Oxford Poverty & Human Development Initiative (OPHI)
Oxford Department of International Development
Queen Elizabeth House (QEH), University of Oxford

OPHI WORKING PAPER NO. 28

A Dissimilarity Index of Multidimensional Inequality of Opportunity

Gaston Yalonetzky*

November 2009

Abstract

A recent literature on inequality of opportunity offers quantitative tools for comparisons and measurement based on stochastic dominance criteria and traditional inequality indices. In this paper I suggest an additional way of assessing inequality of opportunity and operationalizing Roemer's (1998) notion of equality of opportunity with an index of dissimilarity across distributions. The index is based on a traditional homogeneity test of multinomial distributions and is more helpful than other tools when both circumstances and advantages/outcomes are multidimensional. It also highlights the correspondence between dissimilarity in outcomes across sets of circumstances and the degree of association between circumstances and outcomes. An empirical application measures changes in inequality of opportunity from an old to a young cohort in Peru.

Keywords: Inequality of opportunity, dissimilarity indices.

JEL classification: D30, D63.

* Oxford Poverty and Human Development Initiative (OPHI), Department of International Development, University of Oxford

This study has been prepared within the OPHI theme on Multidimensional Poverty Measurement.

OPHI gratefully acknowledges support for its research and activities from the Government of Canada through the International Development Research Centre (IDRC) and the Canadian International Development Agency (CIDA), the Australian Agency for International Development (AusAID), and the United Kingdom Department for International Development (DFID).

ISSN 2040-8188 ISBN 978-1-907194-09-2

Acknowledgements

I would like to thank Francisco Ferreira, Maria Ana Lugo, James Foster, Sabina Alkire, Gary Fields, Casilda Lasso de la Vega, Maria Emma Santos, Suman Seth and participants at the World Bank LCR Seminar Series, Washington D.C., the Third Meeting of ECINEQ in Buenos Aires, the HDCA Conference in Lima, and seminar participants at OPHI, QEH for very helpful comments and fruitful discussions. I am also grateful to Lorenzo Oimas for providing me with the database.

The Oxford Poverty and Human Development Initiative (OPHI) is a research centre within the Oxford Department of International Development, Queen Elizabeth House, at the University of Oxford. Led by Sabina Alkire, OPHI aspires to build and advance a more systematic methodological and economic framework for reducing multidimensional poverty, grounded in people's experiences and values.

Oxford Poverty & Human Development Initiative (OPHI)
Oxford Department of International Development
Queen Elizabeth House (QEH), University of Oxford
3 Mansfield Road, Oxford OX1 3TB, UK
Tel. +44 (0)1865 271915 Fax +44 (0)1865 281801
ophi@qeh.ox.ac.uk http://ophi.qeh.ox.ac.uk/

1. Introduction

The concern for inequality of opportunity has long earned its place in the social science and political philosophy literature. Following Roemer's (1998) influential conceptualization of it, recent research has sought to quantify inequality of opportunity and to compare its extent across societies. For instance, Lefranc *et al.* (2008) compare inequality of opportunity and of outcomes across developed countries[1] using stochastic dominance analysis and proposing a Gini index of inequality of opportunity. Checchi and Peragine (2005) measure inequality of opportunity in Italy based on traditional inequality indices which are decomposable in between-group and within-group elements. Ferreira and Gignoux (2008) extend the same between-group approach to a parametric framework and study inequality of opportunity in Latin America.[2] Barros *et al.* (2009) compile studies of inequality of opportunity in Latin America, including the advocacy of a human opportunity index that measures inequality in the attainment of a certain outcome (e.g. completion of an educational level) or in the access to a welfare-enhancing service (e.g. water, housing, etc.) by bringing together the average level of such outcome/service-access with its distribution across groups measured with a dissimilarity index borrowed from the Sociology literature. Similarly Lanjouw and Rao (2008) have applied new refinements on the decomposition of inequality indices (due to Elbers *et al.*, 2008) to tracking changes in caste-based inequality in two Indian villages during the last half of the past century.

On its own, the intergenerational economic mobility literature has a longer history of development of quantitative tools (e.g. tests, indices, etc.) and it is

[1] Belgium, France, United Kingdom, (West) Germany, Italy, Netherlands, Norway, Sweden and the US.

[2] They consider Brazil, Colombia, Ecuador, Guatemala, Panama and Peru for comparisons.

related both with the study of inequality of opportunity and economic inequality in several ways. However, since in the intergenerational mobility literature usually one parental attribute is related to one of the offspring's, a pair at a time, most of the toolkit is not well suited to studying proper inequality of opportunity with multiple circumstances and multiple outcomes/advantages to consider.

In this paper I contribute methodologically to the quantitative analysis of inequality of opportunity by suggesting the use of dissimilarity indices. The index proposed in this paper is based on the statistic of a traditional test of homogeneity of multinomial distributions. Similar indices can also be applied to compare the degree of dissimilarity across transition matrices.[3] When applied to inequality of opportunity comparisons, this dissimilarity index has the advantage over existing indices that it can compare multidimensional distributions of outcomes (or advantages in Roemer's vocabulary), which is an appealing trait in the burgeoning multidimensional welfare measurement literature.[4] Another interesting trait is that, unlike other indices in the recent literature[5], the dissimilarity index attains its minimum value, representing perfect *equality* of opportunity, *if and only if* the distributions of well-being outcomes conditioned on social groups are identical. Hence the concept of inequality of opportunity as *dissimilarity* across conditional distributions that the index is measuring is perfectly in line with a literalist interpretation of Roemer's characterizations of *equality* of opportunity whereby the latter is said to be achieved "if the cumulative distribution functions of advantages across types are identical" (Roemer, 2006, p. 8). Moreover the dissimilarity index attains its maximum value *if and only if* there is perfect association between

[3] As I propose in another paper. See Yalonetzky, 2009.
[4] Other already existing techniques may also be extended to deal with multiple dimensions in the future.
[5] An exception is Checci and Peragine's index based on their "tranches" approach.

the groups in which societies are partitioned and subsets of the welfare outcomes in consideration. The multinomial index is most suitable for discrete variables whereas for continuous ones it requires prior discretization.

Like other indices of inequality of opportunity, the dissimilarity index proposed in this paper is mostly focused on between-group inequalities, that is, on comparisons across conditional distributions of well-being. There is an existing rich and old literature of indices and measurement of between-group inequality. For instance, the work of Gastwirth (1975), Ebert (1984), Bulter and McDonald (1987), Dagum (1980, 1987) and more recently Handcock and Morris (1999) and Breton *et al.* (2008). Even though they have not been considered explicitly in the context of inequality of opportunity, these indices also declare equality whenever the distributions are identical and generally declare maximum inequality between distributions whenever the latter do not overlap at all. However these measures have been devised for continuous variables,[6] univariate outcomes, and in all the mentioned cases they are meant to work for comparisons of pairs of groups, as opposed to several groups like the dissimilarity index.

In an empirical application I study changes in inequality of opportunity in Peru from older to younger cohorts of adults. Considering as welfare outcomes years of education and quality of education measured by type of school attended (public versus private versus none), I find a small but statistically significant reduction in inequality of opportunity among the younger cohort. The groups of people (or types in Roemer's vocabulary) are defined by combinations of gender, paternal and maternal education levels.[7]

[6] Although they could be adapted in some cases for discrete variables, as in the case of Handcock and Morris (1999).

[7] These grouping criteria are circumstances beyond the control of individuals, as is demanded by the inequality of opportunity approach. See, for instance Roemer (1998).

In the next section the dissimilarity index is introduced and its behaviour is investigated. A comparison of their orderings to those of existing quantitative tools in the literature is also provided. Then an empirical application to Peru follows. The paper finishes with a conclusions section.

2. The dissimilarity index of multidimensional inequality of opportunity

Recent operationalizations of Roemer's (1998, 2006) definitions of inequality of opportunity follow his concepts of circumstances, efforts and advantages. Checci and Peragine (2005) develop two indices of inequality of opportunity based on two interpretations of the definition of equality of opportunity by Roemer (1998, 2006). A first, literalist interpretation says that equality of opportunity is achieved only when the distribution of an advantage across the population is independent of the set of circumstances. That is, circumstances should not affect the advantage either directly or indirectly through effort or random shocks and therefore the cumulative distributions of the advantages should be identical across social groups defined by sets of circumstances. The first index of Checci and Peragine (2005) measures the contribution of inequality of opportunity to total inequality by decomposing the latter into a component that measures the degree of inequality across social groups within percentiles of their respective cumulative outcome distributions and a component that measures inequality between mean values of the outcome for every percentile across the total population. They call it the tranches approach. Like this paper's index, the tranches approach declares equality of opportunity if and only if the group-conditioned cumulative distributions are identical. Hence they measure equality of opportunity in agreement with a literalist interpretation of Roemer's definition.

Roemer (2006) proposes a less literalist operationalization of his definition by suggesting comparing mean achievements across population groups and declaring equality of opportunity whenever these means are identical.[8] Therefore equality of opportunity according to the literalist interpretation implies equality according to his less literalist operationalization but the converse is not true. In relation to Roemer's mean-equalization proposal, Checci and Peragine (2005) and Ferreira and Gignoux (2008), develop a similar index of relative inequality of opportunity in which the contribution to total inequality due to opportunity inequality is captured by differences in the mean attainment across social groups.[9] This index declares equality of opportunity according to Roemer's mean-equalization proposal but may declare equality of opportunity even in situations in which a literalist interpretation of Roemer's definition would not agree. Checci and Peragine call it the types approach.

Lefranc et al. (2008) follow a similar reasoning to propose their Gini index of inequality of opportunity. But they depart from Roemer (1998) when they propose an alternative definition of equality of opportunity according to which the latter is achieved when there are no sets of circumstances which are second-order dominated within a society. Guided by concerns over return and risk of the outcome from different circumstance sets, in the definition of Lefranc *et al.* (2008) circumstances may actually affect advantages differentially but equality of opportunity is still

[8] This definition is important in his later advocacy of tracking economic development by focusing on the growth of the mean outcome of the most disadvantaged social groups in society as oppossed to overall mean per capita achievement.

[9] Both Checci and Peragine, and Ferreira and Gignoux advocate using the between-group inequality indicator alone as a measure of absolute inequality of opportunity. Ferreira and Gignoux and Elbers *et al.* also advocate using the ratio of between-group inequality to total inequality as a relative measure.

deemed to exist as long as individuals can not rank any circumstances according to second-order stochastic dominance among their respective outcome distributions.

