Società italiana di economia pubblica

XIX CONFERENZA

ECONOMIA DEL CAPITALE UMANO
Istituzioni, incentivi e valutazioni

Pavia, Aule storiche Università, 13 - 14 settembre 2007

HIGHER EDUCATION AND EQUALITY OF OPPORTUNITY IN ITALY

VITO PERAGINE AND LAURA SERLENGA

pubblicazione internet realizzata con contributo della COMPAGNIA di San Paolo

società italiana di economia pubblica

dipartimento di economia pubblica e territoriale – università di pavia

Higher education and equality of opportunity in Italy

Vito Peragine*and Laura Serlenga†

VERY PRELIMINARY VERSION

February 19, 2007

Abstract

In this paper we provide a methodology to test for the existence of equality of opportunity in a given distribution and for ranking distributions according to equality of opportunity. Moreover, we provide some empirical applications of these new evaluation tools: in the first exercise, we evaluate the income distributions of South and North of Italy on the basis of different definitions of equality of opportunity. Then, we repeat the exercise using the graduation marks of Italian students.

1 Motivation

Equality of opportunity (EOp) has gained a central role in public discussions about social justice in western liberal societies. Indeed, this idea has been defended by a number of scholars in recent years, both in the area of political philosophy and normative economics (see Arneson 1989, Barry 1991, Cohen 1989, Dworkin 1981a,b, Rawls 1971, Roemer 1993), and is the leading idea of most political platforms in several western countries.

According to the opportunity egalitarian view, the principle of justice does not require equality of individuals' final achievements; once the means or opportunities to reach a valuable outcome have been equally split among citizens, which particular opportunity, from those open to her, the individual chooses, is a matter of individual choice and is outside the scope of justice. The EOp view combines features of libertarianism and egalitarianism. From the former it borrows the requirement that public policies should be neutral with respect to private goals that motivate individuals in their lives. But, out of egalitarian inspiration, it seeks a genuine equality in conditions that are beyond the

*University of Bari. E-mail: v.peragine@dse.uniba.it
†University of Bari. E-mail: l.serlenga@dse.uniba.it

individual control or a public compensatory intervention in order to offset the ex-ante inequalities.

Some first attempts to translate this philosophical conception in formal economic models have presented the goal of *equalizing opportunities* as a simple goal. In particular, some contributions addressing the question of ranking social states according to equality of opportunity (see, *inter alia*, Ok 1997, Kranich 1996), have formulated the following problem: each individual in a society is endowed with a given (abstract) set of opportunities, assumed to be observable and measurable, and the society is represented as a profile of opportunity sets. Therefore, the problem of measuring the degree of opportunity inequality is handled by characterizing inequality measures (or inequality rankings) of multidimensional distributions of individual opportunities. This approach is surely correct in principle; however, its informational requirements seem too strong to be met in empirical applications.

Meanwhile, another part of economic literature (initiated by Bossert, 1995 and Fleurbaey, 1995a,b and reviewed by Fleurbaey and Maniquet, 1999), mainly devoted to the definition of allocation rules rather than social orderings, has presented the problem of allocating (transferable) resources in order to offset the initial unequal distribution of opportunities. The main conceptual contribution of this literature is a clarification of the relevant and distinct ethical principles involved in the opportunity egalitarianism project. In particular, it has highlighted the fact that the opportunity egalitarian goal is made of two subgoals which are totally distinct and possibly antagonistic. The first subgoal is to neutralize the influence over agents' outcomes of the characteristics that elicit compensation: society should eliminate inequalities due to factors that are beyond the control of people (call these factors *circumstances*). This is called the principle of compensation (Barry 1990, Fleurbaey 1995a). The second subgoal, an expression of the ethics of responsibility, says that society should *not* indemnify people against outcomes that are consequences of causes that are within their control (call these factors *effort*[1], for short). This is called the responsibility principle (Barry 1990, Fleurbaey 1995a).

Although the large consensus gained by the opportunity egalitarian view, and the rich theoretical literature, it is a common practice among applied economist that of evaluating social inequities by looking at the degree of outcome inequality: income inequalities, educational inequalities, and so on.

In this paper we make an effort to propose and apply new measurement tools which are coherent with the opportunity egalitarian ethics. The aim of the paper is twofold. The fist goal is to provide a theoretically sound methodology to evaluate opportunity inequality. The second goal is to provide an empirical application of these new evaluation tools. We believe that our analysis is able to shed some light on aspects otherwise undetected and undetectable by previous distributional analysis.

[1] In this paper we use indifferently the expressions "effort" or "responsibility". It should be stressed that effort, in the present context, does not have the usual meaning of labor or intensity of labor; it is instead a multidimensional variable representing all the factors for which the individual is deemed responsible by society.

In the theoretical part, we address the problem of testing for equality of opportunity in different distributions. However, instead of expressing the opportunity egalitarian goal in its simple form (as in Ok's and Kranich's contributions), hence measuring the degree of inequality in a distribution of opportunities, we employ an indirect approach, where the focus is on the distribution on the consequences of a given distribution of opportunities on some form of individual advantage. More precisely, we propose a unified perspective in order to address two different questions: first, we propose some distributional tests in order to verify the existence of EOp in a given distribution[2]; second, we characterize some distributional conditions in order to rank distributions on the basis of EOp.

In the empirical part of the paper we compare two Italian macro-regions, South and Centre-North, according to equality of opportunity. In the first application we focus on individual earnings, while in the second we consider the distribution of graduation marks among Italian students; in both cases we evaluate the actual distributions of the two macro-regions on the basis of different notions of equality of opportunity.

Our approach can be illustrated by the following argument. Consider a given population and a distribution of a particular form of advantage (income, utility, etc.). The advantage, for each individual, is causally determined by two classes of factors: factors beyond the individual's responsibility (circumstances) and factors for which the individual is responsible. Now partition the population into types, a type being the group of people endowed with the same set of circumstances. If we assume that the individual advantage is determined only by circumstances and effort, and that the distribution of effort is independent of circumstances, then the income distribution observed within a type (i.e., the income distribution conditional to circumstances) can be interpreted as the set of income levels open to individuals with the same circumstances. On the other hand, between-types outcome inequalities (a concept, this, to be precisely defined yet) can be interpreted as reflecting inequalities in opportunities. Hence, roughly speaking, the focus of concern of an analysis of equality of opportunity becomes the measurement of inequality, in terms of the selected outcomes, among different types.

To explain in details the strategy we propose, let us introduce a formal model.

