured as follows. Section 2 provides a conceptual framework by deriving a simple class of measures of inequality of opportunity from Roemer's definition of the concept. Section 3 describes four alternative members of that class that can be estimated in practice, and discusses their properties. Section 4 provides some information on the six household survey data sets used in the analysis. Section 5 reports the results of the alternative estimation procedures for labor earnings. Section 6 presents the results for household welfare, based on per capita income and consumption expenditure distributions. Section 7 discusses the opportunity profiles for all six countries, and how they compare with the poverty profiles. Section 8 concludes.

2. A Conceptual Framework

A natural approach to measuring inequality of opportunity would begin from Roemer's (1998) distinction between "circumstance" and "effort" variables. Following Bourguignon et al. (2007), consider a "model of advantage" of the general form:

$$y = f(C, E, u) \qquad (1)$$

where y denotes the outcome of interest (Roemer's "advantage"); C denotes a vector of circumstance variables; E denotes a vector of effort variables; and u denotes pure luck or other random factors. Roemer's theory explicitly requires that circumstances be economically exogenous (in the sense that the individual has no control over them).[7] But it also explicitly allows for the fact that efforts may be endogenous to circumstances. For example: one can not change one's race, or the family one is born into, but those factors can and do affect one's educational and work choices. Incorporating the fact that efforts are endogenous and may thus depend on circumstances, (1) can be rewritten as:

$$y = f\left[C, E(C,v), u\right] \qquad (2)$$

Roemer's definition of equality of opportunity essentially requires that $F(y|C) = F(y)$, which in turn implies three conditions:

(i) $\frac{\partial f(C,E,u)}{\partial C} = 0, \forall C$, i.e. no circumstance variable should have a direct causal impact on y;

(ii) $G(E|C) = G(E), \forall E, \forall C$, each effort variable should be distributed independently from all circumstances.[8]

To measure inequality of opportunity is therefore to measure the extent to which $F(y|C) \neq F(y)$. An obvious first step would be to test for the existence of inequality of opportunity, by examining whether the conditional distributions $F(y|C)$ differ across the elements of C. This is precisely what Lefranc et al. (2006) do, using stochastic dominance concepts and the associated statistical tests to compare the distribution of opportunities across a number of OECD countries. Theirs is a very interesting approach to ascertaining whether or not individual countries could be described as having equality of opportunity. It also allows for a (partial) ranking of *types* (groups with identical circumstances) within each country. As always, though, greater robustness in ranking comes at a price. Testing for dominance across cumulative distribution functions for different types does not permit a quantification of how far those groups are from one another. Consequently it does not really allow for a ranking of inequality of opportunity across countries, beyond a binary classification into "equal" or "unequal".

In this paper, we follow a complementary approach and seek to construct scalar indices of inequality of opportunity, based on partitioning the population by circumstance categories. Given agreement on a particular vector of circumstance variables $\boldsymbol{C}$, define $\{y_i^k\}$ as a partition of the distribution such that $C_i^k = C^k \Leftrightarrow i \in k, k = 1,...,K$.[9] $\{y_i^k\}$ is then a partition of the population into K groups, such that the members of each group are

---

[7] We write "economically exogenous" to distinguish the original meaning of the term from the common econometric usage, which refers to a correlation between the variable and the residual term. In the case of circumstance variables, econometric endogeneity could arise from the existence of omitted variables, but *not* from reverse causation.

[8] A third condition, which holds by assumption, is $H(u|C) = H(u)$, i.e. random factors and luck are also independent from circumstances. F, G and H denote cumulative distributions. For simplicity, we omit subscripts for individual elements of the circumstance and effort vectors, and the corresponding proliferation of notation for the distributions. See Bourguignon, Ferreira and Menéndez (2007) for a related discussion.

[9] It must be the case, of course, that $K \leq N$, where N is the size of the population.

identical with respect to all circumstances in the vector ***C***. The set of individuals $\mathrm{T}_k : i | i \in k$ is simply what Roemer would refer to as *type k*. Defining the partition $\{y_i^k\}$ requires agreement on a vector ***C***, for which the joint distribution $F(y, C)$ is observed, as well as agreement on the specific partitioning within each variable: for example, how finely the vector of mother's years of schooling, or the spatial location of birth, are to be subdivided. We are looking for a scalar measure $\theta : \{y_i^k\} \rightarrow \Re_+$ that captures the degree of inequality of opportunity in the partition.[10]

Next, note that for any meaningful definition of between-group inequality, stochastic independence implies:

$$F(y|C) = F(y) \Rightarrow IB(\{y_i^k\}) = 0 \qquad (3)$$

where $IB(\{y_i^k\})$ denotes the between-group component of inequality over the previously constructed partition of the population.[11] It follows that two natural candidates for $\theta : \{y_i^k\} \rightarrow \Re_+$ would be indices of the form:

$$\theta(\{y_i^k\}) = IB(\{y_i^k\}) \qquad (4)$$

or $$\theta(\{y_i^k\}) = \frac{IB(\{y_i^k\})}{I(F(y))} \qquad (5)$$

Equation (4) defines a measure of inequality of opportunity as the *absolute* level of the inequality between groups in a population, where those groups arise from an agreed partition of the population, so that members of each group share identical circumstances, in Roemer's sense. Equation (5) defines it as the same between-group inequality, *relative* to overall inequality in the population. As a relative measure, (5) is actually a mapping $\theta : \{y_i^k\} \rightarrow [0,1]$, for any decomposable inequality index *I()*.[12]

As in other contexts (like simple poverty and inequality measurement), absolute and relative measures convey different information, and rank populations differently. Both are useful, and should be seen as complementary. In what follows, we focus on the relative-Θ class, largely to economize on space, but both the relative and the absolute measures may be of interest. The methodological points in the remainder of this section can be easily extended to the absolute-Θ class, in a perfectly analogous fashion.

3. Calculating Relative-Θ Measures in Practice

It is well-known from the inequality decomposition literature that $IB(\{y_i^k\})$ is not a uniquely defined object, even if attention is confined to inequality indices that are

---

[10] More formally, one could write $\theta : \Im(y, C) \times \Pi_C \rightarrow \Re_+$, where $\Im(y, C)$ denotes the space of joint distribution functions of *y* and ***C***, while Π_C denotes the set of possible partitions of a population by the elements of ***C***. This recognizes that the notation $\{y_i^k\}$ conflates two components: a joint distribution of *y* and ***C***, and a specific partition of the population by the elements of ***C***.

[11] The converse statement does not hold, as the inexistence of between-group inequality is a much weaker condition than stochastic independence.

[12] On decomposable inequality indices, see Bourguignon (1979) or Shorrocks (1980).

properly decomposable.[13] In fact, for a given $\{y_i^k\}$, estimates of between-group inequality can differ for three reasons: (i) the specific inequality index *I()* used in the decomposition; (ii) the *path* of the decomposition; and (iii) the decomposition *procedure*, i.e. whether it is estimated parametrically or non-parametrically.

Point (i) is well-established in the literature. The decomposition of inequality by population subgroup for a given distribution and partition will differ across different members of the Generalized Entropy or Atkinson families. To see point (ii), following Foster and Shneyerov (2000) and Checchi and Peragine (2005), define:

- a *smoothed distribution* $\{\mu_i^k\}$, corresponding to a particular partition $\{y_i^k\}$, as the distribution that arises from replacing y_i^k with the group-specific mean μ^k.

- a *standardized distribution* $\{v_i^k\}$ corresponding to a particular partition $\{y_i^k\}$ as the distribution that arises from replacing y_i^k with $y_i^k \dfrac{\mu}{\mu^k}$ (where μ is the grand mean).

Since a smoothed distribution eliminates all within-group inequality by construction, a first member of the relative-Θ class immediately suggests itself as $\theta_d = I(\{\mu_i^k\})/I(\{y_i^k\})$. θ_d is simply the ratio of inequality in the smoothed distribution to the inequality in the original distribution. It summarizes between-group inequality in the partition *directly*, hence the subscript *d*.

A standardized distribution, on the other hand, suppresses all between-group inequality, leaving only inequality within-groups. $\theta_r = 1 - I(\{v_i^k\})/I(\{y_i^k\})$, one minus the ratio of inequality in the *standardized distribution* to the inequality in the original distribution, is therefore another perfectly plausible measure of inequality of opportunity. It computes the between-group inequality in the partition *residually*, hence the subscript *r*. [14]

Unfortunately, although there is no obvious reason why either one of these two paths should be preferred to the other, for most inequality indexes they will yield different measures of the share of between-group inequalities (and thus of inequality of opportunity).[15] Foster and Shneyerov (2000) characterize the class of inequality measures

---

[13] The literature on inequality measurement (e.g. Bourguignon 1979, Shorrocks 1980) has established that total inequality is only additively separable into a between-group component and a within-group component for some indices. The best-known family of additively decomposable measures is the generalized entropy class, which includes the mean log deviation (E(0)) and the Theil entropy index (E(1)). The Gini coefficient is not additively decomposable in the same way. See also Elbers et al. (2008).

[14] This straightforward, non-parametric decomposition is very similar to Checchi and Peragine (2005), who proposed either to re-weight the distributions of outcome in order to equalize the means of the different circumstances groups (in a "types approach") or to reweight the means of the individuals who can be considered as having exerted the same efforts (in a "tranches approach").

[15] It is easy to see why the two decomposition paths yield different results for other generalized entropy measures. The decomposition of total inequality for these measures can be written as follows:

$$I(x^1,...,x^m) = \sum_{k=1}^{m} \frac{n^k}{n}\left(\frac{\mu^k}{\mu}\right)^c I_c(x^k) + I_c(x^B) = IW + IB$$

(eq. 3 in Foster and Sneyerov (2000))

where x^k for k=1,…,m, denote the distribution of outcomes within each sub-group, x^B is the smoothed distribution, n the total population, n^k the population of sub-group k, μ the overall mean, and μ^k the for which the two methods yield the same results; namely the "*path-independent decomposable*" class of inequality measures. They show that when the set of inequality indices under consideration is restricted to those that use the arithmetic mean as the reference income, and that satisfy the Pigou-Dalton transfer axiom, this class reduces to a single inequality measure, the mean log deviation, or E(0).

This has a helpful implication for our attempt to measure inequality of opportunity using (4) or (5). It implies that, if we are prepared to adopt the Foster-Shneyerov path-independence axiom, then the first two of the three previously mentioned reasons for estimates of between-group inequality to vary are eliminated: if we focus on path-independent measures, we must use the mean log deviation as our inequality index I(), and $\theta_d = \theta_r$. Differences due both to the use of different aggregation indices and to path-dependence are eliminated simultaneously.[16]

If one is interested only in an overall estimate of $\theta(\{y_i^k\}) = \frac{IB(\{y_i^k\})}{I(F(y))}$, and if one's sample is sufficiently large relative to the number of cells in the partition $\{y_i^k\}$, then we need go no further: the between-group share of inequality for E(0), in a partition defined by a vector of circumstances, is our single scalar estimate of a lower bound for inequality of opportunity. Unfortunately, however, the richer the information set on people's circumstances, the more cells one would like to include in the partition. As cell numbers increase, cell sizes diminish, leading to the classic problem of data insufficiency for non-parametric estimation. This has led some authors to propose parametric alternatives to the estimation of θ_d and θ_r.