Therefore in their definition of equality of opportunity the properties of the outcome lottery (e.g. average return and risk) faced by people with different circumstances matter in the sense that people may find some circumstances more appealing than others in terms of their related outcome or advantage lotteries. In their framework, if people could choose their circumstance before being born on the grounds of such appeal (formally using a second-order stochastic dominance criterion) and in turn they happened to be indifferent between circumstances, based on that criterion, then their society could be deemed as showing equality of opportunity.

The dissimilarity index introduced in this paper relates to a literalist definition of Roemer (1998; 2006, p. 8) in which equality of opportunity is achieved if and only if the conditional distributions of outcomes/advantages are equal across circumstance sets. This particular definition of equality of opportunity can be further characterized conceptually by stating that it relates to a situation in which both Roemer's *assumption of charity* and Fleurbaey's *equal well-being for equal responsibility* (Fleurbaey, 2008) hold. The *assumption of charity* says that individuals belonging to different types (defined by combinations of circumstances beyond their control) would exhibit the same distributions of effort should their defining circumstances be factored out (Roemer, 1998, p. 16).[10] An allocation of resources following the criterion of equal well-being for equal responsibility is characterized by an equalization of a well-being outcome across types for every

[10] Roemer (1998) uses this concept then to justify his proposal for equal remuneration for individuals belonging to different type but exerting the same relative degree of effort within the distributions of their respective types.

different level of dedication.[11] An equivalent way to characterize this definition is to associate perfect equality of opportunity with Fleurbaey's *circumstance neutralization*, a situation in which individual well-being can only be expressed as a function of responsibility characteristics (i.e. dedication, or Roemer's effort), and not of circumstances.[12] Should any of these conditions fail to hold then the distributions of well being conditioned on type-belonging would not be identical and viceversa.

The types approach of Checci and Peragine and Roemer's mean-equalization operationalization, emphasizes the connection between inequality of opportunity and decomposed between-group inequality of outcomes (where groups are defined according to circumstances). By contrast, the dissimilarity index of inequality of opportunity highlights the association between sets of circumstances, so-called types in Roemer's terminology, and sets or values of advantages/outcomes. In fact the index achieves its maximum value, with which it signals maximum inequality of opportunity only whenever there is perfect or maximum association between circumstances and advantage, which does not necessarily imply perfect correlation since the latter is but one form of perfect association between variables. On the other hand the index achieves its minimum value only when the conditional distributions of outcomes are all identical, i.e. homogeneous, which implies that the conditioning factors are irrelevant in determining the advantages. The index therefore measures a concept of inequality of opportunity based on the degree

[11] This is a concept similar to Roemer's effort and refers to a person's use of resources in order to improve his/her wellbeing. See Fleurbaey (2008, chapter 1).

[12] Ferreira and Gignoux (2008) also elaborate on this point. Fleurbaey (2001) highlights that there may be differences across groups in the ability of individuals to attain the same percentile of the wellbeing distribution corresponding to their respective type/group. Should that be the case, arguably the presence of identical conditional distributions may not suffice to establish equality of opportunity even under the conditions described in the paragraph.

of dissimilarity of multinomial distributions, in turn captured by the metric of a Pearson goodness-of-fit statistic used to test homogeneity of such distributions.

To define the index formally, let's start by assuming that societies can be partitioned into a set of individuals' types. Each type itself is defined by a special combination of values taken by a vector of circumstances, i.e. factors over which the individual does not exert control, like parental education, ethnicity or gender. For instance, imagine a society with two circumstances: gender (male or female) and parental education ("low" or "high"). In such a society type "1" individuals could be those who are male (meaning arbitrarily a value of "1" in the gender entry) and whose parents had "low" education (meaning arbitrarily a value of "1" in the parental education entry). By combining the different categories within each and every circumstance, four types are defined in this example.

In general z circumstances are considered, each of which is partitioned into g_i categories (for $i = 1, 2, \ldots z$), making every circumstance a vector, V_i, with g_i elements. (For instance, a gender vector would have just two elements). By combining all the possible values in the vectors of circumstances a vector of types is defined. Formally, types are generated by a function f that transforms combinations of circumstance values into a natural number representing the ensuing type:

$$f : V_1 \times V_2 \times \ldots \times V_z \rightarrow \mathbb{N}_+^T.$$

The ensuing vector of types, $G = \{1, 2, ..., T\}$, has $T = \prod_{i=1}^{z} g_i$ elements.[13] All individuals having the same set of circumstances are said to be of the same

[13] Circumstances could also be continuous, which might require discretization in practical applications of this index.

type (e.g. in the U.S. context assuming $z = 4$, one set determining a type could be being an adult Asian male whose two parents had achieved complete secondary education). Empirically, the absolute frequency of people in a society belonging to type t, such that $t \in G$, is denoted by N^t and the total population sample is N.

Similarly outcomes or advantages can be considered in a multidimensional way. All possible combinations of outcomes (e.g. health status with education achievement and earnings and so on) are in the vector $O = \{1, 2, ..., A\}$. Assuming that there are b outcome vectors, V^j, each having m_j elements (for $j = 1, 2, \ldots, b$), then multidimensional outcomes are generated by a function q that transforms combinations of individual outcomes into multidimensional outcomes:

$$q : V^1 \times V^2 \times \ldots \times V^b \rightarrow \mathbb{N}_+^A.$$

O has $A = \prod_{j=1}^{b} m_j$ elements, each of which represents a combination of outcomes, each one partitioned in the aforementioned m_j elements or categories. For instance an element $\alpha \in O$ and equal to "1" might stand for having tertiary education, excellent health status and the highest earning capacity (i.e. the categories can represent intervals too). The absolute frequency of people in a society attaining outcome α is N_α. Finally, the probability of attaining a given combination of advantages (e.g. $\alpha = k$) conditional on being of type t is: p_k^t. The corresponding absolute frequency of people being of type t and attaining a combination k is N_k^t.[14]

The index of dissimilarity advocated in this paper belongs to a general class of statistics which measure the degree of dissimilarity between distributions as the

[14] Therefore $\sum_{t=1}^{T} \sum_{\alpha=1}^{A} N_k^t = N$

degree of association between row variables and column variables in a contingency table. For instance, the column variable may represent the conditioning variable (e.g., in our context, the types) and the row variable may represent the outcome variable. The general class, X^*, of statistics is defined as follows:

$$X^{\beta}_{T,A} \in X^* \mid X^{\beta}_{T,A} \equiv \sum_{t=1}^{T}\sum_{\alpha=1}^{A} N^t \frac{|p^t_\alpha - p^*_\alpha|^\beta}{p^*_\alpha} \ \forall \beta, A, T \in \mathbb{N}_{++}, \tag{1}$$

Where p^*_α is a weighted average of the group-specific probabilities for outcome state α in which the weights are given by the share of each sample size on the total sum of them. It is the pooled-sample probability of having outcome α and it is calculated the following way:

$$p^*_\alpha = \sum_{t=1}^{T} p^t_\alpha \frac{N^t}{\sum_{t=1}^{T} N^t} = \frac{\sum_{t=1}^{T} N^t_\alpha}{\sum_{t=1}^{T} N^t}. \tag{2}$$

The weighted average probability performs the comparison of the probabilities across the different types' samples. The closer the respective probabilities across samples then the more the weighted average probability resembles each and every of its constituting probabilities (in (2)) and therefore the closer to zero the statistic in (3) is. Before the index is presented notice the following aspects about this general class:

- $T = 1 \rightarrow X^{\beta}_{T,A} = 0$ by construction. This is a trivial case. Therefore the focus of analysis is on a restricted subgroup of the class for which $T > 1$.
- $\sum_{\alpha}^{A} p^t_\alpha = \sum_{\alpha}^{A} p^*_\alpha = 1$ if and only if $A > 1$. This is an important detail to bear in mind when I compare below the behaviour of the index when $A = 2$ with the dissimilarity index used by the Human Opportunity Index (HOI) of

Barros *et al.* (2009). The latter dissimilarity index is based on $X^{\beta}_{T,A}$ when $\beta, A = 1$. When $A = 1$, $0 \leq p^t_\alpha \leq 1 \wedge 0 \leq p^*_\alpha \leq 1$, and their dissimilarity index, D, is defined as: $D \equiv X^1_{T,1}/2N$.

The dissimilarity index introduced in this paper is based on the statistic of a test of homogeneity among multinomial distributions (e.g. see Hogg and Tanis, 1997) that ensues from the general class, X^*, when $\beta = 2$:

$$X^2_{T,A} = \sum_{t=1}^{T}\sum_{\alpha=1}^{A} N^t \frac{(p^t_\alpha - p^*_\alpha)^2}{p^*_\alpha}. \tag{3}$$

The null hypothesis of the test is that the T distributions are homogenous, i.e. identical in a statistical sense. Formally: $H_O : p^1_\alpha = p^2_\alpha = \ldots = p^T_\alpha \; \forall \alpha = 1, \ldots, A$[15]

The statistic in (3) has an asymptotic chi-square distribution with $(T-1)(A-1)$ degrees of freedom under the null hypothesis of homogeneity. Besides being related to a standard test of multinomial distributions, another key advantage justifying the choice of $X^2_{T,A}$ (among many other options from class X^*) to build a dissimilarity index of multidimensional inequality of opportunity is that this statistic also has a maximum value which conveniently depends only on the number of samples/groups (e.g. the number of types), the number of states (e.g. the number of values that the outcome variable can take, that is the categories of multidimensional outcomes) and on the total population size, N. The maximum value is easily found by noticing that the statistic of the homogeneity test of multinomial distributions is a Pearson goodness-of-fit statistic:

[15] An alternative likelihood ratio statistic for the same test is asymptotically identical and has the following form: $LR = 2\sum_{t=1}^{T}\sum_{\alpha=1}^{A} N^t \log\left[\frac{N^t_\alpha}{N^t}\frac{\sum_{t=1}^{T} N^t}{\sum_{t=1}^{T} N^t_\alpha}\right]$

$$X_{T,A}^2 = \sum_{t=1}^{T}\sum_{\alpha=1}^{A} N^t \frac{\left(p_\alpha^t - p_\alpha^*\right)^2}{p_\alpha^*} = \sum_{t=1}^{T}\sum_{\alpha=1}^{A} \frac{\left(N_\alpha^t - \frac{N^t N_\alpha}{N}\right)^2}{\frac{N^t N_\alpha}{N}}. \qquad (4)$$

Intuitively one can bring together all the conditional probability vectors, i.e. the multidimensional distributions of outcomes conditional on a given type, to form a contingency table. In such a table N_k^t would be the observed frequency of individuals from a type "t" exhibiting a level of multidimensional advantage k; whereas the expected frequency for "t" and k under the null hypothesis of lack of association between circumstances and advantages/outcomes would be given by the expression $\frac{N^t N_\alpha}{N}$ (see for instance, Everitt, 1992). Therefore, (4) can be expressed as:

$$X_{T,A}^2 = \sum_{t=1}^{T}\sum_{\alpha=1}^{A} \frac{\left(OB_\alpha^t - E_\alpha^t\right)^2}{E_\alpha^t}. \qquad (5)$$