2 The analytical framework

We have a society of individuals. Each individual is completely described by a list of traits, which can be partitioned into two different classes: traits beyond the individual responsibility, represented by a person's set of circumstances O, belonging to a finite set $\Omega = \left\{O^1, ..., O^n\right\}$; and factors for which the individual is fully responsible, *effort* for short, represented by a scalar variable $w \in \Theta \subseteq \Re_+$.

[2] In this exercise we will follow very closely Lefranc et al. (2005, 2006). See also Dardanoni et al (2005).

We assume that all individuals have the same degree of access to the set Θ of possible values of effort; however, the value actually chosen by each individual is unobservable. Income[3] is generated by a function $g : \Omega \times \Theta \to \Re_+$, that assigns individual incomes to combinations of effort and circumstances: $x = g(O, w)$.

It is crucial to notice that by effort in this paper it is meant not only the extent to which a person exerts himself, but all the other background traits of the individual that might affect his success, but that are excluded from the list of circumstances. Clearly, different partitions of the individual traits into circumstances and effort correspond to different notions of equality of opportunity.

We do not know the form of the function g, hence we do not make any assumption about the degree of substituability or complementarity between effort and circumstances; this issue, which is indeed important at an empirical level, is not specified in order to keep the approach as general as possible. We assume, however, that the function g is fixed and it is the same for all individuals.

A society income distribution is represented by a cumulative distribution function $F : \Re_+ \to [0, 1]$. We denote by Ψ the set of income distributions. We can partition any given population into n subpopulations, each representing a class identified by the variable O. For $O^i \in \Omega$, we call "type i" the set of individuals whose set of circumstances is O^i . Within type i there will be a distribution of income with density function $f_i(x)$, cumulative distribution $F_i(x) = \int_0^x f_i(t)\, dt$ and population share denoted by $q_i^F = \frac{N_i^F}{\sum_i^n N_i^F}$, so that $F(x) = \Sigma_{i=1}^n q_i^F F_i(x)$. Moreover, $\mu_i^F = \int_0^\infty x f_i(x)\, dx$ is the average income of type i. To put it differently, $F_i(x)$ is the distribution if income conditional to circumstances O_i.

The distributions of income will differ across types; note however that the distribution function is a characteristic of the type, not of any individual.

In a sense, the distribution $F^i(x)$ represent the set of income levels which can be achieved - by exerting different degrees of effort - starting from the circumstances O_i. That is to say, the distribution $F^i(x)$ is a representation of the opportunity set - expressed in income terms - open to any individual endowed with circumstances O_i. Hence, comparing the opportunity sets of two individuals endowed with circumstances (O_i, O_j) amount to comparing their type relevant income distribution $F_i(x), F_j(x)$. Moreover, evaluating the distribution of opportunity sets[4] among individuals in a society amount to evaluate the set of distributions $\left\{F^1(x), ..., F^n(x)\right\}$.

In this paper we shell exploit this idea[5], in two different ways. First, we

[3] In this section we will use the term "income" to indicate any form of individual achievement.

[4] For a more general approach to ranking distributions of opportunity sets, see Kranich (1996) and Ok (1997), where the individual opportunity sets are represented by abstract sets and inequality and welfare ranking of profiles of such sets are characterized. For a survey, see Fleurbaey and Maniquet (2000) and Peragine (1999).

[5] In the approach sckeched above we puts special emphasis on ex ante inequalities among people endowed with the same social circumstances: accordingly, one interprets the inequality within types as mainly due to differential efforts, and the inequality between types as generated by the different circumstances.

However this is not the only way to look at equality of opportunity. For a different approach, will propose some distributional tests in order to verify the existence of EOp in a given distribution. Then, we shall propose some distributional conditions in order to rank income distributions on the basis of EOp.

2.1 Evaluating opportunity sets

First, we address the problem of ranking opportunity sets. In the framework we have introduced, this amount to ranking the types specific income distributions. We assume that a preference relation over the types specific distributions $\left\{F^1(x), ..., F^n(x)\right\}$ can be represented by an evaluation function V, and we impose some condition on such function.

A first fundamental assumption in a social evaluation exercise concerns the aggregation issue, that in this is a within-type aggregation. A natural way forward would be that of imposing a utilitarian structure. Hence we propose the following additive evaluation function V for a given type i:

$$V_{Fi} = \int_0^z U(x) f^i(x) dx \tag{1}$$

where $U : [0, z] \rightarrow \Re_+$ is the individual utility function assumed to be twice differentiable (almost everywhere) in x, and $f^i(x)$ is the income density function of type i.

We now try to capture the basic intuition beyond the opportunity egalitarian ethics, by restricting the class of utility functions U.

First, we assume that types welfare increases with income, whatever the type:

$$(C.1)\ \frac{dU(x)}{dx} \geq 0, \forall x \in [0, z].$$

Hence condition $C.1$, which is a common **monotonicity** assumption, guarantees that social welfare does not decrease as a result of an income increment.

Next, we assume that our evaluation function is inequality averse. We require **within-type strict inequality aversion**:

$$(C.2)\ \frac{d^2U(x)}{dx^2} < 0, \forall x \in [0, z]$$

Alternatively, we could require our function V to be indifferent to income inequality within the same type, therefore assuming **within-type inequality neutrality**:

$$(C.3)\ \frac{d^2U(x)}{dx^2} = 0, \forall x \in [0, z]$$

This conditions says that, for fixed O^i, that is, when focusing on the group of people having the same opportunity set, the welfare gain resulting from a given total extra income, however distributed, is constant. Hence, a reduction

see Roemer (1998).

in income inequality within a type, which leaves the mean income of the type unchanged, has no welfare effects. This welfare condition implies that the function U^i is affine.

Some comments are in order with respect to the interpretation of the function V and of the conditions so far introduced.

The function V could be interpreted as an individual utility function over a lottery with distribution $F_i(x)$ and support $[0, z]$; in this case the conditions $(C.2)$ and $(C.2)$ are to be interpreted, respectively, as requirements of strict risk aversion and risk neutrality. Hence the function V would represent the individual preferences over the opportunity sets.

Alternatively, the function V could represent the preference relation of a social planner. In this case, the interpretation of the conditions $(C.2)$ and $(C.2)$ are to be interpreted, as in the discussion above, as requirements of inequality aversion and inequality neutrality.

The isomorphism between the theory of risk measurement and individual risk attitude on the one hand, and the theory of inequality measurement and social inequality attitude on the other, allows this double consistent interpretation.

Now we define different class of evaluation functions V : the class of types evaluation functions constructed as in (1) and with utility functions which satisfy conditions (C.1) is denoted by $\mathbf{V}_1$; the class of EFs constructed as in (1) and with utility functions which satisfy conditions (C.1) and (C.2) is denoted by $\mathbf{V}_{12}$; the class of types evaluation functions constructed as in (1) and with utility functions which satisfy conditions (C.1) and (C.3) is denoted by $\mathbf{V}_{13}$.