In order to construct these alternatives, define a *parametrically standardized distribution* $\{\tilde{y}_i\}$, corresponding to *F(y, **C**)*, as the distribution that arises from replacing y_i with $\tilde{y}_i = f[\overline{C}, E(\overline{C}, v_i), u_i]$, where the upper bar on the vector ***C*** denotes the vector of sample mean circumstances.[17]

To obtain this counterfactual distribution, a specific model of (2) must be estimated. Once this has been done, $\{\tilde{y}_i\}$ is obtained simply by replacing the individual circumstance values in (2) with the sample average for each circumstance variable. A

---

sub-group mean. c is a parameter for each GE measure. IB is the between-group component. IW is a weighted sum of the inequality measured within each sub-group. The problem is that, for $c \neq 0$, the within-group component is affected by changes in μ^k. Standardizing the distribution will therefore not only eliminate IB. It will also affect IW, by changing the weights associated with the inequality within each sub-group.

[16] There is a sense in which measures based on the smoothed distribution correspond to van de Gaer's concept of "min-of-means", because they focus exclusively on information about the mean advantage levels of each type, while the measures derived from the standardized distribution are closer to Roemer's "mean-of-mins" approach, in the sense that it first removes differences between the types, and then compares the ("relative effort") distributions within each type. See Fleurbaey (2008). In this paper, we deliberately gloss over that philosophical debate, other than to note that we identify a particular measure, axiomatically derived, for which both approaches yield identical results. We see that as an added advantage of our approach.

[17] This is (parametrically) analogous to the standardized distribution because, by giving each and every individual the same circumstance variables, it eliminates any inequality between groups that are associated with circumstances.

variety of alternative specifications for (2) is possible, of course. Bourguignon et al. (2007) use a log-linear/linear specification of the form:

$$\begin{aligned} \ln y &= C\alpha + E\beta + u \\ E &= BC + v \end{aligned} \qquad (6)$$

The reduced form of (6) is $\ln y = C(\alpha + \beta B) + v\beta + u$, which can be estimated by OLS as $\ln y = C\psi + \varepsilon$. (7)

Under these functional form assumptions, the parametrically standardized distribution is estimated by $\hat{\tilde{y}}_i = \exp[\overline{C}_i \hat{\psi} + \hat{\varepsilon}_i]$.

Analogously, define a *parametrically smoothed distribution* $\{\tilde{z}_i\}$, corresponding to *F(y,* ***C****)*, as the distribution that arises from replacing y_i with $\tilde{z}_i = f[C, E(C)]$, where the error term of the model is suppressed. This counterfactual distribution is also obtained by estimating a specific parametric model for (2), and suppressing within-group inequality by replacing y_i with its prediction, given the vector of circumstances $\boldsymbol{C}$. In a reduced-form framework, and under the functional form assumptions above, the parametrically smoothed distribution is estimated by $\hat{\tilde{z}}_i = \exp[C_i \hat{\psi}]$.

This therefore allows us to define $\theta_r^P = 1 - I(\{\tilde{y}_i\})/I(\{y_i^k\})$ as a parametric alternative to $\theta_r^N = 1 - I(\{v_i^k\})/I(\{y_i^k\})$; and $\theta_d^P = I(\{\tilde{z}_i\})/I(\{y_i^k\})$ as a parametric alternative to $\theta_d^N = I(\{\mu_i^k\})/I(\{y_i^k\})$, where the superscripts now refer to the (parametric or non-parametric) estimation procedure.[18]

Unlike θ_d^N and θ_r^N, θ_d^P and θ_r^P rely on specific functional form assumptions. In addition to the possible sample-size insufficiency for non-parametric estimation, there is another reason why the costs of such a parametric approximation may be worth bearing: the parametric approach permits the estimation of the partial effects of one (or a subset) of the circumstance variables, controlling for the others, by constructing alternative counterfactual distributions, such as:

$$\hat{\tilde{y}}_i^J = \exp[\overline{C}_i^J \hat{\psi}^J + C_i^{j \neq J} \psi^{j \neq J} + \hat{u}_i] \qquad (8)$$

in the case of a parametrically standardized decomposition. This allows us to compute circumstance J-specific inequality shares:

$$\theta_r^J = 1 - I(\{\tilde{y}_i^J\})/I(\{y_i^k\}) \qquad (9)$$

The existence of this trade-off between parametric and non-parametric methods - with non-parametric decompositions being more flexible (with no functional form assumptions) but more data intensive, while the parametric approach is less data intensive but relies on (potentially restrictive) functional form assumptions – points to two things: the need to investigate their comparative performance on the same data set (as we do below), and the possibility that the methods may usefully complement each other.

A final methodological consideration refers to the interpretation of these measures, given the realistic possibility that not all relevant circumstance variables may be observed. A "true" measure of inequality of opportunity, as conceptually defined in equations (3) – (5), would require that all relevant circumstance variables be included in

[18] The measure computed for Brazil by Bourguignon et al. (2007) was an example of a parametric residual measure: θ_r^P.

the vector $\boldsymbol{C}$. This is unlikely to be the case in practice for almost any conceivable data set, and certainly for the six countries we study below. The empirical estimates defined in this section – whether direct or residual, and parametric or not – should be interpreted as lower-bound estimates of inequality of opportunity. To see why they are lower bounds, note that including an additional element in vector $\boldsymbol{C}$ causes each and every cell in the partition $\{y_i^k\}$ to be further subdivided (into at least another two cells). This can not lower the between-group inequality share and, unless the additional element is orthogonal to the measure of advantage, will raise it.

An additional reason why the measure is a lower-bound is that the partitioning of the population into categories *within* each circumstance variable in $\boldsymbol{C}$ is very coarse in this paper. An example is the classification of parental occupations into only two cells: "agricultural worker" or "other". For most circumstance variables, international comparability required aiming for "common denominator", relatively aggregated classifications. Like adding other circumstance variables, further subdivision of these categories within each circumstance might also increase (but could not reduce) the share of inequality attributed to opportunities.

Similarly, in the parametric case, notice that equation (7) is a reduced-form specification, intended to capture both the direct and indirect effects of circumstances on advantage. Adding another element of the vector $\boldsymbol{C}$ to this specification (or further refining the set of dummy variables for each circumstance) must reduce the variance of the residual and increase the variance (or any other inequality measure) accounted for by the set of observed circumstances.[19]

In what follows, we calculate three of the four members of the relative-Θ class – namely θ_d^N, θ_r^N and θ_r^P - for the distributions of earnings, income and consumption in recent household surveys from six Latin American countries. We show that the two non-parametric indices differ for decomposable inequality measures other than E(0), but focus our discussion on that index, for which $\theta_d^N = \theta_r^N$. We then examine the differences between the non-parametric and parametric estimates. Although, as expected, these differences are larger for smaller sample sizes, they are generally quite small, suggesting that our path-independent estimates of inequality of opportunity are also methodologically robust to the choice of estimation procedure. Before presenting the results for each concept in Sections 5 and 6, the next section briefly describes the data sets.

4. The Data and the Partition by Circumstances

We use data from six nationally representative household surveys in Latin America, namely the Brazilian *Pesquisa Nacional por Amostra de Domicílios* (PNAD)

[19] A possible misunderstanding would be to argue that, because certain omitted circumstances might be negatively correlated with the observed vector $\boldsymbol{C}$, the parametric measure need not be a lower bound. It is of course possible that the share of inequality attributed to a specific set of (observed) circumstances is overestimated. This might happen if omitted circumstance variables are negatively correlated with the observed ones. But the R^2 of regression (7) can not fall by including these other circumstance variables, so that the estimate *is* a lower-bound for the share of inequality attributed to *all* circumstances (rather than to the observed subset), analogously to the non-parametric case.

1996; the Colombian *Encuesta de Calidad de Vida* (ECV) 2003; the Ecuadorian *Encuesta Condiciones de Vida* (ECV) 2006; the Guatemalan *Encuesta Nacional sobre Condiciones de Vida* (ENCOVI) 2000; the Panamanian *Encuesta de Niveles de Vida* (ENV) 2003; and the Peruvian *Encuesta Nacional de Hogares* (ENAHO) 2001. The ENAHO and the PNAD are original national surveys, while the others are LSMS-type surveys. This particular group of surveys were selected for containing information on family background and, more specifically, on parents' education, father's occupation, or both.[20]

In all countries, we restrict the sample to individuals aged 30 to 49, which are the cohorts with the highest proportion of employed persons.[21] Sample sizes for each survey, both before and after excluding observations with missing data, are reported in Table 1. Sample sizes with complete information range from about 6,000 (for Panama) to 72,000 observations (for Brazil), for the analysis of income and consumption, and from about 4,000 to 50,000 for the analysis of earnings (when only employed individuals are retained).[22]

The surveys contain information on a common set of circumstances: (a) three variables related to family background: father's and mother's education and father's occupation during the person's childhood; (b) ethnicity (or race); and (c) region of birth (or type of area of birth). The only exceptions are that the father's occupation variable is not available for Colombia and Peru, and results must be interpreted with this caveat in mind. We also use gender as a circumstance variable in the analysis of earnings. Parental education variables are coded into three categories: no education (or unknown), primary (incomplete or complete, depending on the country), and complete primary or secondary and more.[23] Father's occupation is recoded into two categories: agricultural workers and others. Ethnicity (coded in two categories) is captured either by self-reported ethnicity or by the ability to speak an indigenous language. Region of birth is coded into three broad regions (one being generally the capital area) but is captured by the type of area (urban/rural) for Panama. Table 2 describes the specific definitions of the circumstance variables in each survey in greater detail. Table 3 (Panel A) presents the corresponding descriptive statistics.

The number of categories for each circumstance variable was restricted to three or fewer, so as to reduce the number of "circumstances group" cells with zero or very few observations. As discussed in Section 3, this is important for the non-parametric analysis, which relies on the quality of the estimates for conditional means in these cells and their sampling variation may be very high for cells containing few observations. This greater sampling variance might artificially inflate the estimated inequality between groups,

[20] The *Encuesta Nacional sobre Niveles de Vida de los Hogares* (MxFLS) 2002 contains a similar set of variables. Unfortunately its simple size proved to be too small relative to the partition to generate reliable results.

[21] For Brazil, we further restrict the sample to household heads and their spouses, as the family background information was collected only for these individuals.

[22] Employment rates are defined as the ratio of all employed individuals to all persons in the 30-49 age-group. These rates for men and women (respectively) are: 0.90 and 0.55 in Brazil; 0.91 and 0.62 in Colombia; 0.97 and 0.72 in Ecuador; 0.96 and 0.51 in Guatemala; 0.91 and 0.53 in Panama; and 0.94 and 0.72 in Peru.

[23] Whether complete primary attainment was included as part of the middle or upper grouping for parental education depends on relative group sizes. An effort was made to prevent the top grouping from being too small relative to other countries, to enhance comparability. None of the results is particularly sensitive to these decisions.

thereby inducing an over-estimation of inequality of opportunity. Table 4 shows the maximum number of cells in each survey, the number of cells actually observed (i.e. the complement of the number of empty cells), the mean cell size and the proportion of cells with fewer than five observations. Despite observing only six circumstance variables and exercising considerable parsimony in the partitioning of the population, we still have two surveys – from Guatemala and Panamá – for which over 20% (40%) of cells have fewer than five observations in the income/consumption (earnings) analysis. By contrast, Brazil's national PNAD survey, with a sample size one order of magnitude larger, has 6%-8% of cells with fewer than five observations. This reflects the limited sample sizes of LSMS surveys, and underscores the importance of the parametric estimates in validating (or refuting) the non-parametric results presented below.

