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

OPHI WORKING PAPER NO. 39

A Dissimilarity Index of Multidimensional Inequality of Opportunity

Gaston Yalonetzky *

November 2010

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 with two indices of dissimilarity across distributions. The indices are based on a traditional homogeneity test of multinomial distributions and are similar to the square coe¢ cient of variation (Reardon and Firebaugh, 2002). Their properties are studied, as well as their usefulness and limitations in applications when both circumstances and advantages/outcomes are multidimensional. An empirical application measures changes in inequality of opportunity from an old to a young cohort in Peru. The importance of assessing the sensitivity of the results to group definitions and group proportions is highlighted.

*Oxford Poverty & Human Development Initiative, University of Oxford. E-mail address: gaston.yalonetzky@qeh.ox.ac.uk.

This study has been prepared within the OPHI theme on multidimensional 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 978-1-907194-23-8

A Dissimilarity Index of Multidimensional Inequality of Opportunity

Gaston Yalonetzky*

November 7, 2010

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 with two indices of dissimilarity across distributions. The indices are based on a traditional homogeneity test of multinomial distributions and are similar to the square coefficient of variation (Reardon and Firebaugh, 2002). Their properties are studied, as well as their usefulness and limitations in applications when both circumstances and advantages/outcomes are multidimensional. An empirical application measures changes in inequality of opportunity from an old to a young cohort in Peru. The importance of assessing the sensitivity of the results to group definitions and group proportions is highlighted.

1 Introduction

The concern for inequality of opportunity has long earned its place in the Social sciences and Political Philosophy. Following Roemer (1998)'s influential conceptualization, 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 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. Barros et al. (2009) compile studies of inequality of opportunity in Latin America, including the advocacy of a Human Opportunity Index (the HOI). Similarly Elbers et al. (2008) have applied new refinements on the decomposition of inequality indices to tracking changes in between-group inequality in several countries.

The intergenerational economic mobility literature has a longer history of development of quantitative tools. However, since in that 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 inequality of opportunity with multiple circumstances, let alone multiple outcomes. On the other hand the

*OPHI, University of Oxford. E-mail address: gaston.yalonetzky@qeh.ox.ac.uk. I would like to thank three anonymous referees for very helpful and constructive comments, and 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 fruitful discussions. I am also grateful to Lorenzo Oimas for providing me with the database.

toolkit of the literature on segregation indices, and contingency tables, contains indices that deal explicitly with multiple circumstances. Moreover these indices and those derived from them, are worth considering for applications involving multiple outcomes, as I propose in this paper with two indices.

This paper contributes to the quantitative analysis of inequality of opportunity by suggesting the use of dissimilarity indices related to the statistic of a traditional test of homogeneity of multinomial distributions, which is in turn Pearson's chi-square statistic, but expressed in terms of probabilities.[1] When applied to inequality of opportunity, these dissimilarity indices have the advantage of being readily applicable to comparisons of multidimensional distributions of outcomes, which is an appealing trait in the burgeoning multidimensional welfare measurement literature. Another interesting trait is that the dissimilarity indices attain their 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 indices are measuring, is 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 indices attain their maximum value *if and only if* there is complete association between the groups in which societies are partitioned and the welfare outcomes in consideration. The indices are most suitable for ordered, discrete variables; whereas for continuous variables they requires prior discretization.

In an empirical application I study changes in inequality of educational opportunity in Peru from older to younger cohorts of adults. Considering as educational wellbeing outcomes, levels of educational attainment and quality of education, measured by type of school attended (public versus private versus none), I compute inequality of opportunities for these two outcomes separately and for their joint distribution. The ensuing cross-cohort trends depend on the definition of types and on whether the index used is sensitive to the proportions of adults in every group. For instance, I find evidence of lower inequality of opportunity among the younger cohorts vis-a-vis the oldest ones, for the most refined partitions of adults into groups, for all three estimations. The groups of people (or types in Roemer's vocabulary) are defined by combinations of gender, paternal and maternal education levels.

In the next section the dissimilarity indices are introduced. Then their behaviour is investigated, considering properties discussed in the segregation and inequality of opportunity literatures. A discussion comparing the indices' orderings to those of existing quantitative tools for univariate distributions follows. Then the reach and limitations of using dissimilarity indices to address multiple outcomes are also discussed. Finally, an empirical application to Peru is presented, followed by concluding remarks.

2 The dissimilarity index of multidimensional inequality of opportunity

In Roemer (1998)'s influential conceptualization, advantages (i.e. outcomes) are determined by efforts and circumstances. Combinations of the latter determine people's types. Nobody should be held accountable for belonging to a specific type since it is beyond the individual's control. Hence as a source of inequality, circumstances are deemed *morally irrelevant* in the inequality of

[1] The Pearson chi-square statistic is used to test the null hypothesis of lack of association in a contingency table. It has support on the absolute frequencies of the table. See , e.g. Everitt (1992).

oppportunity literature Hild and Voorhoeve (2004). They should not affect the advantage either directly or indirectly through effort or random shocks. Therefore, in many definitions of equality of opportunity, the distributions of the advantages should be identical across social groups defined by sets of circumstances.

The dissimilarity indices discussed in this paper relate 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 can be further characterized 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 would exhibit the same distributions of effort should their defining circumstances be factored out Roemer (1998, p. 16). 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 different level of dedication.[2] This definition is equivalent to the strong criterion of Lefranc et al. (2009). Another equivalent way to characterize it 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, Romer's effort, or *morally relevant* factors), and not of circumstances.[3] 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 indices presented in this paper are closely related to indices from the rich segregation literature,[4] particularly to the squared coefficient of variation. These indices work on probability space, thereby being specially suited to deal with multivariate distributions of discrete variables. Several of them also handle multiple groups, hence they are applicable to multiple circumstances. To define this paper's indices, assume that societies are partitioned into a set of individuals' types, defined by a combination of values taken by a vector of circumstances; i.e. factors over which the individual does not exert control, e.g. parental education, ethnicity or gender. Generally, 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. 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 vector of types, $G = \{1, 2, ..., T\}$, has then $T = \prod_{i=1}^{z} g_i$ elements. All individuals having the same set of circumstances are said to be of the same type. The absolute frequency of people in a society belonging to type t ($t \in G$) is N^t, and the total population sample is N. Similarly outcomes or advantages can be considered, generally, in a multidimensional way. All possible combinations of outcomes (e.g. health status with education achievement, earnings, etc.) are in vector $O = \{1, 2, ..., A\}$. Assuming 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:

[2] 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).

[3] Ferreira and Gignoux (2008) also elaborate on this point.

[4] For instance, the work of James and Taeuber (1985), White (1986), Hutchens (2001) and Reardon and Firebaugh (2002) Including references there in. I would like to thank an anonymous referee for pointing to me references from this rich literature on segregation measurement.

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

O has $A = \prod_{j=1}^{b} m_j$ elements representing a combination of outcomes, each one partitioned in the mentioned m_j elements or categories.[5] 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. Therefore $\sum_{t=1}^{T}\sum_{\alpha=1}^{A} N_k^t = N$.

The indices of dissimilarity advocated in this paper belong to a general class of statistics which measure the degree of dissimilarity between distributions as the 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. the types) and the row variable may represent the outcome variable. Following Reardon and Firebaugh (2002), these indices can be derived from a family of indices that compute the average of a dissimilarity function, $f : [0,1] \times ... \times [0,1] \to \mathbb{R}_+$, that is sensitive to the divergence between p_k^t and p_α^*; 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 :

$$p_\alpha^* = \sum_{t=1}^{T} p_\alpha^t \frac{N^t}{\sum_{t=1}^{T} N^t} = \frac{\sum_{t=1}^{T} N_\alpha^t}{\sum_{t=1}^{T} N^t}. \tag{1}$$

The family of indices computes the average f over p_α^* $(\forall\alpha)$ and the relative size of each type in the population: $w^t = \frac{N^t}{N}$.Hence this approach of *segregation as disproportionality in group proportions* (Reardon and Firebaugh, 2002, p. 39-40)proposes the following family:

$$W_{T,A} \in W^* \mid W_{T,A} \equiv \sum_{\alpha=1}^{A} p_\alpha^* \sum_{t=1}^{T} w_t f\left(p_\alpha^t; p_\alpha^*\right) \;\; \forall A, T \in \mathbb{N}_{++}, ^6 \tag{2}$$

When: $f\left(p_\alpha^t; p_\alpha^*\right) = \left|\frac{p_\alpha^t - p_\alpha^*}{p_\alpha^*}\right|^\beta \forall\beta \in \mathbb{N}_{++}$ the following subfamily, X^*, of (2) emerges:

$$X_{T,A}^\beta \in X^* \mid X_{T,A}^\beta \equiv \sum_{\alpha=1}^{A} p_\alpha^* \sum_{t=1}^{T} w^t \left|\frac{p_\alpha^t - p_\alpha^*}{p_\alpha^*}\right|^\beta \;\; \forall\beta, A, T \in \mathbb{N}_{++}, \tag{3}$$

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 (1)), rendering the statistic in (3) closer to zero. The dissimilarity indices proposed in this paper are 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 $\beta = 2$:

[5] 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).

[6] In Reardon and Firebaugh (2002) the respective function for the family is: $\sum_{m=1}^{M} \pi_m \sum_{j=1}^{J} \frac{t_j}{T} f\left(\frac{\pi_{jm}}{\pi_m}\right)$, where M is the number of groups in the population, T is the size of the population, j is an organizational unit (e.g. school), π_m is the percentage of the population belonging to group m and π_{jm} is the probability of belonging to group m conditioned on being in organization unit j. Notice that, by contrast, p_α^t stands for the probability of attaining a state of wellbeing α conditioned on belonging to group/ type t.

$$NX^2_{T,A} = \sum_{t=1}^{T}\sum_{\alpha=1}^{A} N^t \frac{\left(p^t_\alpha - p^*_\alpha\right)^2}{p^*_\alpha}. \tag{4}$$

The null hypothesis of the test is that the T distributions are homogenous, i.e. identical in a statistical sense. Besides being related to a standard test of multinomial distributions, another advantage justifying the choice of $X^2_{T,A}$ (among many other options from class X^*) for 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 groups (e.g. the number of types) and the number of states (e.g. the values that the outcome variable takes). The maximum value is easily found by noticing that the statistic of the homogeneity test of multinomial distributions is Pearson's goodness-of-fit statistic:

$$NX^2_{T,A} = \sum_{t=1}^{T}\sum_{\alpha=1}^{A} N^t \frac{\left(p^t_\alpha - p^*_\alpha\right)^2}{p^*_\alpha} = \sum_{t=1}^{T}\sum_{\alpha=1}^{A} \frac{\left(N^t_\alpha - \frac{N^t N_\alpha}{N}\right)^2}{\frac{N^t N_\alpha}{N}}. \tag{5}$$