Thus, we have identified several classes individual utility functions U that implicitly define classes of evaluation functions V. The next step consists in deriving suitable welfare criteria for choosing among opportunity sets, by requiring unanimous agreement among these classes. Hence we have the following definitions of a preference relation over the set of types distribution functions $\left\{F^1(x), ..., F^n(x)\right\}$.

Definition 1 *For all $O_i, O_j \in \Omega$,*

$$F_i \succeq_{V1} F_j \text{ if and only if } V(F_i) \geq V(F_j) \text{ for all } V \in \mathbf{V}_1$$

$$F_i \succeq_{V12} F_j \text{ if and only if } V(F_i) \geq V(F_j) \text{ for all } V \in \mathbf{V}_{12}$$

$$F_i \succeq_{V13} F_j \text{ if and only if } V(F_i) \geq V(F_j) \text{ for all } V \in \mathbf{V}_{13}$$

Standard results in inequality theory allow to identify the distributional conditions corresponding to the welfare criteria above.

Thus, the welfare ranking $\succeq_{W1}$ is equivalent to first order stochastic dominance $(\succeq_{FSD})$. Hence:

Remark 2 *For all distribution functions (F_i, F_j), $F_i \succeq_{V1} F_j$ if and only if*

$$F_i \succeq_{FSD} F_j \iff F_j(x) \geq F_i(x) \text{ for all } x \in [0, z]$$

On the other hand, the welfare ranking $\succeq_{W12}$ is equivalent to second order stochastic dominance $(\succeq_{SSD})$. Hence:

Remark 3 *For all distribution functions* (F_i, F_j), $F_i \succeq_{V12} F_j$ *if and only if*

$$F_i \succeq_{SSD} F_j \iff \int_0^t F_j(x)\,dx \geq \int_0^t F_i(x)]dx \text{ for all } t \in [0, z]$$

As it is well known, second order stochastic dominance is equivalent to Generalized Lorenz dominance (see Shorrocks 1983).

Finally, the welfare ranking $\succeq_{W13}$ is equivalent to higher expected value.

Remark 4 *For all distribution functions* (F_i, F_j), $F_i \succeq_{W13} F_j$ *if and only if*

$$\mu_i \geq \mu_j$$

Given the welfare and distributional criteria discussed above, we can now introduce some criteria in order to test for the existence of EOp.

2.2 Testing for the existence of EOp

So far we have identified several criteria to evaluate opportunity sets.

Now we shall use these criteria in order to test for the existence of equality of opportunity in a distribution of opportunity sets. The idea, first explored by Lefranc et al. (2005) is the following.

In a distribution of opportunity sets $\Omega = \{O^1, ..., O^n\}$ *there is EOp if and only if, for any pair of opportunity sets* O_i, O_j, *neither* O_i *is preferred to* O_j, *nor* O_j *is preferred to* O_i.

The preference relation can be declined in different ways: first, it could be interpreted both as a social preference relation or an individual preference relation over uncertain prospects; second, different relation could be employed. As expected, we will make use of the preference relations characterized in the previous section, to which correspond well defined and suitable distributional conditions.

The fist test is based on the preference relation $\succeq_{V13}$:

Definition 5 ***Weak EOp***. *There is EOp if and only if, for all* $O_i, O_j \in \Omega$, $F_i \not\succeq_{V13} F_j$ *and* $F_j \not\succeq_{V13} F_i$.
Given the remark above, the test can be read in the following way:
There is EOp if and only if, for all $O_i, O_j \in \Omega$,

$$\mu_i = \mu_j$$

A second test is based on the preference relation $\succeq_{V1}$:

Definition 6 ***EOp1 (EOp of the first order)***. *There is EOp if and only if, for all* $O_i, O_j \in \Omega$, $F_i \not\succeq_{V1} F_j$ *and* $F_j \not\succeq_{V1} F_i$.

Given the remark above, the test can be read in the following way:
There is EOp if and only if, for all $O_i, O_j \in \Omega$,

$$F_i \not\succ_{FSD} F_j \quad and \; F_j \not\succ_{FSD} F_i$$

A second test is based on the preference relation $\succeq_{V12}$:

Definition 7 ***EOp2 (EOp of the second order)*** *Weak EOp of the second order. There is EOp if and only if, for all* $O_i, O_j \in \Omega$, $F_i \not\succcurlyeq_{V12} F_j$ *and* $F_j \not\succcurlyeq_{V12} F_i$.
Given the remark above, the test can be read in the following way:
There is EOp if and only if, for all $O_i, O_j \in \Omega$,

$$F_i \not\succ_{SSD} F_j \quad and \; F_j \not\succ_{SSD} F_i$$

Moreover, we could consider a stronger definition of EOp, by requiring that individuals face similar prospects of outcome, regardless of their circumstances

Definition 8 ***Strong EOP****. There is EOp if and only if, for all* $O_i, O_j \in \Omega$,

$$F_i(x) = F_j(x), \; \forall x \in [0, z]$$

The condition above is extremely demanding and will be violated in most case.

These tests allow us to conclude whether there is EOp or not.

Moreover, by considering how may times - *i.e.* for how many pairs of circumstances - we reject the Hypothesis of EOp in distribution F and distinctively in distribution G, a first, rough ranking can be deduced. Moreover, in case there are pairs of circumstances for which the test has different answers in the two distributions, we perform a deeper investigation, by decomposing the SSD dominance condition in mean and inequality (Lorenz) dominance.

2.3 Ranking distributions of opportunity sets

In this section we address the problem of ranking distributions of opportunity set on the basis of EOp. The aim is that of deriving welfare criteria and dominance condition in analogy with the analysis conducted in the previous section; however, now the criteria have to be defined over the set of distributions Ψ.

Again, we assume that a preference relation over Ψ can be represented by a social evaluation function W. A generalization of the evaluation function V discussed in the previous section to the case of income distributions which can be decomposed across homogeneous sub-groups (as the types in our framework), is obtained by aggregating the welfare of each type, weighted by the relevant population share, and using type-specific utility functions. If we opt for an additive aggregation of the types welfare, then we obtain the following utilitarian SEF:

$$W_F = \sum_{i=1}^{n} q_i^F \int_0^z U^i(x) f^i(x) dx \qquad (2)$$

where $U^i(x)$ is the type i specific utility function.