Turning to the advantage variables, labor earnings are measured on an individual basis as monthly earnings from all occupations, including the monetary value of various in-kind payments. Family incomes and consumption are measured as per capita household income (from all sources) and per capita aggregate household consumption. Aggregates for family incomes are computed as the sum of all household members' individual incomes, and include all jobs earnings plus other incomes such as those from assets, pensions and transfers.[24]

Consumption expenditure data is not available for Brazil. Elsewhere, the reference period is the year, but some expenditures are captured on a weekly or monthly basis. Consumption aggregates do differ across surveys in some respects. In particular, income and consumption are adjusted for differences in the local cost of living in most LSMS datasets (Ecuador, Guatemala, and Panama but not Colombia) and in the Peruvian ENAHO dataset. LSMS surveys (Colombia, Ecuador, Guatemala and Panama) and the ENAHO survey also include imputed rents for owner-occupied housing in both consumption and income aggregates, whereas the PNAD does not. Table 3 (Panel B) reports means and standard deviations for these three economic advantage variables.

5. Inequality in Earnings Opportunity

In most societies, remuneration for one's work in the labor market is an important component of overall income, and thus a key determinant of a person's command over private goods. Some argue that it is also directly related to self-esteem and social status. Furthermore, it is affected both by one's own choices and efforts, and by exogenous circumstances. Labor earnings would therefore certainly qualify as an "advantage" concept in Roemer's terms.

As noted by Lefranc et al. (2006), Roemer's concept of equality of opportunity for earnings would require that the distribution of earnings conditional on any circumstance variable be identical to the marginal distribution – i.e. that there should be no difference across the earnings distributions estimated for particular population subgroups defined according to their circumstances.[25] Figure 1 shows the distributions of

[24] The reference period for the earnings of self-employed workers, which is the month in Brazil, Colombia, and Peru; depends on the frequency of payments in Panama; and is the year elsewhere. For wage earnings the reference period is the month in all surveys.

[25] This is the strongest concept of equality of opportunity. Lefranc et al. (2006) allow for weaker concepts, such as there being no first- or second-order stochastic dominance between the conditional distributions.

earnings conditional on mother's education (Panel A) and on ethnicity (Panel B), for our six countries. At least in the case of mother's education, first order dominance should be satisfied in most countries as conditional distribution functions never cross.[26] The distances between circumstance groups appear particularly high in Brazil and Panama. There are also pronounced visual differences in the ethnicity panel, with the distance between ethnic groups being higher in Guatemala and Brazil.

The scalar inequality measures developed in Section 3 can help us quantify these inter-country differences, as well as allowing us to consider the combined (and partial) effects of all six circumstance variables. Table 5 presents our main results for earnings opportunity. The first row contains the estimates of overall earnings inequality in each of the six countries, using three indices from the generalized entropy class of measures: the mean log deviation, E(0); Theil index, E(1); and half the square of the coefficient of variation, E(2).[27] Below each point estimate, we report bootstrapped standard errors, computed using 100 replicates of the sample, taking into account weights, stratification and clustering for each re-sampling. Since they refer only to a selected sub-sample of the working population, as described in Table 1 (which largely corresponds to 30-49 year-olds), these are not national measures of earnings inequality. Nevertheless, the indices are generally high: mean log deviations are 0.572 in Panama, 0.608 in Colombia, 0.616 in Brazil, 0.638 in Ecuador, 0.675 in Peru and 0.786 in Guatemala.[28]

The next two rows of Table 5 present our non-parametric estimates of the opportunity share of inequality in earnings, given by θ_d^N and θ_r^N, on the basis of the partitions described in Section 4 and Table 2. We report the measures for all three inequality indices (E(0), E(1), and E(2)). As noted in Section 3, we focus our discussion on E(0), the only measure for which $\theta_d^N = \theta_r^N$. By this measure, differences in observed opportunity account for 20% of total earnings inequality in Colombia, 21% in Peru, 25% in Panama, 26% in Ecuador, 29% in Guatemala and 35% in Brazil. The differences between Brazil and any other country, with the exception of Guatemala, are statistically significant at 5% (as are the differences between Guatemala and Peru or Colombia).[29] In other words: a fairly coarse partition of the population by six circumstance variables accounts for between one-fifth and one-third of total earnings inequality.

It is interesting to note that this ranking is different from the ranking of overall earnings inequality. In particular, Brazil, which has only the fourth highest earnings inequality in the sample, has by far the largest opportunity share of that inequality. Similarly, Panama, which has the lowest earnings inequality, ranks above Colombia and Peru in opportunity share. Such re-ranking suggests that inequality of opportunity and inequality of outcomes are not simply different ways of measuring the same thing: they capture different features of distribution in a society.

---

[26] These figures are included as an illustration only, and are not the focus of our analysis. Statistical tests for stochastic dominance (Davidson and Duclos, 2000) are therefore not reported in this paper.

[27] As noted earlier, all generalized entropy measures are additively decomposable by population subgroup.

[28] The differences between Ecuador and Peru, as well as between Brazil, Colombia and Ecuador, or Brazil, Colombia and Panama, are insignificant at the 5% level, on the basis of the bootstrapped standard errors.

[29] The opportunity shares tend to be somewhat higher for the Theil index (E(1)); and much lower for E(2). In the case of the latter, this is due in part to its greater sensitivity to income gaps at the top of the distribution. Notice also that the path-dependence is most pronounces for E(2), given the squaring of relative means built into its weighting scheme for the within-group component. See footnote 13.

The remainder of Table 5 presents the results of the parametric decompositions, namely θ_r^P and θ_r^J, for J set to equal each individual circumstance variable in our partition: gender, ethnicity, father's occupation, father's education, mother's education and birthplace.[30] Still focusing on the path-independent measure E(0), the parametric approach yields systematically lower (overall) opportunity shares of inequality, ranging from 17% in Colombia to 34% in Brazil. This is true in all countries, although the difference between θ_r^N and θ_r^P is only approximately 3% (and statistically insignificant) in the case of Brazil. The differences are larger, but either borderline significant or insignificant at the 5% level, in Colombia, Ecuador, Guatemala, and Peru. The country ranking is identical for the two estimation procedures.

These systematic differences are consistent with the expectation (discussed in Section 3) that the large sampling variance within cells with very few observations may cause an upward bias in the non-parametric estimates. Although it can not be ruled out that the (linear) functional form assumption implicit in the parametric estimate might lead to underestimates, the fact that in the only country for which we have a substantially larger sample size (Brazil) the difference almost vanishes provides some support for the suspicion that the bias might come from the sampling variance in small cells in the non-parametric estimates. Nevertheless, given the remaining uncertainty, we make two recommendations: (i) wherever possible, surveys that may be used for measuring inequality of opportunity should collect larger sample sizes; and (ii) where that is not possible, both parametric and non-parametric estimates should be reported to provide a plausible range for the true lower-bound value of inequality of opportunity.

Regarding the effect of each individual circumstance, Θ_r^J is highest for family background variables in all countries. This is particularly true for mother's education which is associated with between 9% and 12% of total inequality. The relative shares of inequalities associated with ethnicity and place of birth vary across countries, with ethnicity being more important in Brazil, Guatemala, and Panama – where it accounts for between 3% and 7% of inequalities – and the geography of birth having more effect in Peru, Brazil and Panama, where it accounts for 4-6% of overall inequality. Finally, inequality of opportunity related to gender ranges from a low of 0-1% in Panama and Colombia, to a high of 5% in Guatemala. In Brazil, Mexico and Ecuador, gender accounts for 3-4% of overall inequality.

6. Inequality in Opportunity for Household Welfare

Earnings are a key component of family incomes, and an important source of individual status, self-esteem and bargaining power but, as a measure of individual well-being, it would be seriously incomplete. Total household income (or consumption expenditure) per capita are better measures of welfare, because they account for incomes

[30] As discussed in Section 3, these indices are based on the OLS estimates of equation (7). The reduced-form estimates for all six countries are presented in Table 6. All coefficients have the expected signs and most are quite strong. Since this is a reduced-form equation, these coefficients can not be interpreted causally, and capture both the partial direct effects of *C* on *y*, and indirect effects through *E*. As noted earlier, labor market participation is almost 100% for men in this age group, but much lower for women. As implied by the reduced-form specification, we are estimating inequality of earnings opportunities *conditional* on being active in the labor market, so it would be inappropriate to correct for selection.

from other sources (such as capital incomes or transfers) and for resource pooling within the household. Unless access to public and publicly provided goods (such as public safety and free public education or health care, respectively) is taken into account, household income or consumption expenditures are also incomplete and partial measures of welfare. Still, they are better measures than earnings and, for many countries, they are the best available indicators of individual welfare available.[31]

Figure 2 depicts the conditional distributions of consumption per capita for circumstance groups defined according to mother's education (in Panel A) and ethnicity (in Panel B), analogously to Figure 1 for earnings. In Panel A, the consumption distances between groups defined by mother's education are larger than the corresponding earnings gaps (shown in Figure 1) for all five countries, and largest for Guatemala and Panama. Panel B exhibits greater variation across countries, with large gaps between ethnic groups in Guatemala and Panama, much more limited (or insignificant) distances in Colombia, and an intermediate pattern in Ecuador and Peru.[32]

Tables 7 and 8 present our relative measures of inequality of opportunity for household income and consumption expenditures per capita, respectively. These tables are analogous to Table 5 (for earnings), and report θ_d^N, θ_r^N, θ_r^P and Θ_r^J for E(0, 1 and 2), along with bootstrapped standard errors taking into account sampling weights, stratification and clustering. Gender is excluded from the set of circumstance variables since these indicators are defined at the level of the household, and the gender of the household head is endogenous (and thus not a circumstance).[33] We therefore work with five circumstances (race, father's and mother's education, father's occupation, and birth place). Income results are reported for all six countries, but consumption data was not available in the PNAD data (for Brazil). Table 8 contains estimates of inequality of opportunity for consumption for Colombia, Ecuador, Guatemala, Panama, and Peru. Tables 9 and 10 report the OLS coefficients of the reduced-form equation (7), for income and consumption expenditures respectively (analogously to Table 6).[34]

In our samples, overall household income inequality is higher than earnings inequality in Brazil and Panama (by all measures) and in Colombia (by E(1) and E(2), but not by E(0)). It is lower than earnings inequality in Ecuador, Guatemala, and Peru (by all measures), and in Colombia by E(0). In all five countries for which consumption data is available, consumption inequality is considerably lower than either income or earnings inequality. This is consistent with the widespread view that income (and earnings, when these include agricultural and informal sector earnings) is measured with greater error than consumption expenditures, as well as with the expectation that consumption is likely

[31] There are two other steps in the mapping from household income or consumption to individual welfare which we overlook here, by using income or consumption per capita. First, we make an extreme assumption about the (in)existence of economies of scale in consumption within the household. Second, we assume that household resources are shared equally, which they may well not be.

[32] Recall that there is no data for consumption in Brazil's PNAD survey.

[33] Endogeneity arises both because in some countries reported headship is an interviewee choice, and because household formation (e.g. whether or not one marries) is endogenous.

[34] All coefficients in these reduced-form regressions have the expected signs, and most are significant at the 1% level. Coefficient sizes are consistent with a reduced-form specification.

to be closer than current income to permanent income (provided households have access to some consumption-smoothing mechanisms).[35]

Focusing once again on the path-independent measure E(0), non-parametric estimates of inequality of opportunity for household incomes range from 25% (in Colombia) to 37% (in Guatemala). As for earnings, the parametric estimates are slightly lower: from 23% in Colombia to 35% in Guatemala. For both estimation procedures, the indices are higher than the corresponding estimates for earnings in all countries except for Brazil, where the difference is quite small: 34% for earnings versus 32% for income per capita (for the parametric estimates). In addition to earnings capacity, pre-determined circumstances affect another three important household income determinants: other incomes (such as capital incomes or transfers); the choice of one's partner; and the composition of the rest of the household (including, most importantly, the number of children). The pattern found in the data suggests that inequality of opportunities in these three domains tends to reinforce the inequality of opportunities that operates through the earnings channel in Colombia, Ecuador, Guatemala, Panama and Peru; but to partially offset them in Brazil.