Where the OB stands for observed and the E for expected frequency. Cramer (1946) showed that the maximum for an expression like (5) is $min(T-1, A-1)N$. Hence the corresponding maximum for the statistic (3) is:

$$X_{T,A,\max}^2 = \min(T-1, A-1)N. \qquad (6)$$

Thus combining (3) and (6) the dissimilarity index is defined as:

$$H_{T,A}^2 = \frac{X_{T,A}^2}{X_{T,A,\max}^2}. \qquad (7)$$

The index is advocated for applications in which $T, A > 1$.[16] It fulfils axioms

[16] For $A = 1$ I advocate using $X_{T,1}^1$. The triviality of $T = 1$ has been discussed above.

of population invariance[17] and scale invariance.[18] Notice that an advantage of the dissimilarity index, based on the statistic in (3), is that it can be used to assess inequality of opportunity with multiple outcomes. It is also normalized in order to take the value of 0 when the samples under comparison (i.e. the conditional probability vectors) are identical, which would reflect equality of opportunity, at least in terms of the types considered. And it takes the value of 1 if and only if there is maximum association between circumstances and outcomes.[19]

In the context of the dissimilarity index (and in general, of contingency tables analysis) maximum association has three related meanings depending on whether $T < A$, $T > A$ or $T = A$. When $T < A$ (more outcome categories or states than types) maximum association means that for any arbitrary partition of the sets G and O into non-overlapping subgroups then:

$\forall k \in G, O_k \subset O : O_k \rightleftarrows k$

$\wedge$

$O_1 \cup O_2 \cup \ldots \cup O_T = O$,

where O_k is a subset of O made of all those outcome elements attained by type k with positive probability. Maximum or perfect association means therefore that for every type there is a vector of outcomes which is a subgroup of the outcome vector and is only attainable by that type. For instance, if type t_1 is associated with outcomes α_3 and α_4 (i.e. that there exists a positive probability of being in outcomes α_3 or α_4 conditional on being type t_1), then no other type is associated

[17] That is, if every individual in society is replicated n times, the value of the index remains unaltered.

[18] That is, if the measurements of outcomes are altered proportionately (or additively) in the same way as the boundaries of the partitions of outcomes are altered (i.e. the boundaries that determine whether for one individual $\alpha = k$), then the index's value remains unaffected.

[19] As mentioned, perfect correlations, positive or negative, are just examples of maximum association.

with those categories, and similarly if type t_2 is associated with outcomes α_5 and α_6 then type t_1 is not associated with those latter outcomes. The concept of maximum association is not a concept of perfect predictability because if a type is associated with more than one outcome grids (as in the aforementioned examples) then one cannot perfectly predict the final outcome (e.g. it could be either α_3 or α_4 if the type is t_1) although one can accurately predict that someone with type t_1 never attains outcomes α_5 or α_6.

When $T > A$ (more types than outcomes) the roles of types and outcome values are reversed, so maximum association means:

$\forall \alpha \in O, G_\alpha \subset G : G_\alpha \rightleftarrows \alpha$

$\wedge$

$G_1 \cup G_2 \cup \ldots \cup G_A = G,$

where G_α is a subset of G made of all those types who attain outcome α with positive probability. Maximum or perfect association means in this case that for every outcome state, or value/category, there is a vector formed by all and only the types that attain that specific outcome state. Any other subgroup of types can not attain that outcome and/or any other outcome is associated with a different, non-overlapping subgroup of types.

When $T = A$ maximum association implies that every type is associated exclusively with only one outcome and the reverse holds true: every outcome is associated exclusively with only one type:

$\forall \alpha \in O, k \subset G : k \rightleftarrows \alpha$

This concept of perfect or maximum association seems suitable for a benchmark of maximum inequality of opportunity when multiple circumstances and multiple outcomes are considered, since there is no natural ordering of multidimensional categories.

For exposition purposes, since in some applications values for (7) may lie far from unity[20], an alternative index stemming from a monotonic transformation of (7), and whose ordinal rankings are perfectly consistent with (7), is:

$$H\checkmark = \sqrt{H^2_{T,A}}. \tag{8}$$

2.1. Another representation of the index[21]

There is a different way of writing the index in (10) which renders the index more familiar to traditional inequality indices. Even though the index handles probabilities, as opposed to values of variables, it can be expressed as a proportion of the weighted sum of the squared coefficient of variations of the probabilities across types for a given outcome, α, where the weights are the average probabilities of being in outcome α, p^*_α. Define the probablity variance of outcome α as: $\sigma^2_\alpha = \sum_{t=1}^{T} w_t \left(p^t_\alpha - p^*_\alpha\right)^2$; and the coefficient of variation, $cv_\alpha = \frac{\sigma_\alpha}{p^*_\alpha}$. Then:

$$H^2_{T,A} = \frac{\sum_{\alpha=1}^{A} p^*_\alpha cv^2_\alpha}{\min\left(T-1, A-1\right)}. \tag{9}$$

2.2. Behaviour of the index

In this subsection the behaviour of the index is fleshed out further to illustrate the sensitivity of the index to association between types and outcomes. A related point to be illustrated here is that most changes in the distribution of outcomes across types (e.g. due to intra or inter-type transfers) have an *a priori* ambiguous

[20] Pearson's coefficient of contingency also has a similar empirical trait (Everitt, 1992, p. 54-5).
[21] I would like to thank James Foster for pointing to me this way of representing the index.

effect on the index. The nature of the effect depends on whether the change brings about an increase or a decrease in the degree of association between types and outcomes, i.e. the criterion by which inequality of opportunity is measured. In the context of probabilities, changes in the distribution come about by changes in the number of units (e.g. individuals) which fall into the cells of the contingency table, i.e. by migration of units from one cell to another. In the application of contingency tables to inequality of opportunity, units can only move across cells representing different outcome values within each type column but not across types because, by definition, people are assumed to be unable to change the very circumstances beyond their control which define the types they belong to. Consider the contingency table 1:

Table 1: Representation of distribution of outcomes across and within types with a contingency table

	Types				Row totals
	N_1^1	$\cdots$	$\cdots$	N_1^T	N_1
	$\vdots$	$\uparrow$	$\cdots$	$\vdots$	$\vdots$
Outcomes	$\vdots$	N_α^τ	$\cdots$	$\vdots$	N_α
	N_A^1	$\downarrow$	$\cdots$	N_A^T	N_A
Column totals	N^1	N^τ	$\cdots$	N^T	N

The minimum change that could occur in the table is that one observation from type τ migrates away from outcome row α. Such change, which in the case of discretized continuous variables could be due to a transfer, is related to a change in at least four variables: $N_\alpha^\tau, N_k^\tau, N_\alpha$ and N_k; and correspondingly in at least four probabilities: $p_\alpha^\tau, p_k^\tau, p_\alpha^*$ and p_k^*. When one unit moves away from outcome α toward outcome k, p_α^τ transfers to p_k^τ the amount of $\frac{1}{N^\tau}$, while p_α^* transfers $\frac{1}{N}$ to

p_k^*. To measure the impact of such a minimum change on the index, $H_{T,A}^2$, the following notation is considered by replacing (7) by the respective formulas of the numerator and denominator:

$$H_{T,A}^2 = \frac{\sum_{t=1}^{T}\sum_{\alpha=1}^{A} w^t \frac{\left(p_\alpha^t - p_\alpha^*\right)^2}{p_\alpha^*}}{\min(T-1, A-1)} \tag{10}$$

where $w_t = \frac{N^\tau}{N}$ is the population share of type t. In this illustration a unit belonging to type τ migrates from outcome state j to outcome state i. The new probabilities are decorated with a hat and the proposed migration implies that: $\widehat{p}_\alpha^t = p_\alpha^t \forall t \neq \tau; \alpha \in O \wedge \widehat{p}_\alpha^\tau = p_\alpha^\tau \forall \alpha \neq i, j \wedge \widehat{p}_i^\tau = p_i^\tau + \frac{1}{N^\tau} \wedge \widehat{p}_j^\tau = p_j^\tau - \frac{1}{N^\tau}$ $\wedge\ \widehat{p}_i^* = p_i^* + \frac{1}{N} \wedge \widehat{p}_j^* = p_j^* - \frac{1}{N}$. Let also $\Delta H_{T,A}^2 \equiv H_{T,A}^2\left(\widehat{p}_y^z\right) - H_{T,A}^2\left(p_y^z\right)$ and $\Delta mH \equiv \min\left\{T-1, A-1\right\}\Delta H_{T,A}^2$. Then:

$$\begin{aligned}\Delta mH \quad = \quad & \sum_{t=1}^{T-1} w^t \frac{\left(p_i^t - \widehat{p}_i^*\right)^2}{\widehat{p}_i^*} + w^\tau \frac{\left(\widehat{p}_i^\tau - \widehat{p}_i^*\right)^2}{\widehat{p}_i^*} + \sum_{t=1}^{T-1} w^t \frac{\left(p_j^t - \widehat{p}_j^*\right)^2}{\widehat{p}_j^*} \\ & + w^\tau \frac{\left(\widehat{p}_j^\tau - \widehat{p}_j^*\right)^2}{\widehat{p}_j^*} - \sum_{t=1}^{T} w^t \frac{\left(p_i^t - p_i^*\right)^2}{p_i^*} - \sum_{t=1}^{T} w^t \frac{\left(p_j^t - p_j^*\right)^2}{p_j^*},\end{aligned} \tag{11}$$

which yields:

$$\begin{aligned}\Delta mH \quad = \quad & \frac{1}{N^2\left(p_i^*+\frac{1}{N}\right)\left(p_j^*-\frac{1}{N}\right)}\left[\frac{1-w^\tau}{w^\tau}\right]\left(p_i^*+p_j^*\right) \\ & +\frac{2}{N\left(p_i^*+\frac{1}{N}\right)}\left(p_i^\tau-p_i^*\right) \\ & -\frac{2}{N\left(p_j^*-\frac{1}{N}\right)}\left(p_j^\tau-p_j^*\right) \\ & -\frac{1}{N\left(p_i^*+\frac{1}{N}\right)}\sum_{t=1}^{T}w^t\frac{\left(p_i^t-p_i^*\right)^2}{p_i^*} \\ & +\frac{1}{N\left(p_j^*-\frac{1}{N}\right)}\sum_{t=1}^{T}w^t\frac{\left(p_j^t-p_j^*\right)^2}{p_j^*}.\end{aligned} \tag{12}$$

As equation (12) shows, a minimum change of one migrating unit generates an *a priori* ambiguous effect on the index. Such a result is reasonable since the index is measuring inequality as increased association and a migration of one unit may or may not bring about more association between types and outcomes. Such migration may or may not bring about more similarity across the probabilities p_j^t and p_i^t $(\forall t \in T)$. Increasing similarity across probabilities from different types related to the same outcome cell, e.g. α, means reducing the degree of association betwen types and outcomes.