Also in the current scenario, the interpretation of the evaluation function W is ubiquitous. In fact, the SEF proposed above can be interpreted as:

$$W_F = \sum_{i=1}^{n} \Pr\left\{k \in O^i\right\} E\left[U^i(x) \left| k \in O^i \right.\right]$$

where $\Pr\left\{k \in O^i\right\}$ is the probability for an individual k of being endowed with the circumstances O^i and, therefore, of facing the prospect $F_i(x)$; $E\left[U^i(x) \left| k \in O^i \right.\right]$ is the expected utility associated to type i. Hence, our SEF can be expressed in Harsanyi (1955)-type terms, as a weighted sum of the expected utility associated to each type weighted by the probability to belong to that type.

We now try to capture the basic intuition beyond the opportunity egalitarian ethics, by restricting the class of utility functions $< U_1(x), ..., U_n(x) >$. Different value judgments are expressed in this framework by selecting different classes of such functions. These implicitly define welfare rankings[6]. On the other hand, inequality rankings will be typically expressed in terms of conditions on the distribution functions $F_i(x)$.

First, we could impose on the types specific functions U_i the properties $(C.1)$, $(C.2)$ and $(C.3)$ already introduced in the previous section: these conditions are not type-specific, hence we can safely apply them.

In addition, we now formulate some type dependent properties.

First, we formulate a condition expressing the aversion to inequality between the opportunity sets. The condition expressing **between-types income inequality aversion** is the following:

$$(C.4)\ \frac{dU^i(x)}{dx} \geq \frac{dU^{i+1}(x)}{dx}, \forall i \in \{1, ..., n\}, \forall x \in [0, z]$$

which says that the marginal increase in welfare due to an increment of income, is a decreasing function of opportunity. Actually, in case of a continuum of opportunity types i and fully differentiable utility function $U(x, i)$, condition C.4 would require that the cross derivatives be non-positive. Condition $(C.4)$ implies that a transfer of income from a richer to a poorer type (in opportunity terms), at a given income level, is welfare improving[7].

To the properties already introduced we now add the following condition[8]:

[6] For a different approach to opportunity egalitarian welfare orderings, which makes use of rank-dependent social evaluation functions, see Peragine (2002).

[7] Conditions $(C.1)$, $(C.3)$ and $(C.4)$ entail cardinal unit comparability (cf Sen, 1970).

[8] An analog condition is introduced by Jenkins and Lambert (1992) in the context of income inequality in presence of differences in needs, in order to extend the "sequential generalized Lorenz dominance" obtained by Atkinson and Bourguignon (1987) to the case of distributions with different types partitions.

$$(C.5) \qquad \forall i,j \in \{1,...,n\}, U^i(z) = U^j(z)$$

where z is the maximum possible income. By introducing condition C.5 any affine transformation such as, for example, $U^i \rightarrow a^i + bU^i$, now is supposed to be able to affect the results of social comparisons[9]. This requirement is necessary in a context with different types population.

Now we define two classes of social evaluation functions: the class of social evaluation functions constructed as in (2) and with utility functions which satisfy conditions (C.1), (C.2), (C.4) and (C.5), denoted by $\mathbf{W}_{EOP1}$; the class of social evaluation functions constructed as in (2) and with utility functions which satisfy conditions (C.1), (C.4) and (C.5), denoted by $\mathbf{W}_{EOP2}$.

Thus, we have identified two different classes of SEFs that implicitly define welfare rankings of distributions. The next step consists in deriving suitable welfare and distributional conditions by requiring unanimous agreement among these classes. Hence we have the following

Definition 9 *For all F,G*$\in \Psi$,

$$F \succeq_{EOP1} G \text{ if and only if } W(F) \geq W(G) \text{ for all } W \in \mathbf{W}_{EOP1}$$

$$F \succeq_{EOP2} G \text{ if and only if } W(F) \geq W(G) \text{ for all } W \in \mathbf{W}_{EOP2}$$

We now identify a range of tests which, if successful, will ensure welfare dominance for appropriate classes of SEFs; these, in turn, correspond to appropriate classes of the weight functions $U_i(p)$, for these define the SEFs. The aim of the analysis is the following: given a class of utility functions $U_i(x)$ expressing our ethical concerns, we seek conditions, expressed in terms of the distribution functions $F_i(x)$ and $G_i(x)$ and population shares q_i^F and q_i^G, which are necessary and sufficient for welfare dominance according to the criteria defined above.

We first propose the following distributional condition (Peragine, 2004):

Theorem 10 *For all F,G* $\in \Psi$, $F \succeq_{IOP1} G$ *if and only if*

$$\sum_{i=1}^{k} q_F^i \mu_F^i \geq \sum_{i=1}^{k} q_G^i \mu_G^i, \forall k \in \{1,...,n\}$$

This test[10] can be interpreted as a second order stochastic dominance (generalized Lorenz dominance) applied to the distribution of the type means weighted by the relevant population shares: $\left(q_1^F \mu_1^F, ..., q_n^F \mu_n^F\right)$.

This result provides us with a first distributional condition and implicitly suggests a clear, and easy to implement, algorithm: first, partition the population into groups homogeneous with respect to some selected circumstances; second, calculate the arithmetic mean of the income distribution of each of

[9] More precisely, by adding condition $(C.5)$ we pass from cardinal unit comparability to cardinal full comparability (cf Sen, 1970).

[10] This condition was characterized in Peragine (2004).

these groups, and weight such a mean income with the relevant population share; finally, apply to the distributions of these "weighted means" the ordinary generalized Lorenz criterion.

In sum, we are implicitly making the following operations: (i) we evaluate the opportunity set of each type by the weighted mean $q_i^F \mu_i^F$; (ii) we evaluate the distribution of opportunity sets by the Generalized Lorenz criterion.

Second, we obtain the following characterization:

Theorem 11 *For all $F, G \in \Psi$, $F \geq_{IOP2} G$ if and only if*

$$\sum_{i=1}^{k} q_i^F G_i(x) \geq \sum_{i=1}^{k} q_i^G F_i(x), \forall x \in [0, z], \forall k \in (1, ..., n)$$

Proof. See the Appendix. ■

The test above is a sequential distributional test, to be checked in ascending order by k, starting from the lowest type ($k = 1$), then adding the second, then the third, and so on. The condition to be satisfied at each stage is a standard stochastic dominance of the first order. Hence, the theorem characterize a *sequential first order stochastic dominance* condition, where each type distribution is weighted by the relevant population share. This condition dictates the following procedure: take first the lowest type of the two distributions, and check for dominance; then add the second lowest type, then the third lowest type and so on, until all the population is included, performing the dominance check at every stage. We have to perform n different tests, starting form the lowest type, until all types are merged. If these tests are always positive, then we have welfare dominance for the family $\mathbf{W}_{EOP2}$; and the converse is also true.