While total inequality is lower in the distribution of consumption expenditures than in the income distribution, the opposite is true for estimates of inequality of opportunity. The (E(0)) opportunity shares of inequality reported in Table 8 are considerably higher than those reported in Table 7 for all five countries, and regardless of whether the estimates are parametric or non-parametric. The differences are in the 20-30% range for Ecuador, Panama and Peru, 40% for Guatemala, but only 6% for Colombia. This supports the notion that income-based measures of inequality of opportunity tend to underestimate lifetime (or permanent income) inequality of opportunity, since transitory income variance (and likely higher measurement error) is effectively counted as inequality due to "efforts and luck".[36] Our non-parametric (parametric) estimates of inequality of opportunity in the distribution of consumption expenditures are: 27% (24%) in Colombia, 34% (32%) in Ecuador, 35% (34%) in Peru, 42% (39%) in Panama, and 52% (50%) in Guatemala.[37]

Despite the sample size limitations (especially for Panama and Guatemala), the parametric and non-parametric estimates turn out to be very close. The differences are smaller for consumption inequality than for earnings, reflecting larger sample sizes, and thus a lower proportion of cells with zero or few observations (see Table 4). Although the parametric estimates remain systematically below their non-parametric counterparts, the differences are now never statistically significant, and the country ranking is identical.

Turning to the analysis of individual circumstance variables, we find that family background characteristics are once again associated with the largest share of inequality of opportunity. The share of inequality accounted for by mother's education alone is higher than 15% in most countries, and as high as 26% in Guatemala. The share of inequality associated with the other variables is usually higher than for earnings, with the

[35] See, e.g. Deaton, 1997, on both of these reasons to prefer consumption to income data in assessing the distribution of welfare in developing countries.
[36] See Bourguignon et al. (2007) for a discussion. The finding is analogous to the well-known fact that inter-generational mobility estimates are much higher when based on single-period wages for parents and children, than when based on longer earnings histories. See, inter alia, Solon (1999) and Mazumder (2005).
[37] With the exception of the difference between Ecuador and Peru, all cross country differences are significant at the 5% level, on the basis of the bootstrapped standard errors.

same broad ranking across different circumstances (parental background more important than either race or birth region). The higher levels of inequality of opportunity observed in Central American countries, however, are associated with larger partial shares for region of birth (which is also important in Peru) and ethnicity.

7. The opportunity-deprivation profile: identifying the least advantaged groups

The analysis has so far focused on scalar measures of inequality of opportunity in each country, largely expressed as shares of total outcome inequality. These indices can be useful to summarize the importance of a set of predetermined circumstances in the structure of inequality in a particular country. Since the relative measures are not closely correlated with measures of outcome inequality, they are also informative of some of the differences in the *nature* of inequality across countries. Ultimately, a country where a smaller share of total inequality is associated with differences in opportunity is likely to be a fairer society, where individual choices and effort (and luck) play a greater role in determining outcomes than family origin, race or gender.

However, the partition of the population $\{y_i^k\}$ into K *types*, which is undertaken to generate these indices, can also be used to yield a byproduct of potentially even greater interest to analysts and policymakers alike. Recall that each cell in the partition corresponds to a Roemerian type $T_k : i | i \in k$, such that $C_i^k = C^k \Leftrightarrow i \in k, k = 1,...,K$. We have seen that equal opportunities attain when $F_k(y) = F_l(y), \forall k \neq l$, which is a different way of writing $F(y|C) = F(y)$ for a discrete partition. Differences in the outcome distributions among types therefore are taken to reveal (or arise from) inequality of opportunity.

At least conceptually, it is not unreasonable to see $F_k(y)$ as an individual *i*'s ($i \in k$) opportunity set for outcome *y*. Given *i*'s circumstances C_i^k, only *i*'s own choices, efforts and luck will determine his final position, $p_i = F_k(y_i)$. If it were possible, therefore, to rank $F_k(y)$ across *k* in a meaningful way, we would obtain a ranking of opportunity sets across types, which we call an *opportunity profile*.

As previously discussed, one obvious such ranking would be given by any (first- or second-order) stochastic dominance relationships between types. However, the stochastic dominance approach to constructing an opportunity profile suffers from two problems. The first, which is conceptual in nature, is that any such ranking is perforce partial and incomplete (see Atkinson, 1970). The second, which is practical in nature, is that the distribution of cell sizes partly summarized in Table 4 makes it impossible to estimate the conditional distributions for the full set of 54 – 216 types in our partitions.

Albeit conceptually less satisfactory, a feasible alternative ranking algorithm would be to use a particular moment of $F_k(y)$, such as the mean, or indeed a particular percentile, such as the median, the first quartile, etc. Because the type's mean outcome, μ^k, was central in constructing smoothed and standardized distributions, and thus for the decomposition exercises reported in Sections 5 and 6, we choose to use it as the ranking criterion for the type-specific opportunity sets $F_k(y)$ in what follows. While this strikes us as a reasonable choice, it is still arbitrary, and the reader is cautioned that alternative criteria are certainly possible, and might imply different rankings.

Furthermore, we choose to focus only on the least-advantaged types in a society – those with the lowest-ranking opportunity sets. We avoid the issue of setting an "opportunity deprivation threshold", and choose simply to consider the types that account for the bottom 10% of the population. In other words, we rank all types in each country in increasing order of mean outcome. The bottom *m* groups are included in that country's opportunity profile, where the population share over *m* sums to 10%. Formally, these are the *m* groups k = 1,…, m such that: $\mu^1 \leq \mu^2 \leq ... \leq \mu^m \leq \mu^j$, for every $j > m$, and $\sum_{k=1}^{m} N_k = \frac{N}{10}$, where N is the overall population size. We refer to the set of types $\{k | k \in (1,...,m)\}$ as the *opportunity-deprivation profile*; and to the individuals *i* that belong to those types as the *opportunity-deprived*.

Opportunity-deprivation profiles for each of our six countries, specified in terms of the full set of circumstances that define them (ethnicity, mother's and father's education levels, father's occupation and birthplace) and constructed from the non-parametric estimates of conditional group means, are available from the authors on request. The number of types in the opportunity-deprivation profile varies across countries: There are 5 such groups in Guatemala and Peru, 6 in Brazil, 10 in Colombia, 16 in Ecuador, 20 in Mexico, and 25 in Panama. Some types represent large populations (there are two groups in the Brazilian profile that represent more than 2 million people each) while others represent only a few hundred individuals.

When presented in their "full" form, opportunity-deprivation profiles are simply a list of the types with the lowest-ranking opportunity sets in each country in our sample. For comparative purposes, however, it may be useful to have a synthetic overview of the opportunity-deprived group as a whole, in each country. Table 11 thus summarizes a number of characteristics of all opportunity-deprived individuals in our six countries. Three common traits are salient. First, members of ethnic minorities form the vast majority of the population in these disadvantaged groups. In three of our six countries, these groups are composed exclusively of members of racial or ethnic minorities: black and mixed-race in Brazil; and native speakers of indigenous languages in Guatemala and Peru. In two other countries, ethnic minorities are still a majority of the opportunity-deprived: 76% of the opportunity profile in Panama consists of native speakers of indigenous languages; and 61% of self-reported indigenous, black or mixed-race ethnicity in Ecuador. Colombia is the only country in our sample where ethnic minorities are not the majority among the opportunity-deprived but, even there, the proportion of minorities, 33%, is still higher than in the population as a whole.

Second, family background is also strongly associated with opportunity-deprivation. In the four countries where this information is available, never fewer than 83% of the opportunity-deprived are daughters and sons of agricultural workers, and this proportion reaches 100% in Guatemala. Almost the same holds for parental education: In all countries, more than 90% of the opportunity-deprived are daughters and sons of women who did not go to school – 99% in Guatemala and Peru, 98% in Ecuador, 96% in Colombia, 93% in Panama, and 91% in Brazil. Similar results hold for father's education, although in Colombia, Ecuador and Panama, father's education is a less powerful predictor of opportunity deprivation than mother's education.

Third, opportunity deprivation is remarkably spatially concentrated. A majority of the opportunity-deprived are often natives of the same specific regions. In Brazil, all persons in our profile were born in the Northeast or North regions; in Colombia, 99% hail from peripheral departments; in Guatemala, 99% come from one of the North and North-western departments; in Panama, 96% were born in a rural area.[38] There is somewhat greater spatial heterogeneity in the opportunity-deprivation profiles for Ecuador and Peru.

Does opportunity deprivation manifest itself in lower economic achievement levels? Qualitative, the answer is "yes" by construction, since the types were ranked by mean economic achievement. Quantitatively, the last row in Table 11 gives the income share of the opportunity deprived in each country. Since they account for 10% of the population in all countries by construction, the distance between their income (or consumption expenditure) share and 10% can be seen as a rough quantitative measure of the economic consequences of opportunity deprivation in each country. The income share of the 10% of the population we classify as opportunity deprived is 2,7% in Panama, 2.9% in Brazil, 3.5% in Guatemala, and 4.4% in Ecuador, 4.8% Peru, and 5.0% in Colombia.[39]

We conclude this section with a brief discussion of the differences between the opportunity-deprivation profile we have introduced, and the more standard concept of a poverty profile. A poverty profile describes the characteristics of individuals with individual incomes below a poverty line, whereas the opportunity profile ranks individuals by the mean income (or consumption) of the *type* they belong to. These are conceptually very different objects. An opportunity-deprivation profile will *include* individuals from very disadvantaged backgrounds, who happened to be successful and have escaped poverty through their own efforts or sheer luck. A poverty profile will not. An opportunity-deprivation profile will *exclude* individuals from more advantaged backgrounds, who did poorly either through bad luck or poor performance, whereas a poverty profile will include them.

Differences between the two profiles may, therefore, contain information on how powerful circumstances are in determining poverty outcomes. If there is very little difference, effort and luck would appear to be largely powerless to compensate for the initial opportunity deprivation individuals inherit. Conversely, if there is limited overlap between the two profiles, one could claim that initial circumstances matter little to a person's chances of escaping poverty. Table 12 describes the poverty profile for our six countries, by arbitrarily fixing the poverty line at the first decile in each distribution. In this fashion both profiles refer to the "bottom" 10% of the population, with the difference arising from the ranking criterion used to define "bottom".

The comparison of the two profiles reveals interesting patterns. Unsurprisingly, the opportunity-deprived are more homogenous than the poor with respect to most

[38] Geographical regions are not reported in the survey for Panama, so an urban-rural subdivision was used instead.

[39] One can also isolate the types that account for the top end of the opportunity profile in each country. Call them "opportunity hoarders". Their income shares are 22.6% in Panama, 23.1% in Ecuador, 23.7% in Peru, 25.8% in Colombia, 28.8% in Brazil, and 29.3% in Guatemala. Details of the "opportunity-hoarding" profile for our six countries is available from the authors on request.

circumstance variables. Although ethnic minorities form the majority of the opportunity-deprived in five countries (and 100% in Brazil, Guatemala and Peru) they account for lower shares of the poor: 70% in Guatemala, 69% in Brazil, 56% in Peru, 54% in Panama, 34% in Mexico, 32% in Ecuador, and 15% in Colombia. A similar pattern arises for place of birth: poverty is less spatially concentrated than economic opportunity: 70% of the poor in Brazil live in the North or Northeast (as compared with 100% of the opportunity-deprived being born there). 65% of the poor live in Colombia's peripheral departments, while 99% of the opportunity-deprived were born there. And so on. Family backgrounds are also more heterogeneous among the poor than among the opportunity deprived, although the share of children of agricultural workers is still very high at 80% in Ecuador and Panama, and 75% in Guatemala.