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^t_k is 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 is given by the expression $\frac{N^t N_\alpha}{N}$ (see, e.g. Everitt, 1992). Therefore, (5) can be expressed as:

$$NX^2_{T,A} = \sum_{t=1}^{T}\sum_{\alpha=1}^{A} \frac{\left(OB^t_\alpha - E^t_\alpha\right)^2}{E^t_\alpha}, \tag{6}$$

where the OB stands for observed and the E for expected frequency. Cramer (1946) showed that the maximum for an expression like (6) is $NX^2_{T,A,\max} = \min(T-1, A-1)N$, and that is, precisely, the maximum for the statistic (4). Thus dividing (4) by $NX^2_{T,A,\max}$, yields the first dissimilarity index:

$$H^2_{T,A} = \frac{NX^2_{T,A}}{NX^2_{T,A,\max}}. \tag{7}$$

The index is advocated for applications in which $A > 1$. It is very similar, but not identical, to the square coefficient of variation by Reardon and Firebaugh (2002).[7]

The second dissimilarity index is a special case of the first index ,(7), in which: $w_t = \frac{1}{T}$; that is, all type weights are set to be equal:

$$\overline{H}^2_{T,A} = \frac{1}{\min(T-1, A-1)T} \sum_{t=1}^{T}\sum_{\alpha=1}^{A} \frac{\left(p^t_\alpha - \tilde{p_\alpha}\right)^2}{\tilde{p_\alpha}}, \tag{8}$$

where $\tilde{p_\alpha} = \frac{1}{T}\sum_{t=1}^{T} p^t_\alpha$. Compared to (7), the index in (8) has additional population invariance properties,[8] with respect to that in (7), rendering it useful whenever it is desired that only changes

[7] The squared coefficient of variation by Reardon and Firebaugh (2002), when expressed in terms of types and outcomes, is equal to: $C = \frac{X^2_{T,A}}{A-1}$. Hence $H^2_{T,A} = C \leftrightarrow A \leq T$ and $1 \geq H^2_{T,A} > C \leftrightarrow A > T$.

[8] These are discussed in the next section.

in the conditional distribution probabilities, and not changes in the proportions of types in the population, can impact on the degree of inequality. Effectively, the index (8) takes a "representative agent" approach to compare the different conditional distributions.

3 Behaviour of the indices

In this section the behaviour of the indices is elucidated by explaining the situations under which the indices attain their maximum value of between-type inequality; by analyzing their fulfillment of other properties discussed in the segregation literature; and by deriving some results that clarify the effects of migrations of individuals, from one outcome state to another, on the value of the indices.

3.1 The indices' maximum inequality

The indices (7) and (8) measure the degree of between-type inequality in terms of the degree of association between types and outcomes. In the context of a contingency table having two variables with two states each, Kendall and Stuart (1973) propose two notions of maximal association:

- *Complete association*: It occurs when all individuals who have an attribute A also have an attribute B, even though not everyone having attribute B may have attribute A. This definition can be extended to row and column variables in a table, each with several possible values. In the latter case, complete association is meant to occur when all individuals having value A of row variable also have value B of column variable although not all those having value B of column variable may have value A of row variable.

- *Absolute association*: All individuals who have attribute A also have attribute B and *viceversa*. In the case of row and column variables each having several values, absolute association means that all individuals having value A of row variable also have value B of column variable and viceversa.

A first property of the two dissimilarity indices (7) and (8) is that they attain their maximum value (of 1) in the following situations:[9]

- When $T < A$ (more outcome categories or states than types) and every type $t \in G$ is associated with a subgroup of outcomes $O_t \subset O$, such that the T subgroups of outcomes do not overlap, i.e. $\cup_{t=1}^{T} O_t = O$.In other words, if an individual's outcome, β, belongs in O_t ($\beta \in O_t$) then the individual is of type t (but not every member of type t has an outcome β) This is a case of complete association.

- When $T > A$ (more types than outcomes) and every outcome $\alpha \in O$ is associated with a subgroup of types $G_\alpha \subset O$, such that the A subgroups of types do not overlap, i.e. $\cup_{\alpha=1}^{A} G_\alpha = G$.In other words, if an individual's type, t, belongs in G_α ($t \in G_\alpha$) then the individual's outcome is α (but not everyone with outcome α is of type t). This is another case of complete association.

[9] The proofs that the indices attain their maximum values under these three situations are in the Appendix 8.

- When $T = A$, the maximum value of the indices is attained under absolute association. That is, every type $t \in G$ is exclusively associated with one outcome $\alpha \in O$.Being of type t implies having outcome α and viceversa.

The lack of overlap required in order to attain $H^2_{T,A} = 1$, when $T < A$, can take different forms. Therefore, in general, indices measuring inequality on probability space do not distinguish between different forms of non-overlap when they reach their extreme values of complete (or absolute) association.[10]

3.2 Properties fulfilled by the indices

The second property fulfilled by the two indices is that they attain their minimum value of 0 if and only if the conditional distributions of outcomes are identical across types. The general property is the following:

Condition 1 *Perfect Equality (PE): Any index of dissimilarity obtained using the approach of segregation as disproportionality in group proportions*[11] *attains its minimum value if and only if* $p^1_\alpha = p^2_\alpha = \ldots = p^T_\alpha \ \forall \alpha = 1, \ldots, A$.

The proof for indices (7) and (8) is in the Appendix 8. Now if the type-conditioned distributions of outcomes are put together to form a table where types define the columns and outcomes define the rows, then the following condition can be proposed:

Condition 2 *Symmetry in Outcomes (SO): A permutation of outcomes (rows in the table) does not alter the value of inequality*

Condition *(SO)* is fulfilled by the two indices, (7) and (8). Such fulfillment reflects a potential limitation of the indices: They are insensitive to the relative desirability of certain outcomes vis-a-vis others.[12] If the table is further expressed as a contigency table depending on absolute frequencies the following condition is also fulfilled by the two indices:

Condition 3 *Symmetry in Types (ST): A permutation of types (columns in the table) does not alter tha value of inequality*

Because the indices are evaluated on probability space they also fulfill a property of monotonic transformation invariance:

Condition 4 *Monotonic Transformation Invariance (MTI): Inequality is not affected by any common monotonic transformations to the values of the outcome variables and to the boundaries of the partitions of outcomes that determine the probability distribution.*

[10] By contrast, approaches like that of Lefranc et al. (2008), explicitly look at second-order dominance. Therefore their Gini of Inequality of Opportunity (GIO), which works on value space, is sensitive to the form of non-overlap when the latter is present. The types approach (Checchi and Peragine, 2005; Ferreira and Gignoux, 2008) is also sensitive to the form of the non-overlap because it affects the conditional mean values.

[11] (Reardon and Firebaugh, 2002, p. 39-40). These authors call this condition the *disproportionality axiom for segregation* (p. 39).

[12] This is a general limitation of existing inequality indices with support on probability space (as opposed to variables' values space), including the indices of the segregation literature. New univariate indices that overcome this limitation working with cumulative probabilities, can be found in Reardon (2009) and Silber and Yalonetzky (2010).

Another important property fulfilled by both indices is population invariance, i.e. they are insensitive to identical replications of each individual:

Condition 5 *Population Invariance (PI): If every individual is multiplied by a factor* $\lambda > 0$*, then inequality is unaffected.*

The proof is straightforward. Let $p^t_\alpha(\lambda)$ be the probability of attaining state α conditioned on being of type t and given a replication of λ. Then $p^t_\alpha(\lambda) = \frac{\lambda N^t_\alpha}{\lambda N} = p^t_\alpha$. Likewise, using a similar definition for $w_t(\lambda)$, it turns out that: $w_t(\lambda) = \frac{\lambda N^t}{\lambda N} = w^t$. Therefore both indices' values do not change. A similar property is fulfilled by index (8) but not by index (7). Reardon and Firebaugh (2002) call it *composition invariance* and it states that inequality should be unchanged when the replication of individuals is type-specific:

Condition 6 *Type Population Invariance (TPI): If every individual belonging to type t is multiplied by a factor* $\lambda_t > 0\,(\forall t \in [1, T])$*, then inequality is unaffected.*

Indices like (8) fufill *(TPI)* because they are independent of the type weights. Therefore they are "margin-free", taking a "representative agent" approach to the measurement of between-group inequality. There are good reasons to choose indices that fulfill *(TPI),* and to choose indices that do not fulfill *(TPI).*[13]

The literature on segregation discusses an important property called *organizational equivalence.* According to this property the measure of segregation should not change when an organizational unit (a type in the context of inequality of opportunities), is divided into k units, each with the same probability distributions as the original unit; or when k units with the same probability distributions merge into a single unit (e.g. type) (Reardon and Firebaugh, 2002, p. 38). This two-sided property can be redefined in the following encompassing way:

Condition 7 *Type Replication Invariance (TRI): If every type t, and its corresponding conditional distribution of outcomes, is multiplied by a factor* $\lambda_t > 0\,(\forall t \in [1, T])$*, then inequality is unaffected.*[14]

The indices (7) and (8) fulfill *(TRI)* under special settings. In the case of (7), let $H^2_{T,A}(\lambda_t) = \frac{1}{\min(S-1, A-1)} \sum_{t=1}^{S} \frac{w_t}{\lambda_t} \sum_{\alpha=1}^{A} \frac{\left(p^t_\alpha - p^*_\alpha\right)^2}{p^*_\alpha}$, where $S = \sum_{t=1}^{T} \lambda_t$. It is easy to show that: for $H^2_{T,A}(\lambda_t) = H^2_{T,A}$, when $\exists k \mid \lambda_k \neq 1$,[15] it is necessary (and sufficient) that $A \leq T$ and $A \leq S$. A sufficient (but not necessary) condition for $H^2_{T,A}(\lambda_t) = H^2_{T,A}$, when $\exists k \mid \lambda_k \neq 1$ is that $T = S$. In the case of (8) the same results follow after w_t is replaced with $\frac{1}{T}$. Alternatively if either $A > T$ or $A > S$, then *(TRI)* is fulfilled if and only if $T = S$.[16]

Finally, the indices fulfill an important condition related to their reaction to some migrations of individuals between outcomes:

[13] Reardon and Firebaugh (2002, p. 38) point to references of a debate in the segregation literature regarding the desirability of the "margin free" property.

[14] If $\lambda_t \in \mathbb{N}_+/\{1\}$, then type t is being subdivided into λ_t subtypes, each with equal conditional distributions of the outcome to the original type. Alternatively if $\lambda_t = 1/\gamma_t \mid \gamma_t \in \mathbb{N}_+/\{1\}$, then several types with identical conditional distributions of the outcome are being conflated into one single type.

[15] $\lambda_t = 1 \forall t \rightarrow H^2_{T,A}(\lambda_t) = H^2_{T,A}$

[16] The squared coefficient of variation by Reardon and Firebaugh (2002) fulfills *(TRI)* because it is normalized by the equivalent of $A - 1$, as opposed to $\min\{T-1, A-1\}$. Therefore the comparison between this coefficient of variation and the indices (7) and (8), highlight a trade-off between fulfillment of (TRI) and reaching a maximum value of 1 when $A > T$ and there is complete association.