We therefore implicitly make the following operations: (i) we evaluate the opportunity set of each type by the weighted c.d.f. $q_i^F F_i(x)$; (ii) we evaluate the distribution of opportunity sets by the Generalized Lorenz criterion.

A final remark is in order. We have focused on unanimous preference orderings for classes of opportunity egalitarian social decision makers, rather than on purely (opportunity) inequality criteria. Consequently, the distributional conditions obtained are expressed in terms of means, c.d.f. and generalized Lorenz dominance, rather than simple Lorenz dominance. In the analysis of the results it should be kept in mind that our rankings reflect both distributive and aggregative value judgments.

2.4 A summary

Let us summarize the conditions and the criteria discussed and characterized so far.

Remark 1 As for the test of existence of equality of opportunity, we have proposed the following tests.

$$\boxed{(1)\ \textbf{Weak EOp}} \quad \Rightarrow \quad \boxed{\forall (i, j), \mu_i = \mu_j}$$

$$\boxed{(2)\ \textbf{Strong EOp}} \quad \Rightarrow \quad \boxed{\forall\,(i,j)\,, F_i\,(x) = F_j\,(x)\,,\ \forall x \in [0,z]}$$

$$\boxed{(3)\ \textbf{EOp1}} \quad \Rightarrow \quad \boxed{\forall\,(i,j)\,,\ F_i \nsucc_{FSD} F_j \ \ and\ F_j \nsucc_{FSD} F_i}$$

$$\boxed{(4)\ \textbf{EOp2}} \quad \Rightarrow \quad \boxed{\forall\,(i,j)\,,\ F_i \nsucc_{SSD} F_j \ \ and\ F_j \nsucc_{SSD} F_i}$$

Remark 2 As for the ranking of distributions according to equality of opportunity, we have proposed the following criteria

$$\boxed{(5)\ F \geq_{IOP1} G} \quad \Rightarrow \quad \boxed{\sum_{i=1}^{k} q_i^F \mu_i \geq \sum_{i=1}^{k} q_i^G \mu_i, \forall k}$$

$$\boxed{(6)\ F \geq_{IOP2} G} \quad \Rightarrow \quad \boxed{\sum_{i=1}^{k} q_i^F G_i(x) \geq \sum_{i=1}^{k} q_i^G F_i(x), \forall x, \forall k}$$

3 Empirical application: higher education in Italy

In this section we apply the theoretical framework proposed in the first part of the paper, in order to analyze equality of opportunity in Italy. We examine whether final graduate students outcome and their first salaries distributions are characterized by equality of opportunity. In the empirical application, individual circumstances are represented by parental education. Moreover, our analysis also extends to consider the existence of regional disparities in Italy. To this end, the conditional distributions of two Italian macro-regions, the North-Center and the South are compared and ranked according to different notions of EOp. Our sample is divided in North-Center and South with respect to the geographical location of the University where the students were matriculated.[11] We first present the data and the empirical methodology and then discuss the results.

3.1 Data description

The data come from a survey "Indagine sull'inserimento professionale dei laureati" on the transition from college to work of a representative sample of Italian graduates conducted by the National Statistical Office (Istat). In the last wave available individuals who graduated in 2001 are interviewed three years after completion of the degree, in 2004. The survey covers school curriculum, labour market experience in the three years after graduation, job search activities, household and individual information. The interviewed sample corresponds to about the 17 percent of the population of graduates of 2001. The sample dimension is considerable as the ratio between sampled person and the universe

[11] Here, North-Center comprehends the following regions Piemonte, Lombardia, Veneto, Liguria, Trentino, Friuli, Emilia Romagna, Toscana, Umbria, Marche whereas the South includes Lazio, Abbruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna.

is roughly 1:6. In total the sample consists of about 26.000 individuals (10,766 from Southern and 15,165 from North-Center Universities).[12]

Our analysis particularly focuses on two outcome variables, namely graduation marks and net monthly income, whereas we use parental education to define individual *circumstances*. [13] The variable parental education corresponds to the highest degree obtained in the couple of parents and is divided in four classes.[14]

3.2 Statistical analysis and methodology

Our sample allows to build outcome (*i.e.* final marks and income) distributions conditional on parental education and to perform Step 1 and 2 as described in the previous Section.

Step 1 concerns with assessing equality or stochastic dominance relationships by non-parametric tests of stochastic dominance tests as developed in Davidson and Duclos (2000). We follow Lefranc, Pistolesi and Trannoy (2006) and implement the following empirical procedure. We pursue a separate analysis for the two macro-regions and, for all the possible pairs of circumstances i and j within the same region, we perform four tests independently. We perform:

- test 1 (**Weak EOp**): *i.e.*, the null of equality of the means of the distribution of types i and j;

- test 2 (**Strong EOp**): *i.e.*, the null of equality of the distributions of types i and j;

- test 3 (**EOp1**): *i.e.*, the null of first order stochastic dominance of the distribution of type i over j and viceversa;

- test 4 (**EOp2**): *i.e.*, the null of second-order dominance of the distribution of type i over j and viceversa.[15]

Reminding that the interpretation of the results is only tentative we pursue the following strategy:

- If the null of test (1) or test (2) is not rejected, we conclude that the EOp is satisfied.
- If test (3) or (4) accepts dominance of one distribution over the other but not the other way around, we say that equality of opportunity is violated.
- If test (3) rejects dominance of each distribution over the other we say that equality of opportunity of the first order is supported

[12] The sample was selected by one stage stratification (for sex, university and course) without relacement.

[13] In Italy the final graduation mark ranges from 66 to 110 cum laude, in this analysis the 110 cum laude was simply transformed in 111

[14] The value 1 corresponds to graduates who have at least one of the parents with an elementary school degree; 2 to graduates who have at least one of the parents with a middle school degree; 3 to graduates who have at least one of the parents with a upper secondary degree; and 4 to graduates who have at least one of the parents with a university degree or higher.

[15] In what follows we refer to steps (1), (2), (3) and (4) here listed as test (1), (2), (3) and (4).

- If test (3) and (4) conclude that the two distributions dominate each other we give priority to the results of test (2).

Accordingly, we proceed by comparing the results obtained for the two macro-regions. In particular, if Step 1 gives a different answer for a certain pair of circumstances in the two regions, we decompose the second order stochastic dominance. This decomposition is performed by comparing separately the mean of the distributions (via a test of equality of means) and their Lorenz curves (by testing for such dominance).