The last row in Table 12, analogously to Table 11, provides the income share of the poor in each country. They are 0.7% in Brazil (using income per capita), 1.5% in Colombia and Panama, 1.8% in Guatemala and Peru, and 1.9% in Ecuador. The ratio of the income share of the opportunity-deprived to the income share of the poor is 1.80 in Panama; 1.94 in Guatemala; 2.31 in Ecuador; 2.66 in Peru, 3.33; in Colombia; and 4.14 in Brazil.[40] Since the income share of the opportunity-deprived is larger when some among them succeed in escaping poverty, these ratios are suggestive indicators of "mobility". The higher the ratio, the less opportunity-deprivation would seem to amount to a sentence of life in poverty, delivered at birth. Nevertheless, more confident statements on the relationship between opportunity-deprivation profiles, poverty profiles, and more standard measures of mobility (which largely rely on the association between outcomes and one particular circumstance, such as father's wage or education), would require further work.[41]

8. Conclusions

This paper has proposed a simple conceptual framework for the measurement of inequality of opportunity, which derives two empirical tools directly from John Roemer's theory of equal opportunities. The first tool is a class of scalar indices that measure inequality of opportunity as the share (or level) of overall inequality in a given population which exists between social groups defined by different initial circumstances (rather than within these groups). The indices are inspired by the observation that, if opportunities were equally distributed, outcomes would be orthogonal to pre-determined morally irrelevant circumstances, so that the between-type inequality share would be zero. Because not all relevant circumstances are observed, the indices provide a lower-bound estimate of inequality of opportunity.

Indices belonging to this class may differ along three dimensions: the inequality aversion parameter in the underlying inequality measure; the path of the decomposition, and the nature of the estimation procedure. We show that if we restrict our attention to

[40] The number for Brazil is *not* comparable to those of the other countries, since it is built on an income, rather than consumption expenditure, distribution.

[41] Van de Gaer et al. (2001) contain a pioneering theoretical discussion of the relationship between mobility and equality of opportunity. See also Gaviria (2007) for a recent survey of intergenerational mobility in Latin America, with some discussion of attitudes to redistribution. Fields et al. (2007) provide a survey of the evidence on intra-generational income mobility in the region.

path-independent decomposable inequality indices, the class collapses to a unique index, which can be estimated either parametrically or non-parametrically. The proposed parametric estimation procedure is a useful complement to the simple non-parametric decomposition both for data-efficiency reasons, and to estimate partial, circumstance-specific indices.

The second empirical tool is an opportunity-deprivation profile: the list of Roemerian *types* (i.e. social groups that share identical circumstances) that account for the lowest-ranked p% of the population, when types are ranked by their mean advantage levels.[42] The profile identifies the types with the lowest-ranked opportunity sets in society. If followed over time, they would allow a practical application of Roemer's (2006) suggestion that economic development might be measured by the rate of progress of the worst-off type.

We applied these concepts to a rich data set for six countries in Latin America, whose surveys contained information on a number of relevant pre-determined, morally irrelevant circumstances, namely: gender, race or ethnicity, birthplace, mother's and father's education, and father's occupation. We calculated our unique path-independent measure of inequality of opportunity both parametrically and non-parametrically, for the distributions or earnings, household per capita income, and household per capita consumption expenditure. As expected, the non-parametric method tended to systematically overstate inequality of opportunity when sample sizes were small. For larger samples, and in particular when using household income or consumption per capita as indicators of advantage, the two estimates were numerically close and statistically insignificant, generating a robust lower-bound estimate of inequality of opportunity.

For labor earnings, the (parametric) estimates ranged from 17% of total inequality in Colombia, to 34% in Brazil. For household income per capita, the range was between 23% in Colombia to 35% in Guatemala. For consumption expenditures, the range was between 24% in Colombia, and 50% in Guatemala. Differences between the indices for the distribution of earnings and those for household welfare are due both to differences in the extent of measurement error, and to differences in the mechanisms through which circumstances affect outcomes (e.g. family formation, labor force participation, etc.). In all cases, family background – proxied for by parental education levels and the father's occupation – was the largest component of the opportunity share of inequality, although ethnicity was also important in Brazil, Guatemala and Panama.

The opportunity profiles provide an X-ray of the opportunity structures in Latin America, at least for those social groups with the most limited opportunity sets in these six countries. These "opportunity-deprived" types were overwhelmingly members of ethnic minorities, and tended to hail from agricultural families with low levels of education, living in poor regions. A comparison of their income shares with those of the poorest 10% of the population in each country reveals that, as expected, many of those with limited initial opportunities do manage to move out of poverty, while others – from more advantaged backgrounds – fall into poverty. Yet, in no country did the 10% most opportunity-deprived people account for more than 5% of total consumption expenditure. In Brazil and Panama, the figure was less than 3%.

---

[42] We set p%=10%, and used consumption expenditure per capita as our indicator of economic advantage (except for Brazil, where income per capita was used instead).

Both the scalar indices - which reveal that the lower-bound for the share of consumption inequality which is due entirely to factors beyond the individuals control is of the order of 30% to 50% - and the opportunity-deprivation profiles - which document the enduring "costs" of being born of certain races, in certain places and to certain families – suggest that unequal opportunities are an important source of the outcome differences we observe in Latin America. This is a part of inequality that can not be explained as a return to effort, or even as the result of random shocks and pure luck.

In this paper, we have sought to lay the foundation for the measurement of inequality of opportunity, relating it to the relevant economic theory. It would be interesting for future work to investigate at least two aspects of these findings: first, how do our comparisons of opportunity-deprivation and poverty profiles relate to more standard measures of intergenerational mobility? Second, do differences among countries in the *nature* (e.g. in the opportunity share) – rather than merely in the *level* – of inequality, affect social and political attitudes, the nature of redistribution systems, and the rate of economic growth?

References

Alesina, Alberto and George-Marios Angeletos (2005): “Fairness and Redistribution”. *American Economic Review*, **95** (4): 960-980.

Arneson, Richard (1989): “Equality of Opportunity for Welfare”, *Philosophical Studies*, **56**: 77-93.

Atkinson, Anthony B. (1970): “On the Measurement of Inequality”, *Journal of Economic Theory*, **2**: 244-263.

Banerjee, Abhijit, and Esther Duflo (2003): "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth*, **8**(3): 267–99.

Barros, Ricardo P., Jose Molinas Vega and Jaime Saavedra (2008): “Measuring Inequality of Opportunities for Children”, Washington, DC: World Bank mimeo.

Bénabou, Roland and Jean Tirole (2006): “Belief in a Just World and Redistributive Politics”, *Quarterly Journal of Economics* **121**(2), 699-746.

Bourguignon, François (1979): “Decomposable Income Inequality Measures”, *Econometrica*, **47**(4): 901-920.

Bourguignon, François, Francisco H.G. Ferreira and Michael Walton (2007): Equity, Efficiency and Inequality Traps: A research agenda”, *Journal of Economic Inequality*, **5**: 235-256.

Bourguignon, François, Francisco H.G. Ferreira and Marta Menéndez (2003): “Inequality of Outcomes and Inequality of Opportunities in Brazil”, Policy Research Working Paper Series # 3174, The World Bank, Washington, DC.

Bourguignon, François, Francisco H.G. Ferreira and Marta Menéndez (2007): “Inequality of Opportunity in Brazil”, *Review of Income Wealth*, **53** (4): 585-618.

Checchi, Daniele and Vitoroco Peragine (2005): “Regional Disparities and Inequality of Opportunity: the Case of Italy”, IZA Discussion Paper No. 1874/2005.

Cogneau, D., T., Bossuroy, Ph., De Vreyer, C., Guenard, V., Hiller, Ph., Leite, S., Mesple-Somps and L., Pasquier-Doumer, C., Torelli (2006): “Inequalities and Equity in Africa”, Working Paper DIAL, 2006-11.

Cogneau, Denis and Jeremie, Gignoux, (2007). “Earnings Inequalities and Educational Mobility in Brazil over two Decades” (with Denis Cogneau), forthcoming in Klasen S. and Nowak-Lehmann (eds.), Poverty, Inequality and Policy in Latin America, *CESifo Seminar Series*, Massachusetts Institute of Technology (MIT) Press.

Cohen, Gerry A., (1989). “On the Currency of Egalitarian Justice”, *Ethics*, **99**: 906-944.

Davidson, Russell and Jean-Yves Duclos, (2000): "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality", *Econometrica*, **68**(6): 1435-1464.

Deaton, Angus, (1997). *The Analysis of Household Suveys: a Microeconometric Approach to Development Policy*. Baltimore: John Hopkins University Press.

Elbers, Chris, Peter Lanjouw, Johan A. Mistiaen and Berk Ozler (2005): “Re-interpreting sub-group inequality decompositions.” Policy Research Working Paper Series 3687, World Bank.

Fields, Gary, Robert Duval, Samuel Freije and M. L. Sanchez-Puerta (2007): “Intragenerational Income Mobility in Latin America”, *Economía* **7** (2): 101-154.

Fleurbaey, Marc (2008): *Fairness, Responsibility, and Welfare*. Oxford: Oxford University Press.

Forbes, Kristin J. (2000). "A Reassessment of the Relationship Between Inequality and Growth." *American Economic Review* **90**(4): 869–87.

Foster, James and Artyom Shneyerov (2000): “Path Independent Inequality Measures”, *Journal of Economic Theory*, **91**: 199-222.

Gaviria, Alejandro (2007): “Social Mobility and Preferences for Redistribution in Latin America”, *Economía* **8** (1): 55-96.

Hild, Matthias and Alex Voorhoeve (2004): “Equality of Opportunity and Opportunity Dominance”, *Economics and Philosophy*, **20:** 117-145.

Lefranc, Arnaud, Nicolas Pistolesi and Alain Trannoy (2006): “Inequality of Opportunities vs. Inequality of Outcomes: Are Western Societies All Alike?”, ECINEQ Discussion Paper # 2006-54 (August).

Mazumder, Bhakshar (2005): "The Apple Falls Even closer to the Tree than We Thought: New and Revised Estimates of the Intergenerational Inheritance of Earnings." In S. Bowles, H. Gintis, and M. Groves (eds.) *Unequal Chances: Family Background and Economic Success.* Princeton, N.J.: Princeton University Press.

Peragine, Vitorocco (2004): “Ranking Income Distributions according to Equality of Opportunity”, *Journal of Economic Inequality*, **2**, 11-30.

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

Roemer, John E. (2006): “Review Essay, ‘The 2006 world development report: Equity and Development”, *Journal of Economic Inequality* **4**: 233-244.

Sen, Amartya (1985): *Commodities and Capabilities*. Amsterdam: North Holland.

Solon, Gary (1999): "Interneneerational Mobility in the Labor Market." In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics.* Amsterdam: North-Holland.

Shorrocks, Anthony F., (1980): “The Class of Additively Decomposable Inequality Measures”, *Econometrica*, **48** (3), 613-625.

Van de Gaer, Dirk, Erik Schokkaert and Michel Martinez (2001): “Three Meanings of Intergenerational Mobility”, *Economica*, **68**: 519-537.