Condition 8 *Sensitivity to migrations that break or restore partial pairwise equality of opportunity (SMPEP): A migration of an individual from outcome state j to outcome state i increases inequality if states j and i were characterized by partial pairwise equality of opportunity before the migration; and, conversely, it reduces inequality if states j and i get to be characterized by partial pairwise equality of opportunity (PPEP) after the migration.*

In condition 8, partial pairwise equality of opportunity (PPEP) is defined as follows: $\exists\,(\alpha, \beta) \in [1, ..., A] \mid p^1_\alpha = p^2_\alpha ... = p^T_\alpha = p^*_\alpha \wedge p^1_\beta = p^2_\beta ... = p^T_\beta = p^*_\beta$. The indices (7) and (8) fulfill *(SMEP)*, i.e. they react by increasing their value to reflect higher association (and inequality) when within-type migrations from one state to another break PPEP and they react by decreasing their value when a similar migration restores PPEP in the departure and arrival state. The proof is in the Appendix 8.

3.3 General reaction of the indices to migrations of individuals between outcome states[17]

In this subsection the behaviour of the index is fleshed out further by analyzing its sensitivity to association between types and outcomes. It turns out that most changes in the distribution of outcomes across types (e.g. due to intra or inter-type transfers) have an *a priori* ambiguous effect on the index. Notable exceptions are migrations related to condition 8, and migrations away from (or toward to) a situation of complete (or absolute) association in the pairs of departure and arrival states. Generally, the nature of the effect depends on whether the migration 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 by the indices. In the context of probabilities, distributional changes 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 to inequality of opportunity, units can only move across cells representing different outcome values within each type column but not across types, because people are, by definition, unable to change own circumstances beyond their control. Consider the contingency table 1:[18]

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

				Types	Row totals
	N^1_1	$\cdots$	$\cdots$	N^T_1	N_1
	$\vdots$	$\uparrow$	$\cdots$	$\vdots$	$\vdots$
Outcomes	$\vdots$	N^τ_α	$\cdots$	$\vdots$	N_α
	N^1_A	$\downarrow$	$\cdots$	N^T_A	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 j, ending up in outcome row i. Such change is related to a change in at least four variables: N^τ_j, N^τ_i, N_j and N_i; and correspondingly in at least four probabilities: $p^\tau_j, p^\tau_i, p^*_j$ and p^*_i. When δN^τ units move away from outcome j toward outcome i, p^τ_j transfers

[17] The analysis in this subsection considers index (7), but the same results ensue for index (8).

[18] White (1986) pioneered the application of contingency tables in the segregation literature.

to p_i^τ the amount of δ,[19] while p_j^* transfers $w^\tau \delta$ to p_i^*. To measure the impact of such a change on the indices, units of value δ belonging to type τ migrate from outcome state j to outcome state i. The new probabilities are decorated with a hat and the proposed migration implies that: $\widehat{p}_\alpha^t = p_\alpha^t \forall t \neq \tau; \alpha \in O$; $\widehat{p}_\alpha^\tau = p_\alpha^\tau \forall \alpha \neq i, j$; $\widehat{p}_i^\tau = p_i^\tau + \delta$; $\widehat{p}_j^\tau = p_j^\tau - \delta$; $\widehat{p}_i^* = p_i^* + w^\tau \delta$; $\widehat{p}_j^* = p_j^* - w^\tau \delta$. 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 $\Delta mH \equiv \min\{T-1, A-1\}\, \Delta H_{T,A}^2$. Then:

$$\begin{aligned} \Delta mH \quad = \quad & \sum_{t=1}^{T-1} w^t \frac{\left(p_i^t - \widehat{p}_i^*\right)^2}{\widehat{p}_i^*} + w^\tau \frac{\left(\widehat{p}_i^\tau - \widehat{p}_i^*\right)^2}{\widehat{p}_i^*} + \sum_{t=1}^{T-1} w^t \frac{\left(p_j^t - \widehat{p}_j^*\right)^2}{\widehat{p}_j^*} \\ & + w^\tau \frac{\left(\widehat{p}_j^\tau - \widehat{p}_j^*\right)^2}{\widehat{p}_j^*} - \sum_{t=1}^{T} w^t \frac{\left(p_i^t - p_i^*\right)^2}{p_i^*} - \sum_{t=1}^{T} w^t \frac{\left(p_j^t - p_j^*\right)^2}{p_j^*}, \end{aligned} \tag{9}$$

which yields:

$$\begin{aligned} \Delta mH \quad = \quad & \frac{1}{N^2\left(p_i^* + w^\tau \delta\right)\left(p_j^* - w^\tau \delta\right)} \left[\frac{1 - w^\tau}{w^\tau}\right]\left(p_i^* + p_j^*\right) \\ & + \frac{2}{N\left(p_i^* + w^\tau \delta\right)}\left(p_i^\tau - p_i^*\right) - \frac{2}{N\left(p_j^* - w^\tau \delta\right)}\left(p_j^\tau - p_j^*\right) \\ & - \frac{1}{N\left(p_i^* + w^\tau \delta\right)} \sum_{t=1}^{T} w^t \frac{\left(p_i^t - p_i^*\right)^2}{p_i^*} + \frac{1}{N\left(p_j^* - w^\tau \delta\right)} \sum_{t=1}^{T} w^t \frac{\left(p_j^t - p_j^*\right)^2}{p_j^*}. \end{aligned} \tag{10}$$

As equation (10) shows, the migration of probability mass, δ, generates an *a priori* ambiguous effect on the index. A reasonable result since the index is measuring inequality as increased association and migration 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 following situations describe the sensitivity of the indices with respect to the 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: The effect of the migration is ambiguous as shown by equation (10).
- 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 the Appendix 8, the effect is again 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 (12).[20] 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.

[19] Such that: $p_j^\tau \geq \delta \geq \frac{1}{N^\tau} > 0$.
[20] See Appendix 8.

2. The probabilities in the arrival state 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. The change in the index due to migration in this situation is in equation (13).[21] A necessary condition for this migration to reduce the value of the index is: $p_i^\tau < p_i^l$, which makes sense to assume if the probabilities in the arrival state are meant to be equalized after the migration. The condition is, however, insufficient. [22]
3. The probabilities in the arrival state are identical before the migration. This case means $p_i^t = p_i^* \wedge p_j^t = p_j^* \forall t \in G$ and the migration increases inequality. This is the case relevant to condition 8.

- When the probabilities across types are such that there is complete, or absolute, 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 $T = A$. The respective proofs are in Appendix 8, and used also to show that the indices attain their maximum in situations of complete, or absolute, association between all types and all outcomes (not just the departure and arrival states):

1. Whenever $T \neq A$ migration can either leave the initial complete association intact thereby not affecting inequality or complete association between the types and the two (departure and arrival) states.
2. Whenever $T = A$ any migration breaks the initial absolute association thereby decreasing inequality.

3.4 The indices and other concepts and indices of inequality of opportunity

This section compares the performance of the dissimilarity indices in ranking societies, in terms of inequality of opportunity, vis-a-vis other well-known concepts and measurements of inequality of opportunity in the literature. The idea is not to show any superiority of one approach over others, especially in an area where "no consensus has been reached" (Lefranc et al., 2009, p. 1189); but to explore the differences in rankings that different approaches to inequality-of-opportunity measurement yield, focusing on the differences that arise from introducing the dissimilarity indices.

3.4.1 The conception of Lefranc *et al.* (2008,2009)

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. Such is a very reasonable definition from the point of view of a hypothetical outsider who has to choose between different types and is concerned about the risk and return mix involved in every type's conditional distribution of outcomes. As Lefranc et al. (2009) explain, the ranking of societies following a second-order dominance criterion may disagree with that based on the notion that

[21] See Appendix 8.
[22] See equation (13) in Appendix 8.

equality of opportunity is only achieved under *circumstance neutralization*, which is a more general version of Roemer's literalist definition, and is followed by this paper's indices. Since lack of second-order dominance can be achieved with different combinations of compared distributions, Lefranc *et al.*'s (2008) criterion may declare two societies to be opportunity equal even when in one all conditional distributions are identical and in the other one both the mean attainment and the degree of inequality in one type may be higher than in another type (in a case of two types in both societies). In an extreme situation, the dissimilarity indices might declare the former society as being opportunity equal while regarding the latter as perfectly opportunity unequal if the latter exhibits complete association between types and outcomes.[23] These discrepancies are best illustrated by looking at Lefranc *et al.*'s Gini index of inequality of opportunity:

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

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. Whenever a society is opportunity-equal according to Roemer's literalist criterion, and the dissimilarity indices, GIO is zero, thereby measuring equality of opportunity according to both Roemer's and Lefranc *et al.*'s definitions. However the reverse is not true: GIO may be zero and imply equality of opportunity in terms of lack of second-order dominance also when distributions of the advantage/outcome are not equal (thereby implying inequality according to alternative definitions). These differences in rankings stem from attempts at answering different, and complementary questions. The most stringent definitions of equality of opportunity (e.g. Roemer's literalist criterion) and indices based on dissimilarities of probabilities, seek to quantify the departure from *circumstance neutralization,* and tend to be insensitive to the location of the differences between types in the conditional distributions thereby being unable to say anything regarding the desirability of one conditional distribution over another. Whereas, by focusing on second-order dominance, Lefranc *et al.* provide a weaker criterion that is sensitive to the presence of between-group inequality in different parts of the distribution, therefore not measuring inequality as circumstance neutralization but as inability to rank type-conditioned distributions in terms desirability (as judged by the second-order dominance criterion).

3.4.2 The indices based on the types approach (Checci and Peragine, 2005; Roemer, 2006; Ferreira and Gignoux, 2008)

The types approach compares across types/groups a standard that represents and summarizes their conditional distributions. Empirical applications of this approach have typically associated increasing inequality of opportunity with increasing between-groups/types inequality in mean outcomes. The desirable properties that indices of the types approach satisfy are discussed in Peragine (2004). Roemer (2006) also considers this approach as a less literalist (and empirically demanding) alternative criterion. Its key advantage is that, by comparing mean outcomes, the ensuing rankings satisfy the axiom of *inequality neutrality within types*, which implies that mean-preserving transfers, that affect within-type inequality, do not affect between-type inequality (Peragine, 2004, p. 6). Other quantitative approaches in the literature do not fulfill this axiom, including the dissimilarity indices.

[23] Notice that the lack of overlap implied in this example is different from the specific, popular form of lack of overlap whereby the poorest person in one group/type is richer than the richest person in another group/type.

That is, in the other approaches, within-type transfers (or migrations in probability space) affect both within and between-type inequality.