The drawback of the characterization of opportunity with such approach is that it does not allow us to rank different situations in which we would reject equality of opportunity. Hence, in case we find evidence of inequality of opportunity we move to the second step of our analysis, that is we look for partial ranking of the distributions of opportunity sets in the two macro-regions. Hence, in order to do so, we rely on the dominance conditions characterized in Theorems 1 and 2. We first verify the existence of the partial ranking $\geq_{IOP1}$ by numerical comparison of the distributions of the type (weighted) means of the two regions [Test (5)]. Next we apply the second criterion ($F \geq_{IOP2} G$) by sequentially testing the following null hypotheses of first order stochastic dominance 1) $q_1^F F_1(x) \leq q_1^G G_1(x)$; 2) $\Sigma_{i=1}^2 \left[q_i^F F_i(x) \leq q_i^G G_i(x)\right]$; 3) $\Sigma_{i=1}^3 \left[q_i^F F_i(x) \leq q_i^G G_i(x)\right]$; 4) $\Sigma_{i=1}^4$ $\left[q_i^F F_i(x) \leq q_i^G G_i(x)\right]$, where F and G are the conditional outcome distributions of North-Centre and South, respectively [Test (6)].

3.3 The results and some concluding remarks

In this section, we first report the tests of equality and stochastic dominance for the outcomes distributions conditional on four classes of parental education. Figures 1 and 2 show the cumulative distribution functions conditional on parental education. A clear ranking of types emerges in the South where the conditional distribution of the graduation final marks of the fourth type dominates over the third, the third dominates over the second, the second dominates over the first. The North-Centre case only shows such a clear pattern for the third and the fourth type, whereas the cumulative distributions of the first and second type seem to often cross each other (see Figure 1). On the other hand, just observing the cumulative distributions of income (Figure 2) we are not able to make such an explicit assessment. This visual ranking is confirmed by the results of the tests of equality and stochastic dominance. Table 1 shows the results of the first test: we note that the null of equivalence of means cannot be rejected at 5% of significance level in more cases in the North-Centre than in the South for the graduation mark. As far as income is concerned, the null cannot be rejected in any case in the North-Centre and is rejected only in a few cases in the South. Turning to the results of stochastic dominance (Table 2), we generally notice that there is more evidence of equality of opportunities in the North-Centre than in the South. In particular, the null of first and second order dominance is rejected only when comparing type three and four of the distribution of income in the South whereas such a distinct pattern cannot be drawn when considering the North-Centre. In this case, no clear cut conclusion can be provided when comparing the distributions of the graduation mark of the first and the second type as well as those of the second and the third type. On the contrary, there is not evidence of dominance of the income distributions of the fourth over the third type (and viceversa) likewise of the third over the second type (and viceversa).

Tables 3 and 4 show the results of the inequality of opportunity comparisons. In table 3, *we first notice that the types mean are generally higher in the South than in the North-Center for the graduation marks distributions. The opposite happens for income distributions.* As for the criterion $\geq_{IOP1}$, table 3 shows a mixed pattern for the final graduation marks: the South dominates the North-Centre in all cases but the third. Hence we cannot conclude for any dominance according to $\geq_{IOP1}$. On the other hand, the distribution of income shows clear evidence of dominance of the North-Centre over the South according to the criterion $\geq_{IOP1}$.

These results are confirmed when testing the criterion $\geq_{IOP2}$, that is the sequential first order stochastic dominance condition (see Table 4). The south dominates the north in almost all the steps of the sequential procedure (all but the third), when considering the marks distributions. The opposite happens with respect with the income distributions: the north dominates the south in all but the first step. That is, in the north there is more equality of opportunity than in the south, when looking at income levels. These figures are consistent with the general view of less intergenerational mobility in the South than in the North. However, the dominance condition required by $\geq_{IOP2}$ is never fully satisfied. The two distributions are not comparable according to $\geq_{IOP2}$.

In the interpretation of the results, it should be reminded that our criteria reflect both distributive and aggregative aspects. Therefore, it is possible that the dominance of the north over the south in income levels is driven by a average effect, rather than a pure inequality effect.

In order to focus on the pure (opportunity) inequality aspect, one could use an inequality index, rather than a welfare ranking approach. Using an inequality index would have a further advantage. In fact, while in this paper we have characterized partial orderings, it would be interesting to investigate complete orderings which are possibly consistent with the rankings characterized here. The idea is that of using an additively decomposable inequality index, then interpreting the inequality *between* types as opportunity inequality, and the inequality *within* types as inequality due to individual responsibility.

This will be the subject of future work.