The behaviour of the index is further elucidated by looking at the following different situations in which migration can take place:

- When the probabilities across types are not identical in both departure and arrival states, j and i before and after the migration. In this situation the effect of the migration is ambiguous as shown by equation (12).
- When the probabilities across types are identical in the departure state, j.

In this case $p_j^t = p_j^* \forall t \in G$. Here there are three sub-situations:

1. The probabilities in the arrival state are not identical before and after the migration. In this case, as shown in equation (13) the effect is ambiguous. The migration away from j does increase the value of the index because originally $p_j^t = p_j^* \forall t \in G$, and that is captured by the first element of the right-hand side of (13). However the effect of the migration on the degree of similarity across probabilities in the arrival state, i, may or may not increase the overall degree of association. Hence the ambiguity.

$$\begin{aligned}\Delta mH &= \frac{1}{N^2\left(p_i^* + \frac{1}{N}\right)\left(p_j^* - \frac{1}{N}\right)}\left[\frac{1-w^\tau}{w^t}\right]\left(p_i^* + p_j^*\right) \\ &+ \frac{2}{N\left(p_i^* + \frac{1}{N}\right)}\left(p_i^\tau - p_i^*\right) \\ &- \frac{1}{N\left(p_i^* + \frac{1}{N}\right)}\sum_{t=1}^{T} w^t \frac{\left(p_i^t - p_i^*\right)^2}{p_i^*}\end{aligned} \tag{13}$$

2. The probabilities in the arrival states are not identical before the migration but are rendered identical afterwards. In this case, besides having $p_j^t = p_j^* \forall t \in G$, the following also holds: $p_i^l = p_i^m \forall l, m \neq \tau \rightarrow p_i^* = w^\tau p_i^\tau + (1 - w^\tau) p_i^l$. The impact on the index is again ambiguous because migration away from j increases association but migration toward i reduces it. The contribution to the change in the value of the index of these opposite impacts to the respective coefficients of variations of states j and i is mediated by the proportion of the total

population in every state as shown in equation (9). The change in the index due to migration in this situation is in equation (14). Notice that a necessary, but insufficient, condition for this migration to reduce the value of the index is: $p_i^\tau < p_i^l$. On the other hand, the complementary, opposite condition, $p_i^\tau \geq p_i^l$, is sufficient to ensure that this migration increases the value of the index.

$$\begin{aligned}\Delta mH &= \frac{1}{N^2\left(p_i^*+\frac{1}{N}\right)\left(p_j^*-\frac{1}{N}\right)}\left[\frac{1-w^\tau}{w^t}\right]\left(p_i^*+p_j^*\right) \\ &+\frac{2}{N\left(p_i^*+\frac{1}{N}\right)}\left(1-w^\tau\right)\left(p_i^\tau-p_i^l\right)\left[2-w^t\left(p_i^\tau-p_i^l\right)\right]\end{aligned} \tag{14}$$

3. The probabilities in the arrival state are identical before the migration. This case means $p_i^t = p_i^* \wedge p_j^t = p_j^* \forall t \in G$. This is a pre-migration situation of *partial equality of opportunity* in both departure and arrival states (i.e. outcome cells). As equation (15) shows, such migration breaks this partial equality and thus increases association.[22] Accordingly the index reacts by increasing its value. By the same token, any minimum migration which restores outcome-pairwise partial equality of opportunity reduces the value of the index thereby reflecting less association between types and outcomes. Notice also that equation (15) holds the proof to the fact that any migration disturbing an initial situation of perfect, total equality of opportunity raises the value of the index. Conversely, any migration that restores or generates partial or

[22] The result assumes $p_j^* > \frac{1}{N} \wedge 0 \leq p_i^* \leq 1 - \frac{1}{N}$.

total equality of opportunity reduces the value of the index.

$$\Delta mH = \frac{1}{N^2\left(p_i^* + \frac{1}{N}\right)\left(p_j^* - \frac{1}{N}\right)}\left[\frac{1 - w^\tau}{w^t}\right]\left(p_i^* + p_j^*\right) > 0 \qquad (15)$$

- When the probabilities across types are such that there is perfect association between types and outcomes in the departure and arrival states, j and i respectively, before migration. This situation depends on whether $T < A, T > A$ or $T = A$. Hence the different sub-situations can be classified into the folllowing three cases:

Case 1 *When $T < A$ there are two possible sub-situations:*

1. A migration of a member of type τ from j to i that leaves perfect association intact. For this migration to be possible it has to be the case that in the initial situation type τ is exclusively associated both with outcomes j and i, which implies $w^\tau p_j^\tau = p_j^* \wedge w^\tau p_i^\tau = p_i^* \wedge p_i^t, p_j^t = 0 \forall t \neq \tau$. It is easy to show that plugging these conditions into equation (12) yields $\Delta mH = 0$. In words, a migration within states exclusively associated to the type to which the migrating unit belongs leaves the index unchanged. Notice that this result also holds when perfect, exclusive association is present across the whole contingency table, i.e. when the index attains its highest value.

2. A migration of a member of type τ from j to i that breaks perfect association. This type of migration requires that type τ is exclusively associated with outcome j but not with i.[23] This situation involves a

[23] The opposite, that the perfect association is with the final state of migration and not with

type k which initially, before the migration, is exclusively associated with outcome i. This type of migration implies: $w^{\tau}p_j^{\tau} = p_j^* \wedge w^k p_i^k = p_i^* \wedge p_i^t = 0 \forall t \neq k$ (including $p_i^{\tau} = 0$). Under these conditions the value of the index decreases (equation (16)). This result also holds if the whole contigency table exhibits perfect association, i.e. when the index attains its maximum value before migration. Equation (16) proves that a pairwise breakup of perfect assocation in the table reduces the value of the index. Likewise an inverse migration that restores or generates these exclusive associations is reflected in the index by an increase in its value.

$$\Delta mH = -\frac{p_i^*}{N\left(p_i^* + \frac{1}{N}\right)}\left[\frac{w^k + w^{\tau}}{w^{\tau}w^k}\right] < 0 \tag{16}$$

Case 2 *When $T > A$ there are again two possible sub-situations:*

1. A migration of a member of type τ from j to i that leaves perfect association intact. When $T > A$ every outcome is associated with a different subset of types. Therefore for perfect association to remain intact after such migration it has to be the case that the association with type τ is given up by outcome j in favour of i. Therefore $p_j^{\tau} = \frac{1}{N^{\tau}} \wedge \widehat{p}_j^{\tau} = 0 \wedge p_i^{\tau} = 0 \wedge \widehat{p}_i^{\tau} = \frac{1}{N^{\tau}}$. This migration leaves the index also unchanged, including when there is perfect association across the whole contingency table. For the proof see Appendix 3.

2. A migration of a member of type τ from j to i that breaks perfect

the original one, is impossible by definition of the example in which the idea is to break initial perfect association.

association. Unlike the first sub-situation of this second case, in this sub-situation $p_j^\tau > \frac{1}{N^\tau} \wedge \widehat{p}_j^\tau > 0 \wedge p_i^\tau = 0 \wedge \widehat{p}_i^\tau = \frac{1}{N^\tau}$. Hence the migration from j to i effectively breaks the exclusive association between outcomes j and i and the vector of types by generating an association of both outcomes with the same type τ (under $T > A$). This migration reduces the value of the index. For the proof see Appendix 3.

Case 3 *When $T = A$ type τ is exclusively associated with departure state j but, by implication of $T = A$, it is not associated with arrival state i. Therefore a migration from j to i is characterized by the same conditions as the second sub-situation of the case when $T < A$,*[24] *and has the same effect: a reduction in the value of the index.*[25]

2.3. Relationship between the index and other concepts and indices of inequality of opportunity

2.3.1. The conception of Roemer (1998)

Roemer suggests at least two ways to quantify inequality of opportunity. In his book *Equality of Opportunity* (1998) he dedicates more attention to defining an equal-opportunity policy than to characterizing situations of equal opportunity or lack thereof. However, he mentions what an equal-opportunity society looks like to him. In his framework, advantages (i.e. outcomes) are determined by efforts and circumstances (for which individuals should not be held accountable since they

[24] That is: $w^\tau p_j^\tau = p_j^* \wedge w^k p_i^k = p_i^* \wedge p_i^t = 0 \forall t \neq k$ $\left(\text{including } p_i^\tau = 0\right)$.

[25] Notice that this analysis assumes that $p_j^* > \frac{1}{N}$. Otherwise the migration under initial perfect association renders state j without observations/individuals and ΔmH becomes indeterminate. This indeterminacy is reasonable since the continengcy table changes shape (it contracts) when this migration happens and initially $p_j^* = \frac{1}{N}$.

lie beyond their control). Circumstances affect both the remuneration to effort but also the distribution of individuals' effort within every type. Roemer then subscribes to the view that: "... if we could somehow disembody individuals from their circumstances, then the *distribution* of the propensity to exert effort would be the same in every type." (Roemer, 1998, p. 15). With such a statement and with Roemer's assumption of a monotonic relationship between effort and advantages (p. 10), Roemer's statement can be made to imply that should circumstances not affect effort or its remuneration then the distribution of advantages/outcomes should be the same across different types. Roemer also states that in an equal-opportunity environment individuals belonging to different types but exerting the same *degree* of effort, i.e. effort relative to their own type, should receive the same remuneration. Under the assumption of monotonicity between effort and remuneration this requirement can be interpreted literally as implying that the values of all percentiles of the well-being distributions are the same across types, i.e. that the type-conditoned distributions are all identical. In his article "Economic development as opportunity equalization" (2006) Roemer provides a more explicit phrasing for a literalist approach to his definition, which for the sake of reference I call the *percentile* approach. He states that: "...Equality of opportunity for the acquisiton of advantage [...] has been achieved if, at every level of effort, the leves of advantage *across types* are the same. [...] if the cumulative distribution functions of advantages across types are identical" (p. 8). The dissimilarity index is faithful to this notion in that it hits its minimum value, representing perfect equality of opportunity, *if and only if* conditional distributions of well-being are identical. The tranches approach of Checci and Peragine (2005) is also in tune with this literalist interpretation. The reason is that identical remunerations across types for the same degree of effort (measured by the percentile in the group-specific cumulative distribution of the outcome under a monotonicity assumption) are attained if and only if the density functions across types are identical, which in turn implies identical cumulative distributions. Equation (17) states it formally where $F_A^{-1}(p)$ is the value of the outcome for group A at percentile p, f is a density function and the outcome can take values in the interval $[x_{\min}, x_{\max}]$.