The dissimilarity indices and indices based on the types approach[24] may disagree in their social rankings in two main instances: Firstly, whenever there is no between-group inequality in the standard of the distribution, the types approach declares equality of opportunity. The dissimilarity indices might not do so because equality of distributional standards (e.g. the means) can occur even when the multinomial distributions are not homogeneous, e.g. when circumstance neutralization is not present. Secondly when two societies are compared and both exhibit complete (or absolute) association but one has between-group inequality, according to a path-independent decomposition, coupled with no within-group inequality and the other one has some within-group inequality and may or may not have between-group inequality. In such a case the relative version of the indices based on 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 complete (or absolute) association between types and outcome categories.[25] This second source of discrepancy disappears if the absolute version of the indices based on the types approach is considered instead.[26] In that case, only the first instance is relevant.

3.4.3 The indices based on the tranches approach (Checci and Peragine, 2005)

Checchi and Peragine (2005) propose an alternative measure of inequality of opportunity also based on inequality indices decomposable into between and within groups. 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 between individuals belonging to different types but exerting the same degree of effort, captured by belonging to the same percentile tranch, thence the name tranches approach. [27] Indices based on the tranches approach agree with the dissimilarity indices in declaring equality of opportunity if and only if conditional distributions of well-being are identical.

By contrast, the relative version of the tranches-approach index used by Checchi and Peragine (2005) does not rank all distributions characterized by complete (or absolute) association as being perfectly opportunity unequal,[28] because it 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 complete (or absolute) association between types and outcome sets,

[24] Typically these inidices belong to families that are decomposable into a between-group and a within-group component in a path-independent way. The preferred one being the mean log deviation because it considers the mean for between-group comparisons. For path-independent decomposability see Foster and Shneyerov (2000).

[25] These two sources of discrepancy remain even if the indices are adjusted according to the suggestion of Elbers et al. (2008).

[26] That is, e.g. quantifying inequality of opportunity using only the between-group components of the decompositions.

[27] Using path-independent decomposition techniques, inequality of opportunity is calculated as the residual from subtracting between-tranch inequality to total inequality (calculated over a smoothed distribution). A relative measure of inequality of opportunity based on the tranches approach can also be constructed by dividing within-tranch inequality measure by total inequality. Desirable properties for indices based on the tranches approach are discussed in Peragine (2004).

[28] The absolute versions, by construction, do not have an upper bound representing a concept of maximum between-group inequality.

which make the dissimilarity indices attain their maximum, is possible with different distributions of the outcome (reflecting effort through the monotoncity assumption) within each conditioning type/group.

3.4.4 The Human Opportunity Index (Barros et al., 2008)

Let p_1^* be the average accomplishment related to a dichotomous outcome (e.g. access to a basic service); p_1^t is defined similarly with respect to a specific group of society, t. Then the human opportunity index (HOI, Barros et al. (2008)) is defined as:

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

where D is a dissimilarity index based on a statistic belonging to the above specified family:

$$D \equiv X_{T,1}^1 / 2p_1^* = \frac{1}{2} \sum_{t=1}^{T} w^t \frac{\left|p_1^t - p_1^*\right|}{p_1^*}.$$

It is worth comparing D and $H_{T,2}^2$, which deals with dichotomous outcomes. [29]

D works with dichotomous variables whereas this paper's indices work with multinomial distributions, including dichotomous variables. Therefore this paper's indices can also be applied to quantify inequality in access to services, although they may not rank societies the same way as D does. D is less useful to compare multinomial distributions because its maximum value (necessary for normalization) does not depend just on the population size. Instead it depends *ad hoc* on the groups' weights.[30]

Barros et al. (2008) extensively discuss the choice of normalization for D: Firstly, it gives D an appealing meaning as "the minimum fraction of the total number of persons with high outcome [...] that needs to be redistributed across circumstance groups in order to [...] ensure equal proportion to low outcome persons in all circumstance groups" (p. 19); secondly, it renders D either insensitive or prone to decrease when p_1^t increases by a common factor for all t.[31]

Regarding similarities, both D and $H_{T,2}^2$ declare equality of opportunity whenever $p_1^t = p_1^* \forall t$. However in the absence of perfect equality of opportunity, D and $H_{T,2}^2$ may not necessarily rank societies consistently among themselves. In particular, they differ in their sensitivity to a special form of "balanced increase in opportunity" considered by Barros et al. (2008): a common additive increase, i.e. a migration that increases the probability of attaining the cherished outcome by the same amount δ for all types. As Barros et al. (2008) show, such migration reduces the value of D, thereby revealing a "pro-growth bias" (p. 29). By contrast, the reaction of $H_{T,2}^2$ depends on whether p_1^* is higher, equal or lower than $\frac{1-\delta}{2}$. If it is higher, then the common additive increases

[29] The HOI follows a tradition of defining welfare indices which account both for the average and the dispersion of a welfare outcome, started by Atkinson (1970) and Sen (1976).

[30] Although this can be overcome by either definining: $w^t = \frac{1}{T}$ (i.e. imposing TPI) or by adopting the Mean relative deviation index, which is normalized by Simpson's interaction measure (Reardon and Firebaugh, 2002, p. 41).

[31] This normalization comes at a cost, acknowledged by the authors. D's maximum value is equal to $\frac{N-1}{N}$ and 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 $p_1^t = 0 \forall t \neq i$. When that happens, $\lim_{N\to\infty} D = 1$. Alternatively, an index $d \equiv D\frac{N}{N-1}$ is equal to 1 whenever there is maximum dissimilarity as measured by D. However d is not population invariant except when there is maximum or minimum dissimilarity. This trade-off between population invariance 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.

raises the value of $H^2_{T,2}$; the opposite happens if $p^*_1 < \frac{1-\delta}{2}$; and the index is insensitive to the migration if both quantities are equal. [32]

The two indices also differ in declaring maximum inequality of opportunity: Maximum dissimilarity for D implies maximum dissimilarity by $H^2_{T,2}$ but the reverse is not true; because for D the only situation in which maximum dissimilarity holds is that of one individual having a value of 1 and the rest of the population having a 0 for p_1. Whereas for $H^2_{T,2}$ 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, $p^t_2 = 0 \forall t \in G_1 \wedge p^t_2 = 1 \forall t \notin G_1$, exhibit complete association. The unique maximum of D is a special case of the set of maxima considered by $H^2_{T,2}$.

These discrepancies emphasize the need for conceptual criteria to decide which type of index to use for dichotomous outcomes. When the dichotomy is about achieving a valuable situation (e.g. access to a service), I advocate using D since it is reasonable to consider that there is more inequality of opportunity when one individual has full certainty of achieving the outcome and all the others have zero probability of attaining it, than when more than one individual has full certainty and the rest of the population has zero probability. I advocate using $H^2_{T,2}$ when the dichotomy reflects two options whose ranking is not immediately obvious in terms of valuability, or when one wants to assess the dissimilarity of conditional distributions, and one is content to declare perfect inequality when one option is exclusively associated with a subset of types and the other option is exclusively associated with the remaining subset. [33]

4 Application to multiple variables

An interesting feature of indices like (7) and (8), measuring inequality of opportunity on probability space, is that they can be applied to situations with multiple variables. The joint consideration of two or more outcomes has advantages over considering inequality on each outcome separately, yet it also bears potential limitations. Both analyses can be deemed complementary.

On the side of the advantages, looking at differences between groups on their joint distributions of outcomes provides information that is not captured by looking at each variable's marginal distribution separately. For instance, because several joint distributions can belong to the same Fréchet class,[34] it is possible that, comparing two societies, a variable-by-variable analysis of inequality of opportunity may yield perfect equality in both societies, i.e. if the types' joint distributions belong to the same Fréchet classes in each and both societies, whereas an analysis focused on the joint distribution differences may yield different rankings. In addition to that, considering inequality over the joint distributions, on top of the marginal distributions, can be justified by extending the concepts of second-order dominance by Lefranc et al. (2008, 2009) to multiple dimensions. That is, individuals may rank certain joint distributions in terms of preferences for mean attainments, risk, and association of outcomes, reflected in certain classes of welfare functions for which dominance conditions could hold. The indices (7) and (8) do not rank inequality of opportunity in terms second-order dominance conditions,[35] but they do declare perfect equality to hold if and only if

[32] See results in Appendix 9.

[33] For instance considering agriculture versus non-agricultural occupations (e.g. Bossuroy and Cogneau, 2008). It is not a priori clear that one of the occupations is better in some meaningful sense; such judgment depends on specific contexts. Another example is to distinguish between blue-collar and white-collar workers, where, as is known (e.g. Giddens, 2006), not all workers qualified as white-collars in developed countries, e.g. low-rank clerks, are, for instance, financially better-off that all workers qualified as blue-collars, e.g. highly skilled artisans.

[34] That is, they share the same marginal distributions (Frechet, 1951).

[35] In order to do that, an extension of the ideas by Lefranc et al. (2008) is required, using multivariate stochastic

the conditional joint distributions are equal across types, thereby extending Roemer's literalist definition, and the circumstance equalization principle, to multiple dimensions. [36]

The differences in rankings arising from focusing on the joint distributions vis-a-vis the marginal distributions of outcomes separately is illustrated in the following example: Consider a society with two types, T_1 and T_2, and two outcomes, Y and X, each taking three values (e.g. $x_1 < x_2 < x_3$). The joint distributions of the two types are the following:

$T_1 =$

	x_1	x_2	x_3
y_1	0.2	0.1	0
y_2	0.1	0.2	0.1
y_3	0	0.1	0.2

$T_2 =$

	x_1	x_2	x_3
y_1	0.1	0.1	0.1
y_2	0.1	0.2	0.1
y_3	0.1	0.1	0.1

If the indices are estimated considering the joint distribution it is clear that perfect equality does not hold (e.g. $p^{T_1}(X = x_1, Y = y_1) \neq p^{T_2}(X = x_1, Y = y_1)$). However if the indices are estimated for each of the marginal distributions separately, perfect equality does hold, for each variable, since the distributions of T_1 and T_2 belong to the same Fréchet class. Now consider a migration within type T_1, such that after the migration: $p^{T_1}(X = x_1, Y = y_1) = 0.1$ and $p^{T_1}(X = x_3, Y = y_1) = 0.1$. Then the indices estimated over the joint distribution decrease in value because the migration generates PPEP. However, after the migration, the joint distributions do not belong anymore to the same Fréchet class: The indices estimated for the marginal distribution of X do not yield perfect equality anymore. Therefore it is possible to decrease inequality of opportunity in the joint distribution of variables even at the expense of increasing inequality of opportunity in (at least one of) the marginal distributions. Hence looking at the joint distributions may provide information different from, and complementary to, an analysis of variable-by-variable.[37]

On the side of the limitations, assessing inequality over probability space carries an *implicit weighting* of the contributions of inequalities in each outcome variable over total inequality. This implicit weighting may be controversial in some situations. For instance, when there is complete (or absolute) association between types and one variable's values/outcomes, and $T \leq m$, where m is the number of categories/values of the variable with whom there is complete (or absolute) association.[38] In such case, complete (or absolute) association is also sufficiently established between the joint variable outcomes and types making the indices attain their maximum value. Generally, if there is complete (or absolute) association between types and each variable belonging to a subset $l \subset b$, where b is the number of variables, then the indices attain their maximum value when applied to the joint distribution of the b variables, if and only if $\exists i \in l \mid T \leq m_i$, where m_i is the number of categories/values of variable i. Thus the partitions of the variables affect the implicit contribution of each variable's inequality over the whole.[39]

In these situations, i.e. when complete (or absolute) association between types and a subset of variables leads to maximum inequality over the whole subset, as measured by the indices (7) and (8), the variable-by-variable analysis can be complemented by an alternative family of composite indicators of inequality of opportunity that aggregates the specific-variable results. For instance:

dominance conditions. The latter have been discussed, among others, by Atkinson and Bourguignon (1982).