4 Tables

Figure 1. Graduate final marks cumulative distribution functions

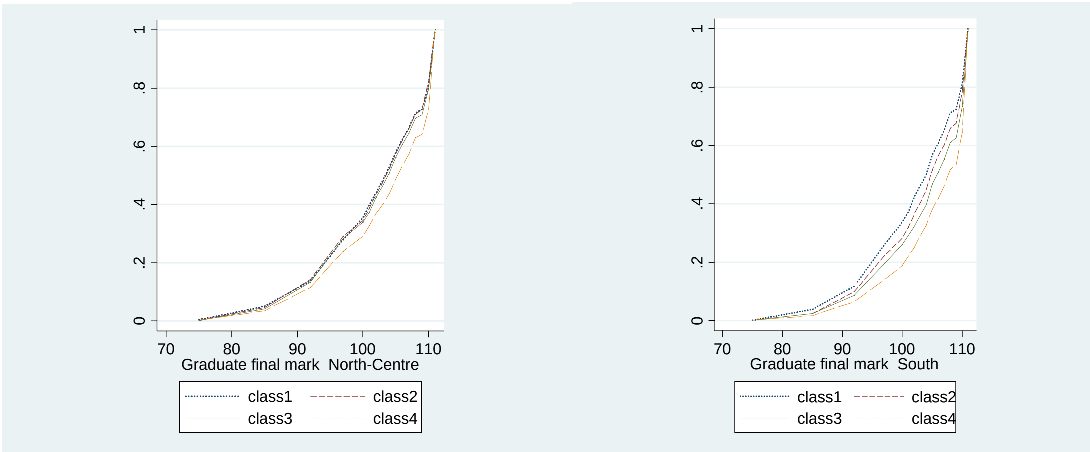

Figure 2. Income cumulative distribution functions

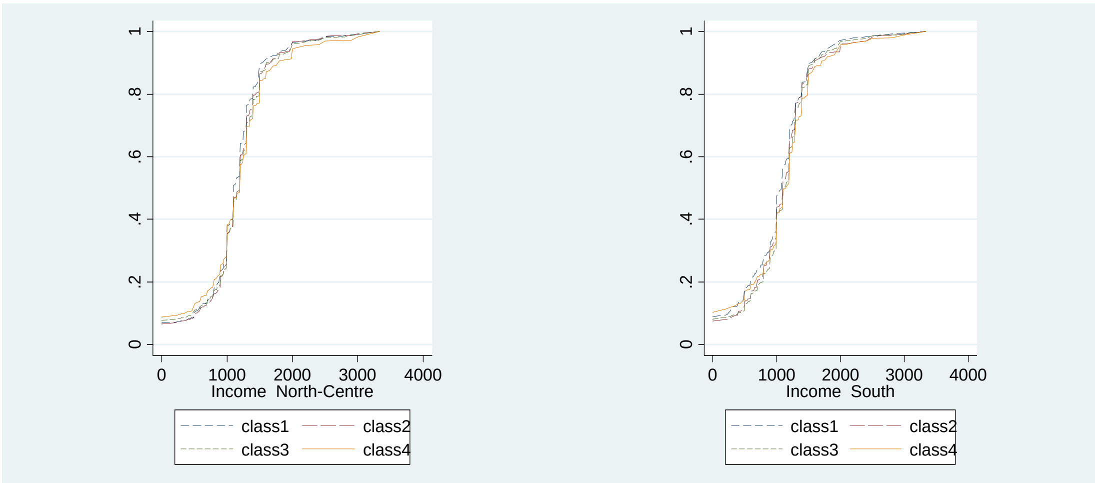

Table 1. Test (1)

Graduation mark								
	North-Centre				South			
	1	2	3	4	1	2	3	4
1	-	$=^*$	$=^*$	$<^*$	-	$=^*$	$<^*$	$<^*$
2		-	$=^*$	$<^*$		-	$<^*$	$<^*$
3			-	$<^*$			-	$<^*$
4				-				-
Income								
1	-	$=^*$	$=^*$	$=^*$	-	$=^*$	$<^*$	$<^*$
2		-	$=^*$	$=^*$		-	$=^*$	$=^*$
3			-	$=^*$			-	$=^*$
4				-				-

Notes: *,**denote 5 and 10% level of significance, respectively. $>$ the mean of the distribution in the row is grater than the distribution in the column; $=$ the means are equal.

Table 2. Test (2), Test(3) and Test(4)

First Order Dominance									Second Order Dominance								
Graduation mark									Graduation mark								
	North-Centre				South					North-Centre				South			
	1	2	3	4	1	2	3	4		1	2	3	4	1	2	3	4
1	-	$\neq^*$	$\neq^*$	$<^*$	-	$<^*$	$<^*$	$<^*$	1	-	$\neq^*$	$\neq^*$	$<^*$	-	$<^*$	$<^*$	$<^*$
2		-	$\neq^{**}$	$<^*$		-	$<^*$	$<^*$	2		-	$\neq^{**}$	$<^*$		-	$<^*$	$<^*$
3			-	$<^*$			-	$<^*$	3			-	$<^*$			-	$<^*$
4				-				-	4				-				-
Income									Income								
	1	2	3	4	1	2	3	4		1	2	3	4	1	2	3	4
1	-	$<^*$	$<^*$	$<^{**}$	-	$<^{**}$	$<^*$	$<^*$	1	-	$<^*$	$<^*$	$<^{**}$	-	$<^{**}$	$<^*$	$<^*$
2		-	$\neq^*$	$\neq^*$		-	$\neq^*$	$\neq^*$	2		-	$\neq^*$	$\neq^*$		-	$\neq^*$	$\neq^*$
3			-	$\neq^*$			-	$\neq^*$	3			-	$\neq^*$			-	$\neq^*$
4				-				-	4				-				-

Notes: see Table 1. $>$ the row dominates the column; $<$ the column dominates the row; $=$

the curves are equal; $\neq$ the curves are different and cannot be ranked

Table 3. Test (5)

Graduation mark		
	North-Centre	South
1	11.41	12.55
1+2	69.75	70.22
1+2+3	221.5	211.8
1+2+3+4	408.95	415.1
Income		
1	132.94	122.67
1+2	813.44	724.32
1+2+3	2567.43	2214.49
1+2+3+4	4504.06	4145.79

Table 4. Test (6)					
		North-Centre			
Graduation mark		1	1+2	1+2+3	1+2+3+4
South	1	$>^*$			
	1+2		$>^*$		
	1+2+3			$<^*$	
	1+2+3+4				$>^*$
Income		1	1+2	1+2+3	1+2+3+4
South	1	$\neq$			
	1+2		$<^*$		
	1+2+3			$<^*$	
	1+2+3+4				$<^*$

5 Appendix

Proof. of Theorem 2

We first state and prove the following Lemma.

Lemma 1 $\sum_{k=1}^{n} v_k w_k \geq 0$ *for all sets of real numbers* $\{v_k\}$ *such that* $v_k \geq v_{k+1} \geq 0$, $\forall k \in \{1, ..., n\}$, *if and only if* $\sum_{i=1}^{k} w_i \geq 0$, $\forall k \in \{1, ..., n\}$.

Proof. of Lemma 1

Applying Abel's decomposition: $\sum_{k=1}^{n} v_k w_k = \sum_{k=1}^{n} (v_k - v_{k+1}) \sum_{i=1}^{k} w_i$. It is obvious that, if $\sum_{i=1}^{k} w_i \geq 0$, $\forall k \in \{1, ..., n\}$, then $\sum_{k=1}^{n} v_k w_k \geq 0$. As for the necessity part, suppose that $\sum_{k=1}^{n} v_k w_k \geq 0$ for all sets of numbers $\{v_k\}$ such that $v_k \geq v_{k+1} \geq 0$, but $\exists\, j \in \{1, ..., n\}$ such that $\sum_{i=1}^{j} w_i < 0$. Consider what happens when $(v_k - v_{k+1}) \searrow 0$, $\forall k \neq j$. We obtain: $\sum_{k=1}^{n} v_k w_k \rightarrow (v_j - v_{j+1}) \sum_{i=1}^{j} w_i < 0$, which is the desired contradiction. $\blacksquare$

We can now prove the theorem, which states that $\Delta W = W(F) - W(G) \geq 0$, for all $W \in \mathbf{W}_{EOP2}$, if and only if

$$\sum_{i=1}^{k} q_i^F G_i(x) \geq \sum_{i=1}^{k} q_i^G F_i(x), \forall x \in [0, z], \forall k \in (1, ..., n).$$