$$F_A^{-1}(p) = F_B^{-1}(p)\,\forall p \in [0,1] \leftrightarrow f^A(x) = f^B(x)\,\forall x \in [x_{\min}, x_{\max}] \qquad (17)$$

However Roemer (2006) also proposes to measure inequality of opportunity in terms of between-group inequality of mean advantage or outcome where the latter is understood as a standard that summarizes the distribution function (p. 8). This latter approach is the one followed by the types approach of Checci and Peragine (2005), Ferreira and Gignoux (2008) as well as Elbers *et al.* (2008) and Lanjouw and Rao (2008). Both the types approach on one hand and the dissimilarity index and the tranches approach, on the other hand, agree in identifying societies where conditional distributions are identical as perfectly opportunity-equal because they both start from Roemer's general notion of equal opportunity. However Roemer leaves open the options for the categorization of and comparison of societies in terms of their relative degree of inequality of opportunity. That fact helps to explain the disagreements, that I show below, on the classification of opportunity-unequal societies between the percentile and the tranches approach on one hand and the types approach on the other. The types approach interprets such inequality in terms of decomposed between-group inequality of the outcome(s), whereas the dissimilarity index takes it to mean the degree of association between types and outcomes, i.e. the dissimilarity between the multinomial distributions and the tranches approach interprets it as the part of total inequality that is not due to differences in mean advantage attainment across percentiles (which Checci and Peragine cluster into tranches).

2.3.2. The conception of Lefranc et al. (2008)

The dissimilarity index highlights a tension between the definition and operationalization of equality of opportunity relying on second-order stochastic dominance and proposed by Lefranc *et al.* (2008) and that of Roemer. The definition of Lefranc *et al.* (2008) declares equality of opportunity whenever there is no second-order dominance across the outcome distributions corresponding to different circumstances. Even though such is a reasonable definition from the point of view of a hypothetical individual outsider who has to choose between different types; its ranking of societies may significantly mismatch with that of an index, like this paper's, based on the notion that equality of opportunity is only achieved when circumstances determining types are irrelevant in the distribution of outcomes within any type; a notion which is based on Roemer's *assumption of charity.*[26] For instance, take societies A and B whose joint distributions of circumstances and outcomes are described below. Considering two circumstances (on the columns) and five possible values for the outcomes (on the rows), society B would exhibit equality of opportunity according to the dissimilarity index. Moreover in the case of the latter society the dissimilarity index would rank A as having per-

[26] "[...] within any type, that distribution [of the propensity to exert effort] would be the same, were we able to factor out the (different) circumstances which define different types." (Roemer, 1998, p.15). Since Roemer assumes a monotonic relationship between effort and outcomes/advantages then the notion of equality of opportunity based on the homogeneity of the outcomes' distributions across types can be derived from this *assumption of charity*.

fect inequality of opportunity. However, if there are even intervals between the values of the outcomes (on the rows) one can easily show that in both societies, A and B, neither circumstance second-order stochastically dominates the other one. Therefore, according to Lefranc *et al.*'s (2008) criterion, equality of opportunity would have to be asserted in both societies. The definition of Lefranc *et al.* agrees with Roemer's every time the latter declares a society to be opportunity-equal, whereas as illustrated by the example, Roemer's does not follow Lefranc *et al.*'s every time the latter declares equality of opportunity.

$$A = \begin{bmatrix} 0 & 0.25 \\ 0.5 & 0 \\ 0 & 0.5 \\ 0.5 & 0 \\ 0 & 0.25 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$$

Consistently with their definition, Lefranc *et al.* also propose a Gini index of inequality of opportunity, which is defined as:

$$GO = \frac{1}{2\mu} \sum_{i=1}^{T} \sum_{j=1}^{T} w^i w^j \left| \mu_i (1 - G_i) - \mu_j (1 - G_j) \right|$$

Where μ is the mean of the welfare measure (e.g. income) over the whole population, μ_i is the respective mean for type i and G_i is the Gini coefficient for type i. Again when comparing GO against Roemer's definition it is easy to verify that whenever a society is opportunity-equal according to Roemer's criterion, and according to the dissimilarity index, GO is zero, thereby measuring equality of opportunity according to Roemer's definition. However the reverse is not true.

GO may be zero even when distributions of the advantage/outcome are not equal. For instance, let $T = 2$, it is not difficult to find values for the two types' means and Gini coefficients such that $\mu_1 > \mu_2$ and $G_1 > G_2$, which imply dissimilarity in the two distributions of the advantage and inequality of opportunity according to Roemer[27], and $GO = 0$, which implies equality of opportunity according to Lefranc *et al*.

2.3.3. The indices based on the types approach (Checci and Peragine, 2005; Ferreira and Gignoux, 2008)

By contrast to the conception of Lefranc *et al.*, both the dissimilarity index and the indices based on the types approach (e.g. Checci and Peragine, Roemer, Ferreira and Gignoux) agree that only societies in which distributions of outcomes are identical across types should be classified as opportunity-equal; because both define equality of opportunity based on Roemer's *assumption of charity*. However there is disagreement on the criterion to declare distributions to be identical. The types approach effectively compares across types/groups a standard that represents the conditional distributions. Empirical applications of this approach have typically associated increasing inequality of opportunity with increasing between-groups/types inequality in mean outcomes. On the other hand, the dissimilarity index more generally associates increasing inequality of opportunity with increasing distance between the multinomial outcome distributions of the different types. When comparing conditioned multinomial distributions, maximum distance between them can be said to hold when maximum association between the condi-

[27] At least if and when such heterogeneity is due to differential remuneration to individuals belonging to different types but exerting the same degree of relative effort within their respective types.

tioning factors (e.g. the types, t, on which the distributions are conditioned) and the partition categories (e.g. the outcome intervals, α) is present. Such discrepancy in the understanding of increasing *inequality* of opportunity leads to at least two instances of disagreement in rankings. The first one is that whenever there is no decomposed between-group inequality[28] the types approach declares equality of opportunity even though the dissimilarity index might not do so because the absence of between-group inequality, as understood by the types approach, can occur even when the multinomial distributions are not homogeneous. In such situation people belonging to different types still face different lotteries and so differential opportunities, which, in the notion upon which the dissimilarity index is based, implies inequality of opportunity. An illustration of this discrepancy is provided again by the comparison of societies A and B (above). Assume further that in both societies rows are associated progressively with the following values for the outcome: 1, 2, 3, 4, 5. Then it is easy to check that in both A and B decomposed between-group inequality is nil.[29] The types approach would rank both societies as exhibiting perfect equality of opportunity. Hence the decomposed between-group inequality index may even declare equality of opportunity when individuals from different types receive different remunerations although they may be exerting the same *degree* of effort,[30] which contradicts Roemer's definition of an equal-opportunity environment in his book, or at least the literalist interpretation

[28] Their notion of between-group inequality is based on path-independent decompositions of traditional decomposable indices into a between-group and a within-group component.

[29] Both Checci and Peragine (2005) and Ferreira and Gignoux (2008) use the mean log deviation index which allows for a decomposition of inequality into between and within-groups without residual, using the groups' arithmetic means, and in a path-independent way. For path-independent decomposability see Foster and Schneyerov (2000).

[30] Roemer defines *degree* of effort, differently from an absolute level of effort, as the relative effort exerted by an individual compared to the effort of other individuals of the same type and measured by the individual's percentile in his/her type's specific effort distribution.

discussed above.[31] The dissimilarity index, on the other hand, accounting for the differential lotteries faced by the two types in society A, would rank the latter as having perfect inequality of opportunity, whereas it would still consider B as exhibiting equality of opportunity.

The second instance happens when two societies are compared and both exhibit maximum association but one has between-group inequality, according to path-independent decomposition, coupled with no within-group inequality and the other one has some within-group inequality and may or may not have path-idependent, decomposed between-group inequality. In such a case the relative version of the types approach, i.e. the between-group component divided by total inequality, ranks the first society as being perfectly opportunity-unequal and the second one as being less opportunity-unequal; whereas the dissimilarity index ranks both as being perfectly opportunity-unequal on the merit of both exhibiting maximum association between types and outcome categories. An illustration of this case is on the comparison between societies C and D. The types approach would rank D has having perfect inequality of opportunity because it does not exhibit any decomposed within-group inequality. By the same token, they would rank C has having less inequality of opportunity than D since the former has some decomposed within-group inequality. By contrast, the dissimilarity index ranks both as being perfectly opportunity-unequal on the grounds of the maximum association between types and outcome levels found in both of them.

[31] An example of this situation is in Appendix 1.

$$C = \begin{bmatrix} 0 & 0 \\ 0.5 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0.5 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$$

These two sources of discrepancy remain even if the indices are adjusted according to the suggestion of Elbers *et al.* (2008) whereby the between-group index is not divided by total inequality (stemming from the implicit assumption that each individual is a group in his/herself) but instead by the maximum value that the between-group component could attain for a given outcome distribution and a given sample share of the different combinations of groups that might be possible to construct. This suggestion in turn is based on Shorrocks and Wan (2005) who show that the maximum between-group inequality possible is attained when the conditional distributions do not overlap.

2.3.4. *The indices based on the tranches approach (Checci and Peragine, 2005)*

Checci and Peragine (2005) propose an alternative measure of inequality of opportunity also based on inequality indices decomposable into between and within groups. Like this paper's dissimilarity index, they follow a literalist interpretation of Roemer's notion that people exerting the same degree of effort, measured by their percentile position in their respective type's effort distribution, should receive an equal amount of the advantage/outcome. Then, assuming monotonicity between (unobservable) effort and observable advantages, they measure inequality of opportunity as inequality in the outcome/advantage between individuals belonging to different types (groups) but exerting the same degree of effort, captured by belonging to the same percentile tranch, thence the name tranches approach. In the implementation, Checci and Peragine (2005) divide the percentile space into tranches and then, as a first stage, they replace the values of the outcome for every individual with that of the mean of the outcome corresponding to the specific group-tranch cell to which individuals belong.[32] With these values total inequality can be computed. Beyond issues of practicality in empirical implementation, Checci and Peragine (2005) claim that this first-stage transformation removes all inequality which is not explained by either circumstances or effort (measured by the tranches). The second stage involves calculating the mean outcome value for every tranche. Finally using similar path-independent decomposition techniques, they use the mean log deviation index, inequality of opportunity is calculated as the residual from subtracting between-tranch inequality to total inequality (calculated over the distribution smoothed in the first stage). A relative measure of inequality of opportunity based on the tranches approach can also be constructed by dividing the within-tranch inequality measure by total inequality.