[36] In the analysis of multivariate poverty dominance, Duclos et al. (2006) pioneered the idea that even when some dominance conditions can not be ascertained over the marginal distributions, it is possible to find significant areas within the joint distribution in which restricted dominance holds. Thus they justify looking at the joint distributions of outcomes.

[37] At the extreme though, perfect equality of the joint distributions implies perfect equality of the marginals.

[38] And, by construction, $A > m$, for two or more outcomes.

[39] Admittedly, this situation is more problematic when continuous variables are discretized.

$$H_\epsilon = \left[\sum_{i=1}^{b} \varphi_i \left(H_{T,A_i}^2 \right)^{\epsilon} \right]^{\frac{1}{\epsilon}}, \tag{11}$$

where H_{T,A_i}^2 measures inequality of opportunity over variable i, which has A_i values; φ_i is the contribution of inequality over variable i to overall inequality and ϵ is a CES parameter that gives more weight to the more unequal variables as it increases. [40] The advantage of using members of the family in (11)[41] is that complete (or absolute) association between types and a subset of variables does not automatically translate into maximum inequality by (11), i.e. when all variables are considered. However the avoidance of this situation comes at the cost of foregoing information on the joint distributions.

5 Empirical application

This empirical application looks at inequality of opportunity of educational outcomes in Peru. Two discrete outcomes are studied: levels of education attained and quality of education attained, proxied by type of school attended. The adult population is divided into 6 types which result from combining the following three circumstances: gender, father's education and mother's education. All these are beyond the control of individuals, as is demanded by the ineuqality of opportunity approach.[42] The purpose of the application is to compare several cohorts of adults in terms of inequality of opportunity over these outcomes, in order to uncover a possible trend. The indices are therefore calculated for educational level and for quality separately, but also calculated for their joint distribution. Using indices (7) and (8), it is possible to assess the impact of the Type-Population-Invariance condition on the cross-cohort comparison. Inequality trends are also estimated for different type partitions; namely, for a simple gender division, and for a division only according to parental educational background. The idea is to assess the robustness of the trend results to different forms of type (dis)aggregation; and also to explore whether trends in inequality based on certain circumstances (e.g. gender) help to explain the trends in inequality based on the types that, in turn, are constructed from those circumstances, interacted with others.[43] In the inequality-of-opportunity literature, Ferreira and Gignoux (2008) discuss the impact of having limited information for the definition of types, on measuring between-group inequality on value space: As more groups (therefore finer partitions) are allowed for, more of the total inequality is explained by the between-group element.[44] When measuring between-group inequality on probability space, the definitions of groups may vary depending on available data. In the case of inequality of opportunity, it depends on whether there is information for the circumstances that are deemed to define the types.[45] Different group definitions may generate different conditional distributions, because

[40] H can also be defined for $\overline{H}_{T,A_i}^2$.

[41] For instance: $H = \frac{1}{b} \sum_{i=1}^{b} H_{T,A_i}^2$.

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

[43] I would like to thank an anonymous referee for suggesting performing this analysis.

[44] Considering that, in applications, it is difficult to gather full information on circumstances determining types, Ferreira and Gignoux (2008) state that empirical estimates provide "lower bounds" of the true/actual degree of inequality of opportunity. In the case of the indices (7) and (8), a finer partition of types leads to unambiguously higher between-group inequality when $T \geq M$, for the minimum T considered. Otherwise the result is ambiguous.

[45] The choice of circumstances to be considered for the definition of types may not just depend on the availability of information (on them), but also on the policy implications of the analysis, i.e. on whether these circumstances are considered worth compensating for.

they consider different choices and combinations of circumstances, each in turn differentially associated with outcomes. Therefore, different inequality patterns may ensue from different group/type partitions, which warrant a sensitivity analysis.[46]

Following the previous section's discussion, a priori it is impossible to say whether inequality over a joint distribution (measured on probability space) is higher or lower than over its related marginal distributions seperately. In fact changes in inequality over joint distributions do not necessarily imply changes in inequality over its marginal distributions and viceversa, even though the latter stem from linear combinations of the former. Hence computing inequality over the joint distribution of well-being outcomes can provide information about inequality of opportunities on top and beyond that provided by the marginal distributions.

5.1 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).

The combination of these two variables yields 7 categories because the "no education" entries only interact with each other. The 6 types ensue from combining the following three variables:

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

Strictly, the combination of these three variables should yield 8 categories but, due to sample size considerations, I consider equally people who have one parent with more than complete primary regardless of whether this parent is the father or the mother.[47] I also perform the analysis for just a gender disaggregation, and for only three types produced by pooling gender in every

[46] In the Geography literature, where types/groups are the equivalent of "modifiable entities", it has long been noticed that different discrete partitions of areas may lead to different statistical patterns, e.g. of the average household size per district. This is the so-called "Modifiable Areal Unit Problem" (MAUP). For a discussion see Openshaw (1984).

[47] By compressing these two categories, thus ending up with 6, as opposed to 8, types, I can exploit the increased gain in sample size and produce indices for a finer partition of cohort age-brackets.

parental-educational-background category. The respective sample sizes, together with the marginal distributions for the two variables, are in the Appendix 7. From the tables it stands out that, the younger the cohort, the higher the schooling attainment: the rates of the two lowest educational outcomes decline monotonically with the youth of the cohort as the proportions of the two highest educational outcomes increase (Figure 1). Within each cohort, adults whose parents had higher levels of education have distributions of educational attainment, and type of school attendend, with higher incidences of better outcomes i.e. that first-order stochastically dominate the distributions of those whose family background exhibits lower educational attainment (Tables 9 and 10). Between males and females, distributional differences are narrower among the youngest cohorts in terms of school attendance (especially on the extreme outcomes), and type of school attendend (Figure 4).

5.2 Results

The values of the indices (7) and (8) for the two educational variables, separately and jointly, appear in Table 2 and 3, respectively. The three estimations agree in showing a pattern of reduction in inequality of opportunity over these educational outcomes from older to younger cohorts. The degree of overlap between the contiguous confidence intervals varies by cohort pairs. In cases like that of type of school attended and the 42-51 and 52-61 cohorts as measured by index (8), the lack of overlap shows a statistically significant difference in the indices, whereas for the same variable, but comparing 52-61 versus 61+ with index (7), the degree of overlap suggests that the differences may not be statistically significant. However, for type school attended Tables 2 and 3 show a monotonic decrease in inequality of opportunity from old to young cohorts.[48] The differences between the indices of the extreme old and young cohort are statistically significant.

In the case of education levels and the joint distribution the patterns are not monotonic for the index not fulfilling *(TPI)*, (Table 2). A reduction trend is apparent in both but it is reversed from the second-to-youngest to the youngest cohort. For education levels and the joint distribution, the youngest three cohorts show less inequality than the oldest two and the differences between the youngest three and the oldest cohort are statistically significant. On the other hand, according to the index fulfilling *(TPI)*, the reduction in inequality over levels and the joint outcome is monotonic. Therefore this application offers empirical evidence in favour of the hypothesis that educational outcomes among younger Peruvian adults are less opportunity unequal vis-a-vis older Peruvian adults. This trend is clearest for the type of school attended, which may be driven more by an increase in enrollment than further homogenization of private school enrollment. For education levels and the joint distribution the reduction is not monotonic for index (7), although even for this index there is a significant reduction in inequality comparing the extreme ends of the cohort spectrum. In other words, the youngest cohorts's joint and marginal distributions are closer to the ideal of *circumstance neutralization* than those of their oldest fellow citizens.

[48] Percentile bootstrapped 95% confidence intervals appear in brackets. In each case 500 resamplings were performed.

Table 2: Inequality of educational opportunities in Peru: Two indicators of educational attainment using an index not fulfilling TPI

Indicators	Cohorts by year ranges				
	22-31	32-41	42-51	52-61	61+
Education level	0.2936	0.2788	0.3017	0.3203	0.3444
	[0.2806,0.3080]	[0.2658,0.2949]	[0.2862,0.3188]	[0.2977,0.3513]	[0.3181,0.3735]
Type school attended	0.2340	0.2440	0.2612	0.3077	0.3151
	[0.2136,0.2586]	[0.2189,0.2701]	[0.2369,0.2909]	[0.2774,0.3451]	[0.2841,0.3535]
Joint distribution	0.2366	0.2245	0.2415	0.2596	0.2784
	[0.2265,0.2470]	[0.2120,0.2396]	[0.2275,0.2570]	[0.2447,0.2860]	[0.2586,0.3097]

Table 3: Inequality of educational opportunities in Peru: Two indicators of educational attainment using an index fulfilling TPI

Indicators	Cohorts by year ranges				
	22-31	32-41	42-51	52-61	61+
Education level	0.3060	0.3265	0.3725	0.4023	0.4223
	[0.2952,0.3163]	[0.3117,0.3407]	[0.3541,0.3901]	[0.3761,0.4270]	[0.3959,0.4471]
Type school attended	0.2187	0.2506	0.2897	0.3699	0.4070
	[0.2017,0.2344]	[0.2280,0.2714]	[0.2634,0.3137]	[0.3341,0.4025]	[0.3759,0.4359]
Joint distribution	0.2459	0.2613	0.2977	0.3250	0.3363
	[0.2372,0.2542]	[0.2489,0.2732]	[0.2828,0.3118]	[0.2988,0.3493]	[0.3139,0.3573]

The robustness and sensitivity of these results is first probed by defining types only in terms of gender. The results are in Appendix Tables 11 and 12. Considering that gender proportions are stable over time in countries, like Peru, not characterized by strong son preferences, the two indices consistently show that, after an increase from the oldest to the second-to-oldest cohort, inequality tends to decrease among the younger cohorts for type of education, level of education and the joint distribution. The youngest cohort exhibits the lowest degree of inequality compared to all the other cohorts, reflecting progress in reducing gender differences in educational opportunities.