World Bank (2006): *World Development Report 2006: Equity and Development*. Washington, DC: The World Bank and Oxford University Press.

Table 1: Survey names, dates and sample sizes

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Survey	PNAD 1996	ECV 2003	ECV 2006	ENCOVI 2000	ENV 2003	ENAHO 2001
Sample selection criteria	30-49 head or spouse	30-49	30-49	30-49	30-49	30-49 head or spouse
Original sample size	85,692	22,517	12,650	6,956	6,339	13,947
Observations with earnings and circumstances	50,560	16,575	9,671	4,661	4,127	9,830
(share of original sample)	0.590	0.736	0.765	0.670	0.644	0.704
Observations with income/consumption and circumstances	71,688	22,436	12,643	6,865	5,653	13,649
(share of original sample)	0.837	0.996	0.999	0.984	0.889	0.979

Table 2: Definition of circumstance variables

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Ethnicity						
category 1	self reported white ethnicity	Other	self-reported ethnicity: white, mixed blood ("mestizo") or other	European maternal language		European maternal language
category 2	self reported black ("negro") and mixed blood ("pardo") ethnicity	self-reported minority ethnicity: "indígena, gitano, archipiélago o palenquero"	self-reported ethnicity: indigenous, black ("negro" or "mulato").	indigenous maternal language	speaks indigenous language	indigenous maternal language
Father's occupation						
category 1	agricultural worker	Missing	agricultural worker or domestic worker	agricultural worker	agricultural worker	missing
category 2	Other		Other	other	other	
Mother's and father's education						
category 1	None or unknown	none or unknown	none or unknown	none or unknown	none or unknown	none or unknown
category 2	completed grade 1 to 4	primary incomplete	Primary	primary incomplete	primary	primary incomplete
category 3	completed grade 5 or more	primary complete or more	secondary or more	primary complete or more	secondary or more	primary complete or more
Birth region						
category 1	Sao Paulo & Federal district	departments at the periphery	Sierra & Amazonia provinces	Guatemala city, North-East departments and El Petén	cities and intermediate urban centers	Inland non-southern departments
category 2	South East, Center-West & South	Central departments(a)	Costa & Insular provinces	North & North-West departments	other urban centers	Southern and other costal departments
category 3	North-East, North or missing	Bogota, San Andres and Providencia islands and foreign country	Pichincha province (with Quito) & Azuay province	South-East, South-West & Center departments	rural areas	Arequipa, Callao & Lima

(a) Central departments are Boyaca, Caldas, Caqueta, Cundinamarca, Huila, Meta, Norte de Santander, Quindio, Risaralda, Santander, Tolima, and Valle del Cauca.

Table 3: Descriptive statistics

a. Circumstances

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Gender						
male	47.4	46.4	48.8	47.2	48.5	47.6
female	52.6	53.6	51.2	52.8	51.5	52.4
Ethnicity						
majority	59.8	90.8	88.3	69.3	92.2	72.3
minority	40.2	9.2	11.7	30.7	7.8	27.7
Father's occupation						
agricultural worker	35.0	missing	51.9	49.5	37.1	missing
other	65.0		48.1	50.5	62.9	
Father's education						
none or unknown	50.2	36.2	27.9	67.3	21.7	30.9
primary	40.2	49.0	56.1	17.3	54.2	32.1
primary complete / secondary	9.7	14.8	16.1	15.5	24.1	37.0
Mother's education						
none or unknown	53.1	31.7	29.3	76.8	24.5	48.7
primary	37.9	53.9	56.4	12.2	54.5	24.9
primary complete / secondary	9.0	14.4	14.4	11.0	21.1	26.4
Birth region						
Region 1	17.6	45.1	31.3	27.5	30.4	45.4
Region 2	47.1	45.8	50.6	20.9	21.1	35.7
Region 3	35.3	9.1	18.1	51.6	48.5	18.8

b. Economic outcomes

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Currency unit	Reais 1996	Pesos 2003	USD 2006	Quetzal 2000	Balboas (USD) 2003	Sols 2001
Individual earnings	905.8 [1,460.1]	544,800 [838,400]	341.4 [516.2]	1,734.2 [4,195.4]	477.0 [620.0]	809.9 [1,542.9]
Per capita total household income	391.9 [708.2]	329,300 [546,400]	199.0 [256.1]	667.0 [1,237.6]	255.3 [376.2]	376.8 [682.9]
Per capita consumption		341000 [485,300]	123.0 [131.8]	603.5 [701.8]	180.3 [187.5]	307.4 [353.2]

Means and standard deviations for economic outcomes in the population. Sources: all six surveys, samples for analysis of per capita income.

Table 4: Description of the disaggregation of the population into circumstances cells

a. Samples for earnings analysis

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Maximum number of groups	216	108	216	216	216	108
Actual number of groups	214	105	193	172	147	102
Mean number of observations per group	236.3	150.2	50.1	27.1	28.1	96.4
Proportion of groups with fewer than 5 observations	0.08	0.14	0.33	0.41	0.44	0.11

b. Samples for income and consumption analysis

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Maximum number of groups	108	54	108	108	108	54
Number of groups observed	108	54	102	96	84	53
Mean number of observations per group	663.8	394.8	124	71.5	67.7	257.5
Proportion of groups with fewer than 5 observations	0.06	0.06	0.17	0.23	0.30	0.08

Table 5: Inequality of Opportunity Indices for Labor Earnings

	BRAZIL			COLOMBIA			ECUADOR			GUATEMALA			PANAMA			PERU		
	E(0)	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)
TOTAL INEQUALITY	**0.616**	0.637	1.271	**0.608**	0.583	1.184	**0.638**	0.587	1.262	**0.786**	0.790	2.927	**0.572**	0.485	0.843	**0.675**	0.679	1.814
	0.009	*0.014*	*0.086*	***0.023***	*0.037*	*0.189*	***0.020***	*0.027*	*0.151*	***0.047***	*0.071*	*0.990*	***0.027***	*0.037*	*0.167*	***0.023***	*0.036*	*0.367*
NON PARAMETRIC ESTIMATES																		
θ_d^N	**0.349**	0.341	0.209	**0.203**	0.235	0.144	**0.256**	0.284	0.158	**0.293**	0.314	0.116	**0.245**	0.266	0.161	**0.212**	0.220	0.095
	0.008	*0.010*	*0.015*	***0.021***	*0.024*	*0.015*	***0.016***	*0.021*	*0.027*	***0.029***	*0.038*	*0.081*	***0.024***	*0.026*	*0.026*	***0.018***	*0.019*	*0.024*
θ_r^N	**0.349**	0.322	0.344	**0.203**	0.258	0.411	**0.256**	0.314	0.498	**0.293**	0.353	0.516	**0.245**	0.272	0.396	**0.212**	0.227	0.000
	0.008	*0.013*	*0.040*	***0.021***	*0.038*	*0.080*	***0.016***	*0.027*	*0.063*	***0.029***	*0.048*	*0.184*	***0.024***	*0.042*	*0.088*	***0.018***	*0.038*	*0.326*
PARAMETRIC ESTIMATES																		
θ_r^P	**0.338**	0.295	0.247	**0.170**	0.197	0.319	**0.214**	0.237	0.381	**0.231**	0.227	0.045	**0.174**	0.148	0.159	**0.174**	0.105	0
	0.009	*0.020*	*0.132*	***0.022***	*0.046*	*0.115*	***0.017***	*0.028*	*0.061*	***0.028***	*0.066*	*0.674*	***0.026***	*0.044*	*0.085*	***0.021***	*0.075*	*1.098*
θ_r^J																		
Gender	**0.036**	0.018	0	**0.003**	0	0	**0.026**	0	0	**0.054**	0.033	0.031	**0**	0	0.012	**0.019**	0	0
	0.005	*0.011*	*0.056*	***0.008***	*0.019*	*0.067*	***0.016***	*0.038*	*0.174*	***0.023***	*0.046*	*0.278*	***0.012***	*0.029*	*0.085*	***0.007***	*0.017*	*0.101*
Race	**0.074**	0.067	0.070	**0.001**	0.001	0.002	**0.008**	0.010	0.017	**0.032**	0.038	0.073	**0.031**	0.020	0.031	**0.023**	0.021	0
	0.004	*0.006*	*0.025*	***0.002***	*0.003*	*0.005*	***0.003***	*0.004*	*0.007*	***0.008***	*0.009*	*0.015*	***0.007***	*0.005*	*0.006*	***0.007***	*0.010*	*0.073*
Father's occupation	**0.068**	0.058	0.062		(a)		**0.061**	0.061	0	**0.016**	0.019	0.043	**0.057**	0.042	0		(a)	
	0.003	*0.005*	*0.019*				***0.009***	*0.016*	*0.075*	***0.006***	*0.009*	*0.026*	***0.013***	*0.031*	*0.127*			
Father's education	**0.110**	0.113	0.162	**0.102**	0.140	0.242	**0.074**	0.101	0.187	**0.086**	0.110	0.195	**0.069**	0.076	0.069	**0.074**	0.060	0
	0.005	*0.008*	*0.024*	***0.017***	*0.028*	*0.061*	***0.013***	*0.018*	*0.043*	***0.022***	*0.039*	*0.154*	***0.018***	*0.024*	*0.057*	***0.013***	*0.020*	*0.106*
Mother's education	**0.123**	0.127	0.187	**0.104**	0.144	0.245	**0.094**	0.127	0.230	**0.092**	0.115	0.224	**0.099**	0.109	0.118	**0.098**	0.106	0.042
	0.006	*0.009*	*0.025*	***0.019***	*0.029*	*0.059*	***0.013***	*0.019*	*0.047*	***0.020***	*0.039*	*0.157*	***0.020***	*0.027*	*0.054*	***0.013***	*0.021*	*0.125*
Birth region	**0.052**	0.035	0.021	**0.017**	0.006	0.000	**0.015**	0.017	0.000	**0.025**	0.036	0.103	**0.056**	0.060	0.093	**0.044**	0.055	0.052
	0.005	*0.007*	*0.025*	***0.010***	*0.020*	*0.062*	***0.008***	*0.015*	*0.047*	***0.014***	*0.022*	*0.086*	***0.015***	*0.021*	*0.040*	***0.011***	*0.019*	*0.057*

Sample individuals 30-49 with positive labor earnings and information on a set of circumstances; standard errors in italics; (a) father’s occupation is missing for Colombia and Peru.

Table 6: Reduced-Form OLS Regression of Earnings on Observed Circumstances.