Of the indices mentioned in this paper, those based on the tranches approach agree with the dissimilarity index in declaring equality of opportunity if and only if conditional distributions of well-being are identical. In the case of the tranches-based mean log deviation index this is easily proved by noticing that:

[32]For instance, if there are just two groups (e.g men and women) and three tranches (e.g. bottom third, middle third, top third), then a woman being at the median level of the women's distribution will have its own value replaced with that of the mean for all women belonging to the middle third of their distribution.

$$I\left(X_W^S\right)=\frac{1}{N}\sum_{p=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{\frac{N_i}{m}}\ln\left[\frac{\mu}{\left(\mu_{ip}^{j}\frac{\mu}{\mu_p}\right)}\right]=\frac{1}{N}\sum_{p=1}^{m}\sum_{i=1}^{n}\left(\frac{N_i}{m}\right)\ln\left[\frac{\mu_p}{\mu_{ip}}\right] \tag{18}$$

where $I\left(X_W^S\right)$ is the inequality index (in this case the mean log deviation) applied to the smoothed distribution of the outcome, denoted by X_W^S. As described above, the smoothing replaces every observation's outcome value with the mean corresponding to its type, i, and its tranch, p. Therefore $\mu_{ip}^{j}=\mu_{ip}^{k}=\mu_{ip}\forall j,k$, where μ_{ip} is the mean of observations belonging to type i and tranch p. Following Checci and Peragine's notation, N is the total number of observations, m is the number of tranches, n is the number of types and N_i is the number of observations in type i. Clearly, $I\left(X_W^S\right)=0\leftrightarrow\mu_{ip}=\mu_p\ \forall i=\{1,...,n\},p=\{1,...,m\}$. That is, the index is equal to zero if and only if the conditional distributions are identical.

However, as in the case with the comparison between the dissimilarity index and the types-approach index, the relative version of the tranches-approach index does not rank all distributions characterized by perfect association as being perfectly opportunity unequal.[33] The reason is that the relative version of the tranches-approach index has total inequality as its maximum, which is a reasonable normalization if the objective is to decompose inequality into effort-led and circumstance-led components. However perfect association between types and outcome sets, which is the standard of perfect inequality of opportunity for the dissimilarity index, is possible with very different distributions of the outcome (reflecting effort through the monotoncity assumption) within each conditioning

[33] The absolute versions do not have an upper bound representing perfect inequality of opportunity like the dissimilarity index or the Gini of inequality of opportunity.

type/group. Therefore two societies exhibiting perfect association are ranked by the dissimilarity index as having equally perfect inequality of opportunity whereas the relative version of the tranches-based index may rank them differently if they exhibit different levels of within-type inequality. For instance, the dissimilarity index ranks E and F as being perfectly opportunity-unequal. By contrast the relative tranches-based index ranks F as having more relative inequality of opportunity than E.

$$E = \begin{bmatrix} 0.5 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \\ 0 & 0.5 \end{bmatrix} \qquad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

2.3.5. The human opportunity index (Barros et al., 2009)

Let p_1^* be the average accomplishment related to an outcome denoted by 1 (e.g. school completion) or the average access to a basic service (e.g. water) defined over a dichotomous variable. And define p_1^t similarly but with respect to a specific group of society denoted by t.[34] Then the human opportunity index (HOI) by Barros et al. 2009 is defined as:

$$HOI = p_t^* (1 - D),$$

where D is a dissimilarity index based on a statistic belonging to the above specified general class, and defined as:

[34] These groups can be defined over sets of Roemer's circumstances, for instance. But D can also be applied over individuals as an index of dispersion in which, for instance, the weights would be given by one over the sample size.

$$D \equiv X^1_{T,1}/2N = \frac{1}{2}\sum_{t=1}^{T} w^t \frac{|p^t_1 - p^*_1|}{p^*_1}.$$

The *HOI* considers together the average attainment of the outcome of interest (e.g. the access to the service) and the relative inequality of such attainment across social groups. It follows a tradition of defining welfare indices which account both for the average and the dispersion of a welfare outcome, started by Atkinson (1970), Sen (1976) and followed, among others, by Yitzhaki (1979), often using the Gini coefficient as a measure of dispersion. Lefranc *et al.* (2008) also define an index of inequality of opportunity, as a Gini coefficient of the areas under the generalized Lorenz curves of the social groups, which in turn depend on the groups' average values of the outcome and the group-specific Gini coefficients.

Since the dissimilarity index of the *HOI* and this paper's index are both based on the same general class of statistics it is worth comparing them by highlighting their similarities and differences. The relevant version of the index for the comparison is $H^2_{T,2}$ since it deals with dichotomous outcomes.

Firstly it is worth comparing the whole *HOI* with the dissimilarity index. They both compare the distribution of an outcome of interest across groups against the group average. In the case of the *HOI* that comparison is embedded in D. However, the dissimilarity index only considers the average level of the outcome as a comparison pivot but not as valuable in itself. In other words, it solely occupies itself with quantifying inequality of opportunity regardless of the average welfare of society (as measured by the outcome).

Secondly, the dissimilarity index of the *HOI* works with dichotomous variables whereas the multinomial dissimilarity index works with multinomial distributions, which include dichotomous variables as a special case. Therefore the multinomial dissimilarity index can also be applied to quantify inequality in access to services, although as I show below it may not rank societies the same way as D does and there may be a conceptual case to apply one or the other index depending on the nature of the outcomes under analysis.

Whereas D is not useful to compare multinomial distributions because its maximum value (necessary for normalization) would not depend just on the population size. Instead it would depend *ad hoc* on the groups' weights.[35] For those reasons D would not be applicable to deal with joint distributions of welfare (i.e. to analyse multidimensional inequality of opportunity).

At this point it is also worth mentioning an important trait of the normalization of an index like D: its maximum value depends on the sample size and is equal to $\frac{N-1}{N}$. This maximum is achieved when only one individual (belonging to a group of size $w^i = 1/N$ has a value of $p_1^i = 1$ and the rest of the population have a value of $p_1^t = 0 \forall t \neq i$. Only when $N \rightarrow \infty$ it happens that $D \rightarrow 1$ under the described situation.[36] Hence multiplying D by $\frac{N}{N-1}$ would correct for this problem, ensuring that it reaches its maximum value whenever there is maximum dissimilarity as measured by D. However the index would no longer exhibit population invariance in all situations which are different from both maximum and minimum dissimilarity. This trade-off between a population invariance axiom and a normalization axiom is absent from $H_{T,2}^2$ (and $H_{T,A}^2$ in general). Both approaches can be defended on theoretical grounds.

[35] This is an inconvenience of assessing dissimilarity between multinomial distributions using absolute deviations as opposed to square deviations.

[36] The traditional Gini coefficient of inequality exhibits a similar trait. It only attains its maximum value in the limit when N tends to infinity and only one individual within the population owns the resource.

A key similarity between D and $H^2_{T,2}$ is that they both declare equality of opportunity whenever $p^t_1 = p^*_1 \forall t$. In the case of $H^2_{T,2}$, $p^t_2 = 1 - p^t_2$, hence $p^t_1 = p^*_1 \forall t \leftrightarrow p^t_2 = p^*_2 \forall t$. However for more general situations different from perfect equality of opportunity, D and $H^2_{T,2}$ may not necessarily rank societies consistently among themselves. This detail is shown more formally in Appendix 4. The two indices also differ in declaring maximum associaton or perfect inequality of opportunity. Maximum dissimilarity as measured by D implies maximum dissimilarity by $H^2_{T,2}$ but the reverse is not true. The reason is that for D there is only one situation in which maximum dissimilarity holds, and that is when one individual has a value of 1 and the rest of the population has a value of 0 for p_1. Whereas in the case of $H^2_{T,2}$ all situations in which there is a subset $G_1 \subset G$ for which $p^t_1 = 1 \forall t \in G_1 \wedge p^t_1 = 0 \forall t \notin G_1$ and, by implication, $p^t_2 = 0 \forall t \in G_1 \wedge p^t_2 = 1 \forall t \notin G_1$, are deemed to exhibit perfect maximum association. The unique maximum of D is a special case of the group of maximum association situations considered by $H^2_{T,2}$.

These discrepancies (on general rankings and on declaring maximum association) emphasize the relevance of the need for conceptual criteria to decide which index to use when faced with dichotomous outcomes. When the dichotomy is about the achievement of a valuable situation (e.g. access to a service) against the lack of it, I advocate using D since it is reasonable to consider that there is more inequality of opportunity in a situation in which one individual has full certainty of achieving the outcome and all the others have zero probability of attaining it, than in a situation in which more than one individual has this full certainty and the rest of the population has zero probability. In empirical applications a relatively large sample should diminish the importance of the above mentioned trade-off between population invariance and normalization axioms.

When the dichotomy is about two options whose ranking is not immediately obvious in terms of valuability or, for instance, when one simply wants to assess the dissimilarity of type-specific distributions across these two categories, and one is content to declare perfect inequality when one of the options is exclusively associated with one subset of types and the other option is exclusively associated with the remaining subset, independently of the different forms that such associations can take, then I advocate using $H^2_{T,2}$. An example would be to consider distributions of occupation with two categories. For instance agriculture versus non-agricultural occupations (e.g. as in Bossuroy *et al.*, 2007). The idea being that it is not immediately obvious that one of the occupations is better in some meaningful sense than the other, but such ranking depends on society-specific contexts. Another example would be to divide between blue-collar and white-collar workers, where, as is known (e.g. see Giddens, 2006), not all workers qualified as white-collars in developed countries may be, for instance, financially better-off that all workers qualified as blue-collars.[37]

3. Empirical application

As an empirical application of the dissimilarity index based on the test of multinomial distributions; I look at changes in inequality of opportunity in Peru from a cohort of adults aged 45 years old or older to a younger cohort of adults aged 22 to 45 years old. I focus on two discrete outcomes: levels of education

[37] Giddens provides the example of highly skilled artisans compared to low-rank clerks.

attained and quality of education attained, proxied by type of school attended.[38] The adult population is divided into 8 types which result from combining three circumstances, each measured with two categories: gender, father's education and mother's education.[39]

3.1. Data

The data come from the Peruvian National Household Survey, ENAHO 2001 which sampled 16,515 households. There are 7 possible multidimensional outcomes stemming from the following values for the outcome variables:

1. Years of education:
   - No education (=1)
   - Some primary, incomplete or complete, but no secondary (=2)
   - Some secondary, incomplete or complete, but no tertiary (=3)
   - Some tertiary education (=4)
2. Quality of education
   - No education (=1)
   - Attended public school (=2)
   - Attended private school (=3)

[38] Private schools are known to be of better quality in Peru. There is, for instance, evidence of an earnings premium from having attended private schools. See Calonico and Nopo (2007).