Tables 13 and 14 show the inequality trends for three types defined on parental background (i.e. by aggregating the six types over gender). Now the trend depends on whether the index satisfies *(TPI)* because the types' proportions changed across cohorts. According to index (7), there is no clear pattern of decline in inequality. For type of education and the joint distribution a U-shaped trend shows up, but not for educational level. This lack of steady decline, combined with the results for an exclusive gender division, may explain the patterns in Table 2, whereby a decline in inequality is slightly reversed at the youngest cohort. By contrast, according to index (8), the decline in inequality is monotonic for type of education, level, and the joint outcome. Hence this result, coupled with the gender-specific results (mild increase followed by steady decline), helps to explain the pattern of monotonic decline found in Table 3. Overall, the conclusions from the empirical assessment depend both on whether a "representative agent" approach is undertaken (i.e. the index fulfills *TPI*) and on the type definition. For instance, the evidence from the most refined type definition used, together with an index fulfilling *(TPI)*, suggest a monotonic decrease in inequality of educational opportunity over the two variables considered.

6 Concluding remarks

This paper proposes dissimilarity indices, based on Pearson's statistic, for the analysis of inequality of opportunity. Influenced by Roemer's and Fleurbaey's conceptions, they measures inequality of opportunity in proportion to the degree of association between sets of circumstances and sets of outcomes. A higher degree of association is related to higher dissimilarity of distributions conditioned on type-belonging; and, in turn, higher inequality of opportunity.

These indices bear the advantage of being able to capture differences in the joint distributions of multiple outcomes. Therefore they provide additional information on inequality, considering that joint distributions may be different even when marginal distributions are identical across groups, and people may have preferences over these differences. On the other hand, under certain conditions involving the number of types and outcome values, maximum inequality over the joint distribution may be attained sufficiently when complete (or absolute) association is present in only one separate dimension. Another contribution is the proposal, and empirical illustration, of two alternative indices: one is sensitive to changes in group composition; the other one is not, thereby measuring inequality with a "representative agent" approach.

This paper's index, the *HOI*'s D, and the types and tranches approaches, do not judge which group or type is the most advantaged. For that an analysis of risk, return and stochastic dominance is required, as per Lefranc et al. (2008). All approaches, including this paper's, agree in classifying societies as opportunity equal when conditional distributions are identical. However the dissimilarity indices and the types approach may disagree in ranking opportunity-unequal societies since the latter relates inequality of opportunity to inequality over a distributional standard of the variable. 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 despite having the same distributional standard.

On the other hand the dissimilarity indices agree with the tranches approach in declaring inequality of opportunity *if and only if* conditional distributions of well-being are identical. However the indices and the relative version of the tranches approach may disagree on the ranking of distributions characterized by complete (or absolute) association. The indices rank all these distributions as being perfectly opportunity unequal; whereas in the tranches approach different between-tranch inequality may yield different values of the relative indicator for different societies characterized by complete (or absolute) association.

The dissimilarity indices belong to a family of inequality-of-opportunity indices, including those of the types and the tranches approaches, which do not account explicitly for average attainments. Among those which do, the most prominent are the *GIO* and the *HOI*.[49] A major difference between the dissimilarity indices and the *GIO* is that the latter declares equality of opportunity not only when conditional distributions are identical. Regarding the *HOI*, both its D and this paper's indices agree on declaring equality if and only if conditional distributions are equal. They also agree on declaring perfect inequality whenever one individual has full certainty of attaining an outcome while the rest of his/her society has zero chance of attaining it. However they may disagree on: (1) ranking societies in intermediate situations of inequality with imperfect association; (2) ranking societies with different forms of complete (or absolute) association between types and outcomes;

[49] The dissimilarity indices discussed in this paper, as well as others from the segregation literature, could also be used to generate indices similar in spirit and shape to the *HOI*. For instance, for a continuous variable one such index could be: $H = \overline{y}\left(1 - H^2_{T,A}\right)$, where $\overline{y}$ is the mean of a continuous variable y. As another example, for a discrete variable it could be: $H = \frac{\overline{x}}{x_{\max}}\left(1 - H^2_{T,A}\right)$, where $\overline{x}$ and $x_{\max}$ are respectively the mean and maximum value of a discrete variable x. The suitability of these measures may be worth discussing in future research, along the lines laid out by Barros et al. (2008).

and (3) their sensitivity to "balanced increases in opportunity".

In the empirical application to educational opportunity in Peru the indices provide evidence of a statistically significant reduction in educational inequality of opportunity among the younger cohorts of adults, vis-a-vis the oldest ones, considering the most refined type definition that combines gender and parental educational background. The reduction trend is monotonic for type of school attended (which makes a difference in terms of quality of education and labour market outcomes in Peru). It is also observed, but not monotonically for index (7), for educational levels and the joint distribution of these two outcomes. 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 occupation or ethnicity could have been considered with richer and larger samples. The definition of types is important not only because, in principle, a more refined partition may lead to higher between-group inequality (Ferreira and Gignoux, 2008), but also because, on probability space, different combinations of circumstances (when defining types) may exhibit different degrees of association with the same outcomes. Likewise, the empirical application illustrates the potential differences in results ensuing from using indices that are (or not) sensitive to type/group compositions. For instance, in the application to Peru, only the index that satisfies TPI declares a monotonic reduction in inequality of opportunity, from older to younger cohorts, over type of school, level of education and joint outcomes, that is robust to the three types' partitions.

This paper seeks to emphasize the value of the dissimilarity approach, from the segregation literature, to measuring inequality of opportunity. Future work ought to explore further the potential and limitations of measuring inequality over multiple dimensions of well-being on probability space; as well as focus on finding quantitative tools to measure multidimensional inequality of opportunity using continuous variables (and combinations with discrete variables).

7 Appendix

Table 4: Sample sizes for the empirical application

Type	**Cohorts by year ranges**				
	22-31	32-41	42-51	52-61	61+
Male, both parents had up to complete primary	2,715	2,969	2,418	1,608	1,805
Male, one parent more than complete primary	586	328	200	101	106
Male, both parents had more than complete primary	876	410	232	119	103
Female, both parents had up to complete primary	3,057	3,207	2,390	1,655	1,526
Female, one parent more than complete primary	625	431	218	96	71
Female, both parents had more than complete primary	875	422	245	92	97

Table 5: Distributions of educational attainment

Cohort	Type	Educational level			
		No education	No secundary	No tertiary	Tertiary
22-31	M, both parents <= complete primary	1.07	32.49	50.31	16.13
	M, one parent > complete primary	0.51	7.00	50.68	41.81
	M, both parents > complete primary	0.11	1.94	35.84	62.10
	F, both parents <= complete primary	5.14	42.30	37.23	15.34
	F, one parent > complete primary	0.16	10.56	47.52	41.76
	F, both parents > complete primary	0.34	2.40	33.71	63.54
32-41	M, both parents <= complete primary	1.72	37.82	44.22	16.23
	M, one parent > complete primary	1.22	7.32	47.87	43.60
	M, both parents > complete primary	0.49	2.93	31.46	65.12
	F, both parents <= complete primary	9.67	45.15	31.15	14.03
	F, one parent > complete primary	1.39	14.39	44.78	39.44
	F, both parents > complete primary	0.47	3.79	27.01	68.72
42-51	M, both parents <= complete primary	2.98	45.95	35.73	15.34
	M, one parent > complete primary	0	6.50	41.50	52.00
	M, both parents > complete primary	0	2.16	31.03	66.81
	F, both parents <= complete primary	19.12	47.66	23.68	9.54
	F, one parent > complete primary	2.29	20.64	47.25	29.82
	F, both parents > complete primary	1.22	4.90	32.65	61.22

Table 6: Distributions of educational attainment

Cohort	Type	Educational level			
		No education	No secundary	No tertiary	Tertiary
52-61	M, both parents <= complete primary	7.71	57.15	22.89	12.25
	M, one parent > complete primary	0	27.72	43.56	28.71
	M, both parents > complete primary	0	4.20	26.05	69.75
	F, both parents <= complete primary	36.44	45.68	13.29	4.59
	F, one parent > complete primary	4.17	39.58	33.33	22.92
	F, both parents > complete primary	2.17	15.22	40.22	42.39
62+	M, both parents <= complete primary	17.89	64.88	12.80	4.43
	M, one parent > complete primary	1.89	30.19	44.34	23.58
	M, both parents > complete primary	0	12.62	40.78	46.60
	F, both parents <= complete primary	49.08	42.86	6.16	1.90
	F, one parent > complete primary	1.41	49.30	32.39	16.90
	F, both parents > complete primary	2.06	21.65	47.42	28.87

Table 7: Distributions of type of school attended

Cohort	Type of individual	Type of school		
		No education	Public	Private
22-31	M, both parents <= complete primary	1.07	94.33	4.60
	M, one parent > complete primary	0.51	86.69	12.80
	M, both parents > complete primary	0.11	72.03	27.85
	F, both parents <= complete primary	5.14	89.43	5.43
	F, one parent > complete primary	0.16	83.84	16.00
	F, both parents > complete primary	0.34	69.94	29.71
32-41	M, both parents <= complete primary	1.72	93.40	4.88
	M, one parent > complete primary	1.22	85.37	13.41
	M, both parents > complete primary	0.49	70.98	28.54
	F, both parents <= complete primary	9.67	84.75	5.58
	F, one parent > complete primary	1.39	83.53	15.08
	F, both parents > complete primary	0.47	64.93	34.60
42-51	M, both parents <= complete primary	2.98	92.93	4.09
	M, one parent > complete primary	0	89.50	10.50
	M, both parents > complete primary	0	77.59	22.41
	F, both parents <= complete primary	19.12	76.61	4.27
	F, one parent > complete primary	2.29	83.49	14.22
	F, both parents > complete primary	1.22	71.43	27.35

Table 8: Distributions of type of school attended

Cohort	Type of individual	Type of school		
		No education	Public	Private
52-61	M, both parents <= complete primary	7.71	87.87	4.42
	M, one parent > complete primary	0	93.07	6.93
	M, both parents > complete primary	0	73.95	26.05
	F, both parents <= complete primary	36.44	61.69	1.87
	F, one parent > complete primary	4.17	86.46	9.38
	F, both parents > complete primary	2.17	78.26	19.57
62-+	M, both parents <= complete primary	17.89	79.56	2.55
	M, one parent > complete primary	1.89	83.96	14.15
	M, both parents > complete primary	0	81.55	18.45
	F, both parents <= complete primary	49.08	48.49	2.42
	F, one parent > complete primary	1.41	85.92	12.68
	F, both parents > complete primary	2.06	73.20	24.74