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Female	-0.589***	-0.487***	-0.799***	-1.028***	-0.399***	-0.638***
	[0.009]	[0.028]	[0.030]	[0.060]	[0.042]	[0.032]
Member of an ethnic minority	-0.364***	-0.009	-0.111***	-0.261***	-0.758***	-0.175***
	[0.009]	[0.043]	[0.038]	[0.059]	[0.101]	[0.036]
Father agricultural worker	-0.366***		-0.322***	-0.090	-0.320***	
	[0.009]		[0.031]	[0.056]	[0.046]	
Father primary education	0.206***	0.204***	0.103***	0.174***	0.125*	0.240***
	[0.011]	[0.032]	[0.035]	[0.069]	[0.065]	[0.042]
Father secondary education	0.559***	0.600***	0.420***	0.396***	0.369***	0.456***
	[0.019]	[0.057]	[0.058]	[0.121]	[0.079]	[0.049]
Mother primary education	0.243***	0.220***	0.291***	0.349***	0.303***	0.165***
	[0.011]	[0.033]	[0.035]	[0.097]	[0.063]	[0.041]
Mother secondary education	0.644***	0.608***	0.634***	0.689***	0.603***	0.486***
	[0.019]	[0.061]	[0.059]	[0.125]	[0.081]	[0.051]
Birth region 2	-0.353***	0.197***	-0.183***	-0.195**	-0.008	0.076**
	[0.013]	[0.030]	[0.030]	[0.077]	[0.056]	[0.033]
Birth region 3	-0.597***	0.427***	0.133***	-0.245***	-0.287***	0.365***
	[0.015]	[0.047]	[0.041]	[0.070]	[0.058]	[0.042]
Constant	6.450***	12.262***	5.084***	7.005***	5.263***	6.709***
	[0.014]	[0.032]	[0.034]	[0.075]	[0.079]	[0.038]
Observations	50560	16575	9259	4661	4127	9830
R-squared	0.35	0.15	0.20	0.22	0.22	0.19

Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.
Omitted categories are: male, ethnic majority, father and mother with no or unknown education, and birth region 1 (see Table 2 for the country-specific definitions).

Table 7: Inequality of Opportunity Indices for Household Income (per capita)

	Brazil			Colombia			Ecuador			Guatemala			Panama			Peru		
	E(0)	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)
Total inequality	**0.695**	0.710	1.595	**0.559**	0.626	1.377	**0.417**	0.452	0.828	**0.619**	0.683	1.722	**0.630**	0.609	1.085	**0.557**	0.635	1.642
	0.009	*0.015*	*0.152*	***0.027***	*0.038*	*0.141*	***0.017***	*0.027*	*0.108*	***0.035***	*0.040*	*0.320*	***0.032***	*0.037*	*0.105*	***0.024***	*0.041*	*0.222*
Non parametric estimates																		
θ_d^N	**0.329**	0.337	0.191	**0.250**	0.261	0.157	**0.290**	0.287	0.187	**0.373**	0.386	0.209	**0.346**	0.335	0.213	**0.292**	0.271	0.124
	0.008	*0.010*	*0.015*	***0.020***	*0.021*	*0.013*	***0.020***	*0.022*	*0.024*	***0.032***	*0.040*	*0.061*	***0.023***	*0.025*	*0.024*	***0.019***	*0.017*	*0.016*
θ_r^N	**0.329**	0.319	0.416	**0.250**	0.287	0.397	**0.290**	0.315	0.421	**0.373**	0.419	0.587	**0.346**	0.322	0.304	**0.292**	0.337	0.418
	0.008	*0.014*	*0.054*	***0.020***	*0.033*	*0.068*	***0.020***	*0.026*	*0.047*	***0.032***	*0.035*	*0.079*	***0.023***	*0.037*	*0.098*	***0.019***	*0.033*	*0.125*
Parametric estimates																		
θ_r^P	**0.322**	0.305	0.382	**0.233**	0.259	0.350	**0.269**	0.284	0.365	**0.345**	0.371	0.498	**0.315**	0.274	0.233	**0.279**	0.302	0.321
	0.009	*0.016*	*0.065*	***0.019***	*0.034*	*0.083*	***0.020***	*0.027*	*0.055*	***0.031***	*0.041*	*0.108*	***0.022***	*0.039*	*0.113*	***0.018***	*0.030*	*0.135*
θ_r^J																		
Race	**0.086**	0.079	0.107	**0**	0	0.001	**0.022**	0.020	0.028	**0.082**	0.077	0.102	**0.066**	0.036	0.037	**0.044**	0.038	0.035
	0.003	*0.004*	*0.016*	***0.002***	*0.002*	*0.004*	***0.004***	*0.004*	*0.006*	***0.012***	*0.011*	*0.014*	***0.009***	*0.006*	*0.007*	***0.008***	*0.008*	*0.026*
Father's occupation	**0.047**	0.044	0.062				**0.095**	0.091	0.113	**0.052**	0.052	0.079	**0.061**	0.053	0.057			
	0.002	*0.003*	*0.011*				***0.010***	*0.013*	*0.026*	***0.011***	*0.012*	*0.019*	***0.011***	*0.013*	*0.035*			
Father's education	**0.132**	0.142	0.222	**0.152**	0.178	0.276	**0.117**	0.126	0.173	**0.145**	0.164	0.257	**0.101**	0.096	0.116	**0.120**	0.113	0.112
	0.006	*0.010*	*0.039*	***0.017***	*0.023*	*0.047*	***0.014***	*0.018*	*0.038*	***0.024***	*0.029*	*0.062*	***0.016***	*0.017*	*0.034*	***0.014***	*0.017*	*0.049*
Mother's education	**0.145**	0.156	0.244	**0.153**	0.179	0.278	**0.154**	0.167	0.227	**0.203**	0.225	0.341	**0.163**	0.154	0.159	**0.172**	0.184	0.249
	0.006	*0.008*	*0.034*	***0.017***	*0.023*	*0.048*	***0.014***	*0.020*	*0.044*	***0.023***	*0.031*	*0.081*	***0.020***	*0.028*	*0.070*	***0.016***	*0.020*	*0.048*
Birth region	**0.079**	0.049	0.043	**0.031**	0.019	0.00	**0.029**	0.030	0.036	**0.046**	0.051	0.098	**0.085**	0.078	0.096	**0.076**	0.092	0.158
	0.005	*0.007*	*0.024*	***0.012***	*0.019*	*0.046*	***0.010***	*0.015*	*0.035*	***0.015***	*0.018*	*0.041*	***0.015***	*0.017*	*0.036*	***0.011***	*0.016*	*0.045*

Sample: individuals 30-49 with positive income and information on a set of circumstances; standard errors in italics; father's occupation missing for Colombia and Peru.

Table 8: Inequality of Opportunity Indices for Household Consumption Expenditures (per capita)

	COLOMBIA			ECUADOR			GUATEMALA			PANAMA			PERU		
	E(0)	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)	**E(0)**	E(1)	E(2)
TOTAL INEQUALITY	**0.449**	0.503	1.013	**0.354**	0.375	0.574	**0.409**	0.436	0.676	**0.381**	0.374	0.539	**0.351**	0.384	0.660
	0.018	*0.024*	*0.079*	***0.015***	*0.018*	*0.047*	***0.024***	*0.023*	*0.039*	***0.016***	*0.019*	*0.042*	***0.015***	*0.022*	*0.076*
NON PARAMETRIC ESTIMATES															
θ_d^N	**0.265**	0.275	0.177	**0.344**	0.347	0.270	**0.524**	0.536	0.440	**0.417**	0.385	0.285	**0.348**	0.339	0.229
	0.017	*0.017*	*0.013*	***0.021***	*0.025*	*0.028*	***0.023***	*0.026*	*0.031*	***0.016***	*0.018*	*0.020*	***0.017***	*0.017*	*0.016*
θ_r^N	**0.265**	0.304	0.456	**0.344**	0.353	0.427	**0.524**	0.542	0.630	**0.417**	0.405	0.475	**0.348**	0.389	0.533
	0.017	*0.023*	*0.035*	***0.021***	*0.024*	*0.033*	***0.023***	*0.023*	*0.022*	***0.016***	*0.024*	*0.044*	***0.017***	*0.024*	*0.040*
PARAMETRIC ESTIMATES															
θ_r^P	**0.244**	0.271	0.408	**0.321**	0.326	0.389	**0.503**	0.519	0.606	**0.386**	0.362	0.417	**0.340**	0.375	0.512
	0.017	*0.023*	*0.041*	***0.022***	*0.028*	*0.042*	***0.020***	*0.020*	*0.021*	***0.016***	*0.023*	*0.046*	***0.017***	*0.022*	*0.036*
θ_r^J															
Race	**0.001**	0.001	0.002	**0.032**	0.027	0.036	**0.141**	0.123	0.136	**0.121**	0.065	0.047	**0.054**	0.051	0.065
	0.002	*0.002*	*0.003*	***0.006***	*0.005*	*0.006*	***0.013***	*0.011*	*0.013*	***0.014***	*0.012*	*0.016*	***0.008***	*0.007*	*0.008*
Father's occupation				**0.106**	0.103	0.120	**0.073**	0.071	0.088	**0.071**	0.069	0.090			
				0.010	*0.011*	*0.018*	***0.011***	*0.011*	*0.015*	***0.011***	*0.010*	*0.013*			
Father's education	**0.154**	0.179	0.288	**0.141**	0.148	0.192	**0.202**	0.219	0.285	**0.108**	0.109	0.146	**0.142**	0.147	0.199
	0.016	*0.019*	*0.028*	***0.015***	*0.019*	*0.032*	***0.021***	*0.021*	*0.027*	***0.014***	*0.016*	*0.023*	***0.013***	*0.014*	*0.019*
Mother's education	**0.166**	0.194	0.314	**0.186**	0.193	0.244	**0.256**	0.279	0.353	**0.167**	0.177	0.241	**0.204**	0.222	0.306
	0.014	*0.018*	*0.028*	***0.015***	*0.019*	*0.032*	***0.019***	*0.021*	*0.028*	***0.017***	*0.020*	*0.029*	***0.014***	*0.017*	*0.025*
Birth region	**0.040**	0.032	0.035	**0.040**	0.044	0.058	**0.102**	0.103	0.137	**0.096**	0.096	0.127	**0.109**	0.124	0.195
	0.011	*0.016*	*0.033*	***0.010***	*0.014*	*0.025*	***0.014***	*0.015*	*0.020*	***0.015***	*0.016*	*0.023*	***0.013***	*0.019*	*0.038*

Sampl: individuals 30-49 with positive household consumption and information on a set of circumstances; standard errors in italics; father's occupation is missing for Colombia and Peru.

Table 9: Reduced-Form OLS Regression of Household Income on Observed Circumstances

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Member of an ethnic minority	-0.414***	-0.015	-0.236***	-0.350***	-0.922***	-0.230***
	[0.009]	[0.032]	[0.025]	[0.036]	[0.057]	[0.023]
Father agricultural worker	-0.295***		-0.291***	-0.192***	-0.210***	
	[0.009]		[0.019]	[0.034]	[0.032]	
Father primary education	0.248***	0.218***	0.135***	0.166***	0.211***	0.189***
	[0.010]	[0.022]	[0.021]	[0.041]	[0.044]	[0.026]
Father secondary education	0.696***	0.689***	0.446***	0.386***	0.473***	0.442***
	[0.018]	[0.039]	[0.035]	[0.080]	[0.057]	[0.031]
Mother primary education	0.299***	0.245***	0.254***	0.362***	0.298***	0.215***
	[0.010]	[0.023]	[0.021]	[0.058]	[0.043]	[0.026]
Mother secondary education	0.789***	0.703***	0.630***	0.893***	0.768***	0.572***
	[0.018]	[0.039]	[0.035]	[0.082]	[0.059]	[0.034]
Birth region 2	-0.378***	0.176***	-0.124***	-0.278***	-0.044	0.058***
	[0.012]	[0.022]	[0.018]	[0.051]	[0.043]	[0.021]
Birth region 3	-0.740***	0.460***	0.170***	-0.258***	-0.361***	0.376***
	[0.013]	[0.028]	[0.025]	[0.045]	[0.042]	[0.030]
Constant	5.372***	11.578***	4.415***	5.857***	6.977***	5.786***
	[0.013]	[0.022]	[0.020]	[0.046]	[0.053]	[0.023]
Observations	71,670	22,436	12,643	6,847	5,649	13,649
R-squared	0.32	0.19	0.24	0.28	0.34	0.25

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Omitted categories are: male, ethnic majority, father and mother with no or unknown education, and birth region 1 (see Table 2 for the country-specific definitions).