[39] Ideally a finer division would have been desirable but I am clustering educational categories due to sample size concerns.

The combination of these two variables yields 7 categories because the "no education" entries only interact with each other. The 8 types ensue from combining the following three variables:

1. Gender
   - Male
   - Female
2. Father's education
   - Up to complete primary
   - More than complete primary
3. Mother's education
   - Up to complete primary
   - More than complete primary

The respective sample sizes are in the table in Appendix 2.

3.2. Results

The value of the index for the old cohort is 0.044698 (or 0.21142 using the transformation in (8)) while for the younger cohort it is 0.039811 (or 0.199526 using the same transformation). Using the percentile bootstrap technique (e.g. see Mooney and Duval, 1993), the 99% confidence interval for the old cohort's estimate is $[0.040367, 0.05012]$. For the young cohort the respective 99% confidence interval is $[0.037614, 0.042935]$. Since both point estimates fall outside the confidence interval of the other sample's estimate the evidence is in favour of a statistically significant reduction in inequality of opportunity from the older to the younger cohort.

4. Concluding remarks

This paper proposes the use of a dissimilarity index for the analysis of inequality of opportunity. A similar index has been proposed for the measurement of heterogeneity across transition matrices (Yalonetzky, 2009). In the opportunity literature the index is an additional quantitative tool influenced by Roemer's (1998) conception. The dissimilarity index measures inequality of opportunity in proportion to the degree to which sets of circumstances associate with sets of outcomes. A higher degree of association, of which correlation is one type, is related to higher heterogeneity and dissimilarity of distributions conditioned on type-belonging; and, in turn, higher inequality of opportunity. Both this paper's index, the one used by the *HOI* and the decomposition approaches, either by between-types or between-tranches, do not judge which group (as defined by sets of circumstances) is the most advantaged. For that an analysis of risk, return and stochastic dominance is required, as the one performed by Lefranc *et al.* (2008). By contrast, the equality of opportunity criterion put forth by Lefranc *et al.* (2008) is not consistent with Roemer's and hence neither with the decomposition approaches and this paper's. The dissimilarity index agrees with the types approach (and with Roemer's different concepts) in classifying societies when equality of opportunity is present, according to the literalist view of identical conditional distributions. However the dissimilarity index and the types approach, as it has been put forward by Cecchi and Peragine (2005), disagree in the ranking of opportunity-unequal societies since the latter relates inequality of opportunity to path-independent, decomposed between-group inequality based on equalization of a distributional standard of the variable (e.g. arithmetic means); whereas the dissimilarity index relates inequality of opportunity to association between types and outcome/advantage values. If inequality of opportunity is understood in terms of association, or distance between multinomial distributions, then one can find distributions of outcomes which are different without there being any between-group inequality according to the traditional path-independent decomposition.

On the other hand the dissimilarity index agrees with the tranches approach, also proposed by Cecchi and Peragine (2005) based on Roemer (1998), in declaring inequality of opportunity *if and only if* conditional distributions of well-being are identical. However the index and the relative version of the tranche approach, whereby within-tranche inequality is divided by total inequality, may disagree on the ranking of distributions characterized by perfect association. The index ranks all these distributions as being equally perfectly opportunity unequal. By contrast, in the tranches approach different between-tranch inequality may yield different values of the relative indicator for different societies, all of them characterized by perfect association. The dissimilarity index declares perfect inequality of opportunity if and only if there is perfect maximum association between types and wellbeing outcomes.

The dissimilarity index belongs to a family of inequality-of-opportunity indices, along with those of the types and the tranches approaches, which do not account for average attainments in societies. Among those which do, the most prominent are the Gini of Lefranc *et al.* (2008) and the Human Opportunity Index, *HOI*. This paper shows that a major difference between the dissimilarity index and the Gini of opportunities is that the latter declares equality of opportunity not only when conditional distributions are perfectly identical. Regarding the *HOI*, this paper shows that both its dissimilarity index and this paper's are based on statistics belonging to the same general class. These two indices also agree on declaring perfect equality if and only if conditional distributions are equal. They also agree on declaring perfect inequality whenever one individual has full certainty of attaining a valuable outcome while the rest of his society has zero chance of attaining it. However they may disagree on rankings of societies in intermediate situations of inequality with imperfect association. They also disagree on ranking societies with different forms of perfect association between types and outcomes.

An advantage of the dissimilarity index is that it is suited to cope with multiple advantages or outcomes. But as any index based on multinomial distributions, it is most suitable for discrete variables. In the case of continuous variables the implementation of the index requires robustness checks in order to assess whether and how the discretization of continuous variables affects the index's rankings.[40]

In the empirical application to educational opportunity in Peru the index proves useful in providing evidence of a statistically significant reduction in multidimensional educational inequality of opportunity among a younger cohort of adults (22-45 year olds vis-a-vis 46 or more years old). In this application several types are defined by combining gender with parental education. These types do not exhaust all the groups of people which can be defined in the Peruvian sample according to circumstances beyond the adults' control. For instance parental oc-

[40] There are heuristic recommendations as to how to partition a continuous space into a discrete multinomial one. See for instance Stuart and Ord (1991; Vol 2).

cupation or ethnicity could have been considered with richer and larger samples. Therefore the types thus defined conflate groups and the estimation of inequality of opportunity can be interpreted as yielding a lower bound, as argued by Ferreira and Gignoux (2008). Alternatively the index plainly measures inequality of opportunity based on the specific definition of types used, which has to be borne in mind when performing comparisons.

This paper has sought to emphasize the value of the dissimilarity approach to measuring inequality of opportunity. Further work ought to focus also on finding quantitative descriptive tools to measure multidimensional inequality of opportunity of continuous variables (and combinations of discrete and continuous variables). Exploring the tranches approach may prove uself for the purpose of this research.

5. Appendix 1

Example of a discrepancy between an index of inequality of opportunity based on the type approach and a literalist interpretation of Roemer's notion.[41]

The following example shows how an index of inequality of opportunity based on a perfect and path-independent decomposition of inequality in between-group and within-group elements may rank a society as being perfectly opportunity-equal even though its individuals are not being remunerated equally when they

[41] Checchi and Peragine (2005) also provide an example showing that a between-group approach may generate different rankings of distributions from those of an index based on a between-tranche approach. The latter approach agrees with Roemer (1998) in detecting inequality of opportunity whenever individuals exerting the same degree of effort but belonging to different types enjoy different values of the outcome/advantage.

exert identical degrees of effort, which is stressed by Roemer as a key principle behind his conception of equality of opportunity (e.g. Roemer, 1998, chapter 3).

Imagine a society with two types: 1 and 2. Assume a monotonic and linear relationship between effort and outcome/advantages.[42] The distribution of the advantage, x, for type 1 is given by a fragmented uniform distribution:

$$f^1(x) = \left\{ \begin{array}{ccc} 0 & , & x < a \vee b < x < c \vee x > d \\ \frac{p}{b-a} & , & a \leq x \leq b \\ \frac{1-p}{d-c} & , & c \leq x \leq d \end{array} \right\}$$

Where $a < b < c < d$ are values that x can take and $0 < p < 1$. By contrast, the distribution of advantage x among type 2 individuals is given by a non-fragmented uniform distribution:

$$f^2(x) = \left\{ \begin{array}{ccc} 0 & , & x < b \vee x > c \\ \frac{1}{c-b} & , & b \leq x \leq c \end{array} \right\}$$

Then if the means of x for both types are identical then the between-group inequality of opportunity indicator would rank this society as being opportunity-equal. That is, if:

$$\int_a^b \frac{p}{b-a} x dx + \int_c^d \frac{1-p}{d-c} x dx = \int_b^c \frac{1}{c-b} x dx$$

Several combinations of a, b, c, d and p could satisfy the equality of means. For instance: $a = 1$, $b = 2$, $c = 3$, $d = 4$ and $p = 0.5$. Considering these numbers, and assuming a monotonic and linear relationship between effort and the outcome, the remuneration of the type 1 individual exerting effort at the 75^{th} percentile

[42] Roemer assumes a monotonic relationship between advantages and effort.

of his/her respective distribution earns $c + (d - c)/4 = 3.25$ whereas the type 2 individual exerting the same degree of effort would earn $b + 3(d - c)/4 = 2.75$. Therefore such a situation would not be regarded as one of equality of opportunity by Roemer's definition even though the between-group inequality index would find no between-group inequality.

6. Appendix 2

Table A2.1. Sample sizes for the empirical application

Type	Younger cohort: 22-45 years old	Older cohort: 46 years old or older
Male, both parents had up to complete primary	6744	4771
Male, only father more than complete primary	817	268
Male, only mother more than complete primary	199	37
Male, both parents had more than complete primary	1402	338
Female, both parents had up to complete primary	7324	4511
Female, only father more than complete primary	967	250
Female, only mother more than complete primary	193	31
Female, both parents had more than complete primary	1419	312

7. Appendix 3

Proofs that migration from a state j to a state i either decreases or leaves the value of the index unchanged when there is initial perfect association and $T > A$

The following are the proofs for the statements made regarding the situation of migration when the probabilities across types are such that there is perfect association between types and outcomes in the departure and arrival states, j and i respectively, before migration, and $T > A$. In a situation of perfect association when $T > A$ state j is exclusively associated with a subset $G_j \subset G$. Therefore $p_j^t = 1 \forall t \in G_j \wedge p_j^t = 0 \forall t \notin G_j$. Similarly, before the migration, state i is exclusively associated with a subset $G_i \subset G$ such that $p_i^t = 1 \forall t \in G_i \wedge p_i^t = 0 \forall t \notin G_i$. Notice further that perfect association means that $G_j \cap G_i = \{\varnothing\}$ and $G_j \cup G_i \subset G$ (i.e. unless $T = 2$ there are other states, $\alpha \neq j, i$, which may or may not be perfectly associated with the rest of types in G).