Figure 1: Distribution of educational attainment by cohorts

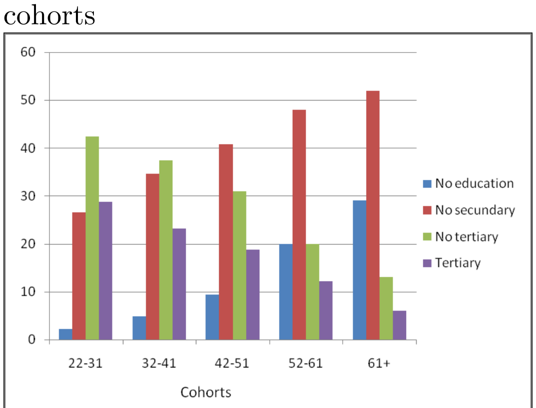

Figure 2: Distribution of school attendance by cohorts

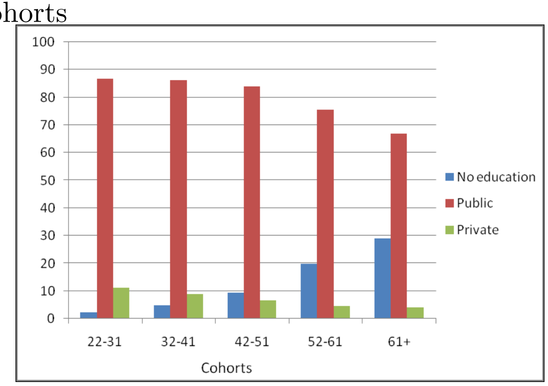

Figure 4: Distribution of school attendance by gender and cohorts

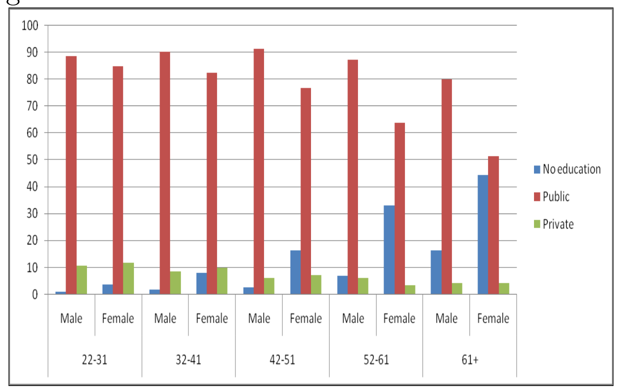

Figure 3: Distribution of educational attainment by gender and cohorts

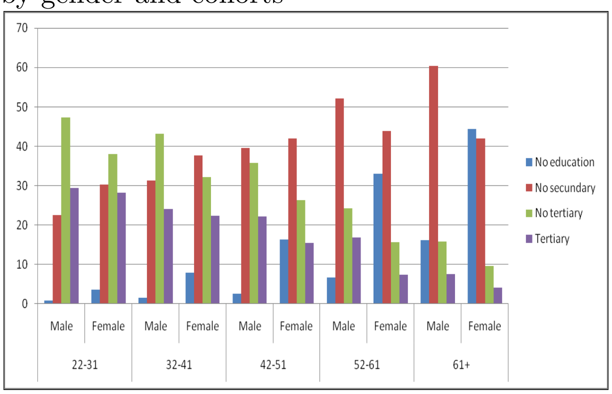

Table 9: Cumulative distributions of educational attainment by parental educational background

Cohort	Type	Educational level			
		No education	No secundary	No tertiary	Tertiary
22-31	Both parents <= complete primary	3.23	40.91	84.29	100
	One parent > complete primary	0.33	9.17	58.26	100
	Both parents > complete primary	0.22	2.39	37.17	100
32-41	Both parents <= complete primary	5.85	47.47	84.91	100
	One parent > complete primary	1.32	12.65	58.77	100
	Both parents > complete primary	0.48	3.85	33.05	100
42-51	Both parents <= complete primary	11.00	57.80	87.54	100
	One parent > complete primary	1.19	15.07	59.57	100
	Both parents > complete primary	0.63	4.19	36.06	100
52-61	Both parents <= complete primary	22.28	73.61	91.64	100
	One parent > complete primary	2.03	35.53	74.11	100
	Both parents > complete primary	0.95	9.95	42.18	100
62-+	Both parents <= complete primary	32.18	86.97	96.73	100
	One parent > complete primary	1.70	39.55	79.10	100
	Both parents > complete primary	1.00	18.00	62.00	100

Table 10: Cumulative distributions of type of school attended by parental educational background

Cohort	**Type of individual**	**Type of school**		
		No education	Public	Private
22-31	Both parents <= complete primary	3.23	94.96	100
	One parent > complete primary	0.33	85.55	100
	Both parents > complete primary	0.22	71.21	100
32-41	Both parents <= complete primary	5.85	94.76	100
	One parent > complete primary	1.32	85.64	100
	Both parents > complete primary	0.48	68.39	100
42-51	Both parents <= complete primary	11.00	95.82	100
	One parent > complete primary	1.19	87.56	100
	Both parents > complete primary	0.63	75.05	100
52-61	Both parents <= complete primary	22.28	96.87	100
	One parent > complete primary	2.03	91.88	100
	Both parents > complete primary	0.95	76.78	100
62-+	Both parents <= complete primary	32.18	97.50	100
	One parent > complete primary	1.70	86.44	100
	Both parents > complete primary	1.00	78.50	100

Table 11: Inequality of educational opportunities in Peru: Men versus women using an index not fulfilling TPI

Indicators	Cohorts by year ranges				
	22-31	32-41	42-51	52-61	61+
Education level	0.1392	0.1780	0.2521	0.3427	0.3116
	[0.1178,0.1578]	[0.1573,0.1965]	[0.2305,0.2721]	[0.3137,0.3693]	[0.2774,0.3425]
Type school attended	0.0950	0.1506	0.2391	0.3300	0.3133
	[0.0749,0.1115]	[0.1306,0.1682]	[0.2179,0.2586]	[0.2997,0.3578]	[0.2814,0.3422]
Joint distribution	0.1422	0.1831	0.2619	0.3459	0.3194
	[0.1212,0.1605]	[0.1620,0.2021]	[0.2400,0.2822]	[0.3188,0.3711]	[0.2885,0.3475]

Table 12: Inequality of educational opportunities in Peru: Men versus women using an index fulfilling TPI

Indicators	Cohorts by year ranges				
	22-31	32-41	42-51	52-61	61+
Education level	0.1402	0.1798	0.2522	0.3430	0.3092
	[0.1185,0.1590]	[0.1589,0.1985]	[0.2305,0.2721]	[0.3141,0.3696]	[0.2755,0.3396]
Type school attended	0.0963	0.1528	0.2392	0.3303	0.3110
	[0.0757,0.1132]	[0.1324,0.1708]	[0.2179,0.2587]	[0.3004,0.3578]	[0.2796,0.3394]
Joint distribution	0.1432	0.1849	0.2620	0.3462	0.3173
	[0.1220,0.1616]	[0.1635,0.2040]	[0.2401,0.2822]	[0.3191,0.3714]	[0.2869,0.3451]

Table 13: Inequality of educational opportunities in Peru: Adults with differentl parental education background using an index not fulfilling TPI

Indicators	Cohorts by year ranges				
	22-31	32-41	42-51	52-61	61+
Education level	0.3415	0.3113	0.3121	0.2974	0.3463
	[0.3290,0.3537]	[0.2957,0.3261]	[0.2938,0.3294]	[0.2680,0.3242]	[0.3180,0.3724]
Type school attended	0.2193	0.2122	0.1870	0.1904	0.2123
	[0.2023,0.2350]	[0.1901,0.2321]	[0.1595,0.2110]	[0.1549,0.2203]	[0.1792,0.2409]
Joint distribution	0.3554	0.3241	0.3206	0.3115	0.3597
	[0.3430,0.3674]	[0.3071,0.3402]	[0.3001,0.3398]	[0.2792,0.3408]	[0.3299,0.3872]

Table 14: Inequality of educational opportunities in Peru: Adults with differentl parental education background using an index fulfilling TPI

Indicators	Cohorts by year ranges				
	22-31	32-41	42-51	52-61	61+
Education level	0.3621	0.3797	0.4107	0.4314	0.4574
	[0.3489,0.3749]	[0.3618,0.3967]	[0.3885,0.4318]	[0.3963,0.4638]	[0.4283,0.4848]
Type school attended	0.2044	0.2257	0.2335	0.3032	0.3456
	[0.1880,0.2196]	[0.2018,0.2473]	[0.2062,0.2579]	[0.2641,0.3377]	[0.3196,0.3697]
Joint distribution	0.3747	0.3911	0.4185	0.4424	0.4623
	[0.3613,0.3876]	[0.3713,0.4099]	[0.3937,0.4419]	[0.4070,0.4752]	[0.4327,0.4902]

8 Appendix

Proofs

1. **The ambiguous impact of a migration when the probabilities in the departure state are identical but the probabilities in the arrival state are different.**

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 (12), which stems from plugging $p_j^t = p_j^* \forall t \in G$ into (10). 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 (captured by the other two elements of the right-hand side). Hence the ambiguity.

- 1.

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

2. **The ambiguous impact of a migration when the probabilities in the departure state are identical and the probabilities in the arrival state are rendered identical *after* the migration.**

In this case, besides having $p_j^t = p_j^* \forall t \in G$, the following also holds: $p_i^l = p_i^m \forall l, m \neq \tau \rightarrow p_i^* = w^\tau p_i^\tau + (1 - w^\tau)\, p_i^l$. Plugging these two conditions into (10) yields (13). A necessary condition for this migration to reduce the value of the index is: $p_i^\tau < p_i^l$. This condition must be assumed if probabilities in the arrival state are expected to be equalized after the migration. However the condition is not sufficient: The first element of the right-hand side of (13) shows the positive effect of the migration on the index due to the breaking of the initial equality of the probabilities in the departure state, while the second element of (13) shows the negative effect of the migration on the index due to the instituion of equality among the arrival state probabilties after the migration. The net effect of these two contradictory effects is ambiguous a priori.

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

3. **Proof that the indices fulfill condition 8, i.e.** *(SMPEP)* **and condition 1, i.e.** *(PE).*

This case means $p_i^t = p_i^* \wedge p_j^t = p_j^* \forall t \in G$, i.e. there is PPEP involving the departure and arrival states. Plugging the PPEP condition into (10) yields equation (14). Such migration breaks PPEP and thus increases association.[50] Accordingly the index reacts by increasing its value, as shown by (14). Conversely, any migration that restores PPEP reduces the value of the index. Now, when $p_i^t = p_i^* \wedge p_j^t = p_j^* \forall t \in G, \forall i, j \in O \times O$, equation (14) proves that any migration disturbing an initial situation of perfect equality raises the value of the index, i.e. the index reaches its minimum value in that situation of perfect equality.

$$\Delta mH = \frac{1}{N^2\left(p_i^* + w^\tau \delta\right)\left(p_j^* - w^\tau \delta\right)}\left[\frac{1 - w^\tau}{w^t}\right]\left(p_i^* + p_j^*\right) > 0 \tag{14}$$

4. **Proof that the indices' values do not increase when a migration breaks complete or absolute assocation between types and the departure and arrival states; and therefore that they attain their maximum value if and only if there is complete or absolute association.**

There are three cases depending on whether $T < A, T > A$ or $T = A$:

- **Case 9** *When $T < A$ there are two possible sub-situations:*

  1. A migration of a member of type τ from j to i that leaves complete 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 (10) yields $\Delta mH = 0$. In words, a migration within states exclusively associated to the type

[50] The result assumes $p_j^* > w^\tau \delta \wedge 0 \leq p_i^* \leq 1 - w^\tau \delta$.

to which the migrating unit belongs leaves the index unchanged.This result also holds when complete 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 complete association between the types and the departure and initial states. This sort of migration requires that type τ is exclusively associated with outcome j but not with i.[51] This situation involves another type k which initially, before the migration, is exclusively associated with outcome i. This type of migration implies: $w^{\tau}p_j^{\tau} = p_j^*$, $w^k p_i^k = p_i^*$, $p_i^t = 0 \forall t \neq k$ (including $p_i^{\tau} = 0$). Plugging these conditions into equation (10) yields equation (15), which shows that the value of the index decreases. This result also holds if the whole contigency table exhibits complete association, i.e. when the index attains its maximum value before migration. Equation (15) proves that a pairwise breakup of complete 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.

$$\Delta mH = -\frac{p_i^*}{N\left(p_i^* + w^{\tau}\delta\right)}\left[\frac{w^k + w^{\tau}}{w^{\tau}w^k}\right] < 0 \tag{15}$$

Case 10 *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 $T < A$,*[52] *and has the same effect: a reduction in the value of the index.*[53]

Case 11 *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 complete 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 $p_j^{\tau} = \delta, \widehat{p}_j^{\tau} = 0, p_i^{\tau} = 0, \widehat{p}_i^{\tau} = \delta$. This migration leaves the index also unchanged, including when there is complete association across the whole contingency table.