Table 10: Reduced-Form OLS Regression of Household Consumption Expenditures on Observed Circumstances

	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Member of an ethnic minority	-0.043 [0.028]	-0.300*** [0.023]	-0.411*** [0.025]	-1.020*** [0.039]	-0.195*** [0.017]
Father agricultural worker		-0.254*** [0.017]	-0.190*** [0.023]	-0.202*** [0.023]	
Father primary education	0.196*** [0.019]	0.159*** [0.018]	0.130*** [0.031]	0.171*** [0.031]	0.145*** [0.019]
Father secondary education	0.615*** [0.031]	0.470*** [0.032]	0.472*** [0.046]	0.368*** [0.042]	0.364*** [0.022]
Mother primary education	0.234*** [0.019]	0.275*** [0.018]	0.374*** [0.038]	0.189*** [0.030]	0.196*** [0.019]
Mother secondary education	0.693*** [0.033]	0.672*** [0.031]	0.725*** [0.052]	0.554*** [0.043]	0.498*** [0.025]
Birth region 2	0.202*** [0.018]	-0.094*** [0.016]	-0.416*** [0.033]	-0.028 [0.032]	0.051*** [0.016]
Birth region 3	0.454*** [0.023]	0.215*** [0.022]	-0.283*** [0.028]	-0.257*** [0.030]	0.411*** [0.023]
Constant	14.216*** [0.020]	3.964*** [0.018]	8.520*** [0.029]	6.990*** [0.039]	5.877*** [0.017]
Observations	22,487	12,643	6,865	5,686	13,649
R-squared	0.22	0.31	0.47	0.42	0.31

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Omitted categories are: male, ethnic majority, father and mother with no or unknown education, and birth region 1 (see Table 2 for the country-specific definitions).

Table 11: Opportunity-deprivation profiles

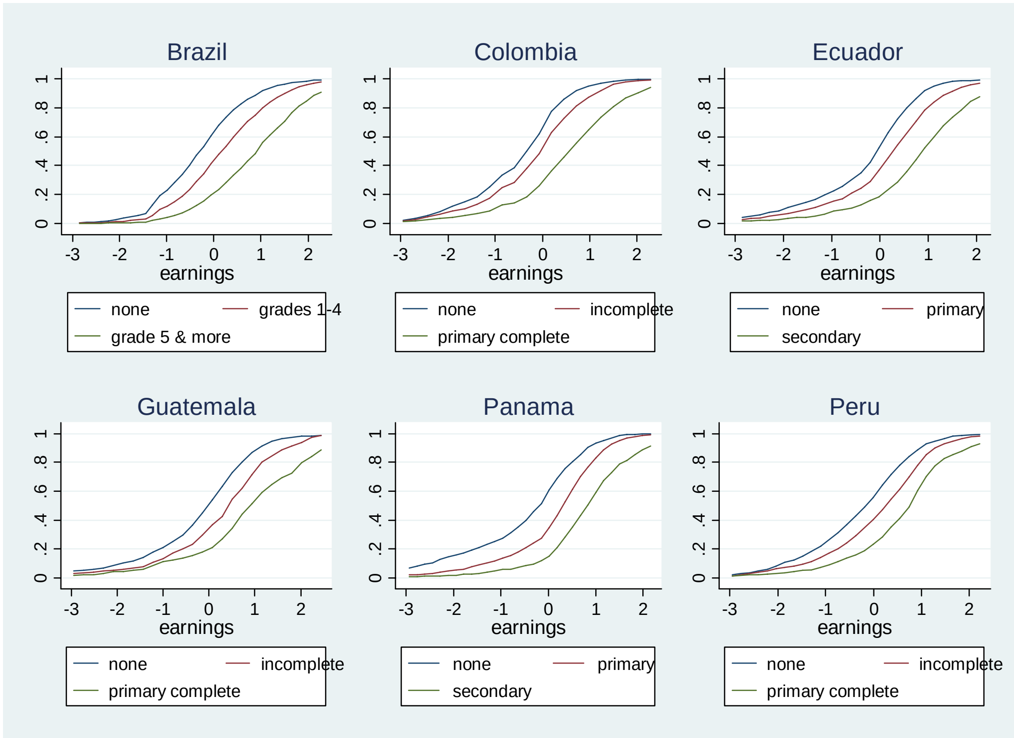
	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Member of ethnic minority	100.0	32.8	61.0	100.0	75.9	100.0
Father's agricultural occupation	87.9		93.4	99.9	83.5	
Other father's occupation	12.1		6.6	0.1	16.5	
Father without education	89.2	76.6	86.9	99.4	58.0	99.8
Father's primary education	10.5	23.4	11.2	0.3	37.0	0.2
Father's secondary education (or complete primary)	0.3	0.0	1.9	0.3	5.0	0.0
Mother without education	90.7	96.0	98.3	99.1	92.6	99.4
Mother's primary education	9.3	3.8	1.1	0.3	5.7	0.0
Mother's secondary education (or complete primary)	0.0	0.2	0.6	0.6	1.7	0.6
Birth regions	Northeast and North (100%)	Periphery (99%)	Coast and insular (51%), Sierra and Amazonia (48%)	North and Northwest (99%)	Rural areas (96%)	South and Coast (58%), inland (42%)
Share of total outcome	2.9	5.0	4.4	3.5	2.7	4.8

Table 12: Poverty profiles

	BRAZIL	COLOMBIA	ECUADOR	GUATEMALA	PANAMA	PERU
Member of ethnic minority	68.5	14.9	31.8	70.2	53.7	56.4
Father's agricultural occupation	56.3		80.0	75.7	80.3	
Other father's occupation	43.7		20.0	24.3	19.7	
Father without education	77.2	57.8	55.0	90.1	66.8	59.9
Father's primary education	21.8	40.3	42.2	9.0	29.0	31.8
Father's secondary education (or complete primary)	1.0	1.9	2.8	0.9	4.2	8.3
Mother without education	79.4	53.5	59.7	96.3	75.2	82.5
Mother's primary education	19.4	44.6	38.5	3.0	23.7	15.4
Mother's secondary education (or complete primary)	1.1	1.9	1.8	0.7	1.1	2.2
Birth regions	Northeast and North (70%), Southeast, Center-west, South (28%)	Periphery (65%), Center (34%)	Sierra and Amazonia (48%), Coast and insular (45%)	North and Northwest (49%), South and Center (46%)	Rural areas (91%), small towns (5%)	Inland (59%), South and Coast (40%)
Share of total outcome	0.7	1.5	1.9	1.8	1.5	1.8

Figure 1: The distribution of earnings conditional on selected circumstance variables

a. By mother's education



b. By ethnicity

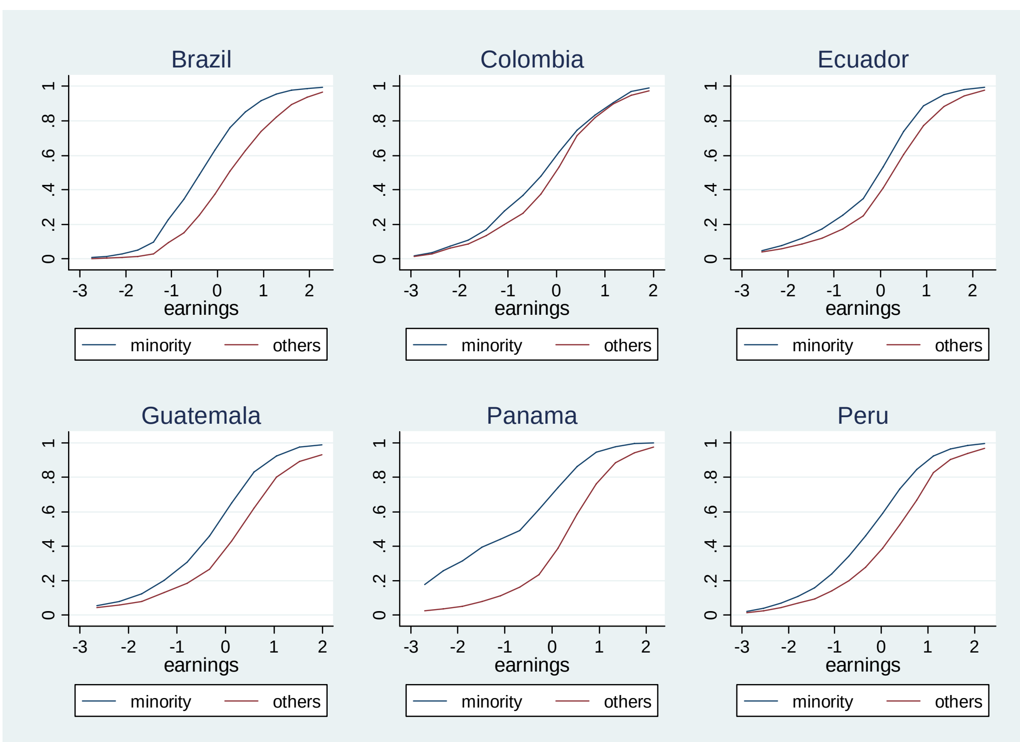
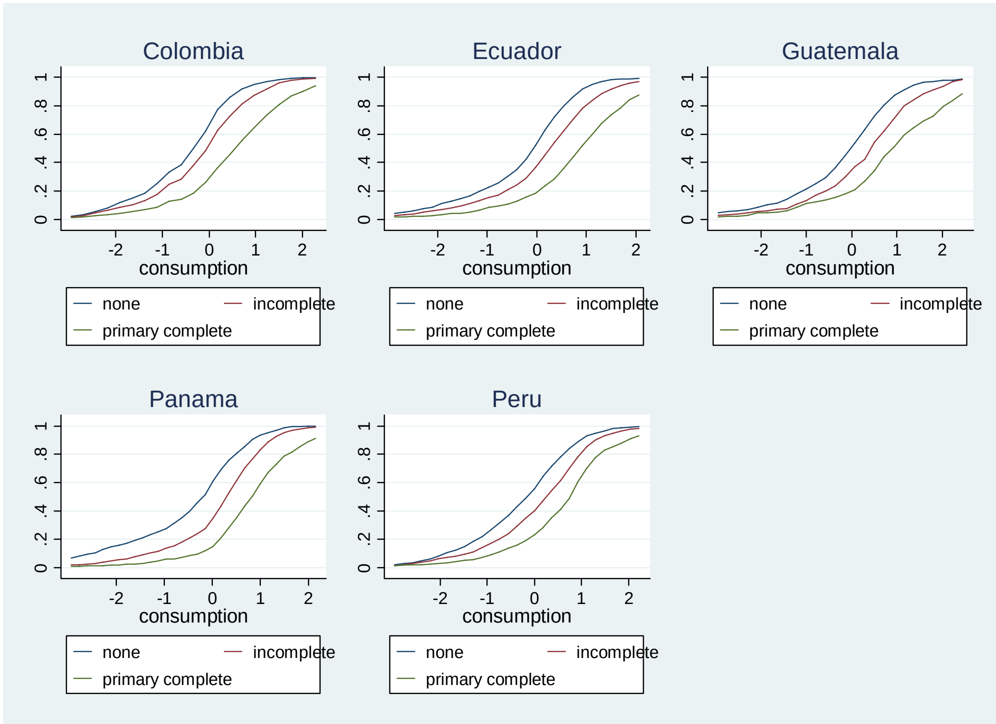


Figure 2: The distribution of consumption conditional on selected circumstance variables

a. By mother's education



b. By ethnicity