Now define $w^j = \sum_{t=1}^{T} w^t I\left(t \in G_j\right)$. That is, w^j is the sum of the weights of all the types which are perfectly associated with j (I is an indicator function that takes the value of 1 whenever the expression in parenthesis is true, and the value of zero otherwise). Similarly define $w^t = \sum_{t=1}^{T} w^t I\left(t \in G_i\right)$. Hence before migration $p_j^* = w^j \wedge \; p_i^* = w^i$. In this context, suppose that a migration of individuals belonging to type τ ($\tau \in G_j$) takes place from j to i. Such migration renders $\widehat{p}_j^\tau = 1 - \delta \wedge \widehat{p}_i^\tau = \delta \wedge \widehat{p}_j^* = w^j - \delta w^\tau \wedge \widehat{p}_i^* = w^i + \delta w^\tau$, that is, after the migration.[43] Following equation (11) the change in the index is:

$$\begin{aligned}
\Delta mH \; = \; & \left(w^j - w^\tau\right)\left[\frac{\left(1-\left(w^j-\delta w^\tau\right)\right)^2}{w^j-\delta w^\tau} + \frac{\left(0-\left(wi+\delta w^\tau\right)\right)^2}{w^i+\delta w^\tau} - \frac{\left(1-w^j\right)^2}{w^j} - \frac{\left(0-w^i\right)^2}{w^i}\right] \qquad (19)\\
& + w^i\left[\frac{\left(0-\left(w^j-\delta w^\tau\right)\right)^2}{w^j-\delta w^\tau} + \frac{\left(1-\left(wi+\delta w^\tau\right)\right)^2}{w^i+\delta w^\tau} - \frac{\left(0-w^j\right)^2}{w^j} - \frac{\left(1-w^i\right)^2}{w^i}\right]\\
& + w^\tau\left[\frac{\left(1-\delta-\left(w^j-\delta w^\tau\right)\right)^2}{w^j-\delta w^\tau} + \frac{\left(\delta-\left(wi+\delta w^\tau\right)\right)^2}{w^i+\delta w^\tau} - \frac{\left(1-w^j\right)^2}{w^j} - \frac{\left(0-w^i\right)^2}{w^i}\right].
\end{aligned}$$

[43] Of course, $0 \leq \delta \leq 1$.

After some manipulation equation (19) is reduced to the following expression:

$$\Delta mH = -2 + \frac{w^j - \delta w^\tau (2-\delta)}{w^j - \delta w^\tau} + \frac{w^j + \delta^2 w^\tau}{w^j + \delta w^\tau} \leq 0. \tag{20}$$

Hence any such migration can not increase the value of the index. If $\delta = 1$ then $\Delta mH = 0$, i.e. perfect association involving states j and i with $G_j \cup G_i$ is kept intact but type τ has changed groups from G_j to G_i. Otherwise if the migration breaks perfect association, i.e. if $0 < \delta < 1$ then $\Delta mH < 0$.[44]

8. Appendix 4

A more formal illustration of potential discrepancies in rankings of societies between the multinomial dissimilarity index, $H^2_{T,2}$, and D

Imagine a migration of a percentage δ of individuals from type τ from state 1 to state 2, and those two states are the only ones under consideration, i.e. $p_1+p_2 = 1$.. Hence $\widehat{p}_1^\tau = p_1^\tau + \delta$ and $\widehat{p}_1^* = p_1^* + w^\tau\delta$. The dissimilarity index of the *HOI*, D, changes the following way:

$$\begin{aligned} \Delta mD \quad &= \quad w^\tau \left[\frac{|p_1^\tau + \delta - p_1^* - w^\tau\delta|}{p_1^* + w^\tau\delta} - \frac{|p_1^\tau - p_1^*|}{p_1^*} \right] \\ &\quad + \sum_{t\neq\tau}^{T} w^t \left[\frac{|p_1^t - p_1^* - w^\tau\delta|}{p_1^* + w^\tau\delta} - \frac{|p_1^t - p_1^*|}{p_1^*} \right], \end{aligned} \tag{21}$$

where $\Delta mD \equiv 2\left[D\left(\widehat{p}_1^t\right) - D\left(p_1^t\right)\right]$, i.e. ΔmD measures the change in D due to the migration. Considering that $p_1 + p_2 = 1$, the change in $H^2_{T,2}$ due to the

[44] The case $\delta = 0$ als renders $\Delta mH = 0$ but it trivially means that no migration took place.

migration is (after some manipulation):

$$\begin{aligned}\Delta mH &= w^{\tau}\left[\frac{\left(p_1^{\tau}+\delta-p_1^{*}-w^{\tau}\delta\right)^2}{\left(p_1^{*}+w^{\tau}\delta\right)\left(1-p_1^{*}-w^{\tau}\delta\right)}-\frac{\left(p_1^{\tau}-p_1^{*}\right)^2}{\left(p_1^{*}\right)\left(1-p_1^{*}\right)}\right] \\ &\quad+\sum_{t\neq\tau}^{T} w^{t}\left[\frac{\left(p_1^{\tau}-p_1^{*}-w^{\tau}\delta\right)^2}{\left(p_1^{*}+w^{\tau}\delta\right)\left(1-p_1^{*}-w^{\tau}\delta\right)}-\frac{\left(p_1^{t}-p_1^{*}\right)^2}{\left(p_1^{*}\right)\left(1-p_1^{*}\right)}\right].\end{aligned} \tag{22}$$

Now define:

$$D_t^F \equiv \frac{\left|\widehat{p}_1^{t}-\widehat{p}_1^{*}\right|}{\widehat{p}_1^{*}} \text{ and } D_t^I \equiv \frac{\left|p_1^{t}-p_1^{*}\right|}{p_1^{*}}.$$

Then expressions (21) and (22) can be rewritten in terms of D_t^F and D_t^I as:

$$\Delta mD = \sum_{t=1}^{T} w^{t}\left[D_t^F - D_t^I\right], \tag{23}$$

$$\Delta mH = \sum_{t=1}^{T} w^{t}\left[\left(D_t^F\right)^2\frac{\widehat{p}_1^{*}}{1-\widehat{p}_1^{*}} - \left(D_t^I\right)^2\frac{p_1^{*}}{1-p_1^{*}}\right]. \tag{24}$$

Notice the differences between (23) and (24): in (24) D_t^F and D_t^I are squared, and they are each multiplied by different weights, respectively $\frac{\widehat{p}_1^{*}}{1-\widehat{p}_1^{*}}$ and $\frac{p_1^{*}}{1-p_1^{*}}$. Therefore there is no guarantee that, for instance, both ΔmD and ΔmH have the same sign in every occasion. Both the squaring and the different weighting can make them disagree in the nature of the change due to the same migration.

References

Anderson, T.W., and L. Goodman (1957), 'Statistical inference about Markov chains', *The Annals of Mathematical Statistics,* 28: 89-110.

Anderson, G. (1996), 'Nonparametric tests of stochastic dominance in income distributions', *Econometrica*, 64: 1183-1193.

Asadullah, N. (2006), 'Intergenerational economic mobility in rural Bangladesh', Manuscript.

Atkinson, A. (1970), 'On the measurement of economic inequality', *Journal of Economic Theory*, 2: 244-63.

Breton, M. le, Michelangeli, A. and E. Peluso (2008), 'Wage discrimination measurement: in defense of a simple but informative statistical tool', Manuscript.

Bossuroy T., D. Cogneau and V. Hiller (2007), 'Intergenerational mobility in five African countries', Manuscript.

Barros, R.P. de, Ferreira, J. M. Vega and J. S. Chanduvi (2009), *Measuring inequality of opportunity in Latin America and the Caribbean*, Washington D.C.: The World Bank.

Butler, R. and J. McDonald (1987), 'Interdistributional income inequality', *Journal of Business and Economic Statistics"*, 5: 13-7.

Calonico, S. and H. Nopo (2007), 'Where did you go to school? Private-public differences in schooling trajectories and their role on earnings', *Well-being and Social Policy*, 3: 25-46.

Checchi, D. and V. Peragine (2005), 'Regional disparities and inequality of opportunity: the case of Italy', Discussion Paper Series, 1874, IZA.

Cramer, H. (1946), *Mathematical methods for Statistics*, Princeton: Princeton University Press.

Dagum, C.(1980), 'Inequality measures between income distributions with applications', *Econometrica*, 48: 1791-803.

Dagum, C. (1987), 'Measuring the economic affluence between populations of income receivers', *Journal of Business and Economic Statistics*, 5: 5-8.

Elbers, C., P. Lanjouw, J. Mistiaen and B. Ozler (2008), 'Reinterpreting between-group inequality', *Journal of Economic Inequality*, 6: 231-45.

Ebert, U. (1984), 'Measures of distance between income distributions', *Journal of Economic Theory*, 32: 266-74.

Everitt, B.S. (1992), *The analysis of contingency tables, Monographs on Statistics and Applied Probability*, 45, Second Edition: Chapman & Hall.

Fleurbaey, M. (2001), 'Egalitarian opportunities', *Law and Philosophy*, 20: 499-530.

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

Ferreira, F. and J. Gignoux (2008), 'The measurement of inequality of opportunity: theory and application to Latin America', Manuscript.

Foster, J. and A. Sheyerov (2000), 'Path independent inequality measures', *Journal of Economic Theory*, 91: 199-222.

Gastwirth, J. (1975), 'Statistical measures of earnings differentials', *The American Statistician*, 29: 32-5.

Giddens, A. (2006), *Sociology*, Fifth Edition: Polity Press.

Handcock, M. and M. Morris (1999), *Relative distribution methods in the social sciences*. New York: Springer.

Hogg, R. and E. Tanis (1997), *Probability and statistical inference*, New Jersey: Prentice Hall.

INEI (2001), *Encuesta Nacional de Hogares (ENAHO) IV trimestre: condiciones de vida y pobreza*, Lima: INEI.

Lanjouw, P. and V. Rao (2008), 'Revisiting between-group inequality measurement: an application to the dynamics of caste inequality in two Indian villages', Manuscript.

Lefranc, A., N. Pistolesi and A. Trannoy (2008), 'Inequality of opportunities vs. inequality of outcomes: Are Western societies all alike?', *The Review of Income and Wealth*, 54: 513-46.

Mooney, C. and R. Duval (1993), *Bootstrapping. A nonparametric approach to statistical inference*, Series: Quantitative applications in the social sciences, 95: A Sage University Paper.

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

Roemer, J. (2006), 'Economic development as opportunity equalization', Cowles Foundation Discussion Paper, 1583.

Sen, A. (1976), 'Real national income', *The Review of Economic studies*, 43: 19-39.

Shorrocks, A. and G. Wan (2005), 'Spatial decomposition of inequality', *Journal of Economic Geography*, 5: 59-81.

Stuart, A. and K. Ord (1991), *Kendall's Advanced Theory of Statistics, Volume 2*, Fifth Edition: Edward Arnold.

Yalonetzky, G. (2008), *Essays on Economic Mobility*, DPhil Dissertation, University of Oxford.

Yalonetzky, G. (2009), 'Comparing economic mobility with an heterogeneity index: an application to education in Peru', OPHI Working Paper, Oxford Poverty & Huma Development Initiative.

Yitzhaki, S. (1979), 'Relation deprivation and the Gini coefficient', *Quarterly Journal of Economics*, 93: 321-4.