  2. A migration of a member of type τ from j to i that breaks complete association between types and the departure and arrival states. Unlike the first sub-situation, now $p_j^{\tau} > \delta, \widehat{p}_j^{\tau} = 0, p_i^{\tau} = 0, \widehat{p}_i^{\tau} = \delta$. This migration breaks the complete association and reduces the value of the index. The same result ensues if complete association is prevalent across the whole contingency table.

The joint proof for these latter two sub-situations is the following: In a situation of complete association, when $T > A$, state j is exclusively associated with a subset of types: $G_j \subset G$. 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

[51] The opposite, that the type is exclusively associated with the final state of migration and not with the original one, is impossible by definition of the example in which the idea is to break initial complete association.

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

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

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.[54] Following equation (9) the change in the index is:

$$\begin{aligned}
\Delta mH &= \left(w^j - w^\tau\right)\left[\frac{\left(1-\left(w^j-\delta w^\tau\right)\right)^2}{w^j-\delta w^\tau} + \frac{\left(0-\left(wi+\delta w^\tau\right)\right)^2}{w^i+\delta w^\tau} - \frac{\left(1-w^j\right)^2}{w^j} - \frac{\left(0-w^i\right)^2}{w^i}\right] \quad (16)\\
&+ w^i\left[\frac{\left(0-\left(w^j-\delta w^\tau\right)\right)^2}{w^j-\delta w^\tau} + \frac{\left(1-\left(wi+\delta w^\tau\right)\right)^2}{w^i+\delta w^\tau} - \frac{\left(0-w^j\right)^2}{w^j} - \frac{\left(1-w^i\right)^2}{w^i}\right]\\
&+ w^\tau\left[\frac{\left(1-\delta-\left(w^j-\delta w^\tau\right)\right)^2}{w^j-\delta w^\tau} + \frac{\left(\delta-\left(wi+\delta w^\tau\right)\right)^2}{w^i+\delta w^\tau} - \frac{\left(1-w^j\right)^2}{w^j} - \frac{\left(0-w^i\right)^2}{w^i}\right].
\end{aligned}$$

After some manipulation equation (16) is reduced to the following expression:

$$\Delta mH = -2 + \frac{w^j - \delta w^\tau\left(2-\delta\right)}{w^j - \delta w^\tau} + \frac{w^j + \delta^2 w^\tau}{w^j + \delta w^\tau} \leq 0. \quad (17)$$

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 $\Delta mH < 0$.[55] Equation (17) is sufficient to prove that the index attain its maximum value if and only if there is complete association in the whole table and $T > A$.

9 Appendix

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

Imagine a migration of a percentage δ of individuals from type τ from state 2 to state 1, 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:

$$\begin{aligned}
\Delta mD &= w^\tau\left[\frac{\left|p_1^\tau + \delta - p_1^* - w^\tau\delta\right|}{p_1^* + w^\tau\delta} - \frac{\left|p_1^\tau - p_1^*\right|}{p_1^*}\right] \quad (18)\\
&+ \sum_{t\neq\tau}^{T} w^t\left[\frac{\left|p_1^t - p_1^* - w^\tau\delta\right|}{p_1^* + w^\tau\delta} - \frac{\left|p_1^t - p_1^*\right|}{p_1^*}\right],
\end{aligned}$$

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

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

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 migration is (after some manipulation):

$$\begin{aligned} \Delta mH &= w^\tau \left[\frac{\left(p_1^\tau + \delta - p_1^* - w^\tau \delta\right)^2}{\left(p_1^* + w^\tau \delta\right)\left(1 - p_1^* - w^\tau \delta\right)} - \frac{\left(p_1^\tau - p_1^*\right)^2}{\left(p_1^*\right)\left(1 - p_1^*\right)}\right] \\ &+ \sum_{t \neq \tau}^{T} w^t \left[\frac{\left(p_1^t - p_1^* - w^\tau \delta\right)^2}{\left(p_1^* + w^\tau \delta\right)\left(1 - p_1^* - w^\tau \delta\right)} - \frac{\left(p_1^t - p_1^*\right)^2}{\left(p_1^*\right)\left(1 - p_1^*\right)}\right]. \end{aligned} \tag{19}$$

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 (18) and (19) can be rewritten in terms of D_t^F and D_t^I as:

$$\Delta mD = \sum_{t=1}^{T} w^t \left[D_t^F - D_t^I\right], \tag{20}$$

$$\Delta mH = \sum_{t=1}^{T} w^t \left[\left(D_t^F\right)^2 \frac{\widehat{p}_1^*}{1 - \widehat{p}_1^*} - \left(D_t^I\right)^2 \frac{p_1^*}{1 - p_1^*}\right]. \tag{21}$$

Notice the differences between (20) and (21): in (21) 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.

The impact of a "balanced increase in opportunity", in the form of a common migration, on $H^2_{T,2}$, and D

If the migration of a percentage δ of individuals from state 2 to state 1 is now common to all types, then D changes the following way:

$$\Delta mD = \sum_{t=1}^{T} w^t \left[\frac{\left|p_1^t - p_1^*\right|}{p_1^* + \delta} - \frac{\left|p_1^t - p_1^*\right|}{p_1^*}\right] < 0 \tag{22}$$

Condition (22) shows the "pro-growth bias" of D (Barros et al., 2008, p. 29). The same migration affects $H^2_{T,2}$ the following way:

$$\Delta mH = \sum_{t=1}^{T} w^t \left(p_1^t - p_1^*\right)^2 \left[\frac{\delta\left(2p_1^* - 1 + \delta\right)}{\left(p_1^* + \delta\right)\left(1 - p_1^* - \delta\right)\left(p_1^*\right)\left(1 - p_1^*\right)}\right]. \tag{23}$$

Hence $p_1^* > \frac{1-\delta}{2} \rightarrow \Delta mH > 0$, $p_1^* < \frac{1-\delta}{2} \rightarrow \Delta mH < 0$, $p_1^* = \frac{1-\delta}{2} \rightarrow \Delta mH = 0$.That is, the relationship between the percentage of people with access to the valuable outcome before the migration, i.e. p_1^*, and the magnitude of the migratio, δ, determines the magnitude and nature of the sensitivity of $H^2_{T,2}$.

References

Atkinson, A. (1970). On the measurement of economic inequality. *Journal of Economic Theory 2*(3), 244–63.

Atkinson, A. and F. Bourguignon (1982). The comparison of multi-dimensioned distributions of economic status. *Review of Economic Studies XLIX*, 183–201.

Barros, R. P. d., F. Ferreira, J. Molinas, and J. Saavedra (2009). *Measuring inequality of opportunity in Latin America and the Caribbean*. The World Bank.

Barros, R. P. d., J. Molinas, and J. Saavedra (2008). Measuring inequality of opportunities for children. Unpublished, World Bank, Washington.

Bossuroy, T. and D. Cogneau (2008). Social mobility and colonial legacy in five africa countries. Dial Working Paper 2008-10.

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 University Press.

Duclos, J.-Y., D. Sahn, and S. Younger (2006). Robust multidimensional poverty comparisons. *The Economic Journal 116*(514), 943–68.

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

Everitt, B. (1992). *The analysis of contingency tables* (Second ed.). Chapman & Hall.

Ferreira, F. and J. Gignoux (2008). The measurement of inequality of opportunity: theory and application to latin america. Policy Research Working Paper Series.

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

Foster, J. and A. Shneyerov (2000). Path independent inequality measures. *Journal of Economic Theory 91*, 199–222.

Frechet, M. (1951). Sur les tableaux de correlation dont les marges sont donnes. *Annales de l'Universite de Lyon II 4*.

Giddens, A. (2006). *Sociology* (5 ed.). Polity Press.

Hild, M. and A. Voorhoeve (2004). Equality of opportunity and opportunity dominance. *Economics and Philosophy 20*, 117–45.

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

Hutchens, R. (2001). Numerical measures of segregation: desirable properties and their implications. *Mathematical Social Sciences 42*, 13–29.

James, D. and K. Taeuber (1985). Measures of segregation. *Sociological Methodology 15*, 1–32.

Kendall, M. and A. Stuart (1973). *The Advanced Theory of Statistics* (3 ed.), Volume 2. Griffin.

Lefranc, A., N. Pistolesi, and A. Trannoy (2008). Inequality of opportunities vs. inequality of outcomes: are western societies all alike? *Review of Income and Wealth 54*(4), 513–43.

Lefranc, A., N. Pistolesi, and A. Trannoy (2009). Equality of opportunity and luck: definitions and testable conditions, with an application to income in france. *Journal of Public Economics 93*, 1189–1207.

Openshaw, S. (1984). *The modifiable areal unit problem.* Number 38 in Concepts and techniques in modern geography. Geo Books.

Peragine, V. (2004). Ranking income distributions according to equality of opportunity. *Journal of Economic Inequality 00*, 1–20.

Reardon, S. (2009). Measures of ordinal segregation. In Y. Fluckiger, S. Reardon, and J. Silber (Eds.), *Occupational and Residential Segregation*, Volume 17 of *Research on Economic Inequality.*

Reardon, S. and G. Firebaugh (2002). Measures of multigroup segregation. *Sociological Methodology 32*, 33–67.

Roemer, J. (1998). *Equality of Opportunity.* 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*(1), 19–39.

Silber, J. and G. Yalonetzky (2010). Measuring inequality in life chances when one variable is ordinal.

White, M. J. (1986). Segregation and diversity measures in population distribution. *Population index 52*(2), 198–221.