By definition, $\Delta W \geq 0, \forall W \in \mathbf{W}_{EOP2}$, if and only if

$$\sum q_i^F \int_0^z U^i(x) f^i(x) dx - \sum q_i^G \int_0^z U^i(x) g^i(x) dx \geq 0$$

for all the functions U^i satisfying conditions C.1 and C.4. Using integration by parts, we obtain that $\Delta W \geq 0$ if and only if

$$\sum q_F^i \left[U^i(x) F^i(x)\right]_0^z - \sum q_i^F \int_0^z \frac{dU^i}{dx} F^i(x)\, dx - \sum q_i^G \left[U^i(x) G^i(x)\right]_0^z +$$

$$+ \sum q_i^G \int_0^z \frac{dU^i}{dx} G^i(x)\, dx \geq 0.$$

Now we know that $F^i(z) = G^i(z) = 1$, hence the above expression reduces to:

$$\sum \left[q_i^F - q_i^G\right] U^i(z) + \sum \int_0^z \frac{dU^i}{dx} \left[q_i^G G^i(x) - q_i^F F^i(x)\right] dx \geq 0.$$

Now, considering that, by condition (C.4) $U^i(z) = U^j(z)$, and that $\sum\limits_{i=1}^{n} q_i^F = \sum\limits_{i=1}^{n} q_i^G = 1$, we obtain that $\Delta W \geq 0$ if and only if

$$\sum_{i=1}^{n} \int_0^z \frac{dU^i}{dx} \left[q_i^G G^i(x) - q_i^F F^i(x)\right] dx \geq 0$$

or, equivalently,

$$\int_0^z T(x)\, dx \geq 0$$

where

$$T(x) = \sum_{i=1}^{n} \frac{dU^i}{dx} \left[q_i^G G^i(x) - q_i^F F^i(x)\right]$$

Now considering that, by conditions (C.1) and (C.3), $\frac{dU^i(x)}{dx} - \frac{dU^{i+1}(x)}{dx} \geq 0$, we can apply Lemma 1. Hence we obtain that $T(x) \geq 0$, $\forall$ U satisfying C.1 and C.3, if and only if

$$\sum_{i=1}^{k} q_i^F G_i(x) \geq \sum_{i=1}^{k} q_i^G F_i(x), \forall x \in [0, z], \forall k \in (1, ..., n)$$

Clearly, if $T(x) \geq 0$ $\forall x$, then $\int_0^z T(x)\, dx \geq 0, \forall x$, which proves the sufficiency part of the theorem.

As for the necessity part, suppose, for a contradiction, that $\Delta W \geq 0, \forall W \in \mathbf{W}_{EOP2}$ and $\forall F, G \in \Psi$, but $\exists h \in \{1, ..., n\}$ and $\exists I \equiv [a, b] \subseteq [0, z]$ such that $\sum_{i=1}^{h} \left(q_i^F G_i(x) - q_i^G F_i(x)\right) < 0, \forall x \in I$. Then, by Lemma 1, $\exists$ a set of functions $\{U_i : [0, z] \to \Re_+, i \in \{1, ..., n\}\}$ such that $\sum_{i=1}^{n} \frac{dU^i}{dx} \left[q_i^G G^i(x) - q_i^F F^i(x)\right] < 0$ $\forall x \in I$. Thus we have $\Delta W = \int_0^1 T(x)dx$, where $T(x) < 0 \, \forall x \in I$. Clearly, $\int_a^b T(x)dx < 0$. Now I can select a function $T(x)$ (i.e., sets of functions U_i and distributions $F_i(x)$ and $G_i(x)$) such that $T(x) \to 0$ $\forall x \in [0, z] \backslash I$. In this case we obtain that $\Delta W = \int_0^1 T(x)dx \to \int_a^b T(x)dx < 0$. A contradiction.

■

References

[1] Arneson R. (1989) Equality of Opportunity for Welfare. *Philosophical Studies*, 56: 77-93.

[2] Barry B(1991) Chance, choice and justice. In his *Liberty and Justice: Essays in Political Theory*, Volume 2. Oxford: Oxford University Press.

[3] Beach, C. M., and R. Davidson (1983) Distribution-free statistical inference with Lorenz curves and income shares, Review of Economic Studies, pp. 723–735

[4] Betts JR and Roemer JE(1999) Equalizing opportunities through educational finance reform, University of Califormia, Davis, mimeo.

[5] Checchi D. and V. Peragine (2005) Regional disparities and inequality of opportunity: the case of Itlaly. IZA Working Paper No. 1874

[6] Cohen G. A. (1989) On the currency of egalitarian justice. *Ethics* 99, 906-944.

[7] Davidson, R., and J.-Y. Duclos (2000): "Statistical inference for stochastic dominance and for the measurement of poverty and inequality," Econometrica, vol 68, 1435–1464.

[8] Dworkin R. (1981a) What is equality? Part1: Equality of welfare. Part2: Equality of resources. *Philos Public Affairs* 10, 185- 246; 283-345.

[9] Foster, J.E. and A.A. Shneyerov (2000). Path independent inequality measures. Journal of Economic Theory, vol. 91, pp. 199-222.Fleurbaey M and Maniquet F (1999) Compensation and responsibility, in K Arrow, A Sen and K Suzumura (eds), *Handbook of Social Choice and Welfare*, (forthcoming).

[10] Kranich L. (1996) Equitable opportunities: an axiomatic approach. *J Econ Theory* 71, 131-147.

[11] Lambert, P.J. and J.R. Aronson (1993). Inequality decomposition analysis and the Gini coefficient revisited. Economic Journal, vol. 103, pp. 1221-27.

[12] Lefranc A., N. Pisotlesi and A.Trannoy (2005) Inequality of opportunities vs inequality of outcomes: Are Western Societies all alike? ISER Working paper 2005-15

[13] Lefranc A., N. Pisotlesi and A.Trannoy (2006) Equality of Opportunity: definitions and testable conditions with an application to France. mimeo

[14] Ok E.A. (1997) On the measurement of opportunity inequality. *J Econ Theory 77: 300-329.*

[15] Peragine V. (1999) The Distribution and Redistribution of Opportunity. *J. Econ. Surveys*, 13 , pp. 37-69.

[16] Peragine V. (2002) Opportunity egalitarianism and income inequality: the rank-dependent approach. *Mathematical Social Sciences*, vol 44, pp.45-64.

[17] Peragine V. (2004) Measuring and implementing equality of opportunity for income. *Social Choice and welfare*, vol 22, pp.187-210.

[18] Roemer JE (1993) A pragmatic theory of responsibility for the egalitarian planner. Philosophy and Public Affairs, vol. 22, 146-166.

[19] Roemer JE (1998) Equality of opportunity, Cambridge, MA: Harvard University Press.

[20] Shorrocks A. (1983) Ranking income distributions. *Economica*, 50, pp.1-17.