IZA DP No. 1874

Regional Disparities and Inequality of Opportunity: The Case of Italy

Daniele Checchi
Vitorocco Peragine

December 2005

Regional Disparities and Inequality of Opportunity: The Case of Italy

Daniele Checchi
University of Milan and IZA Bonn

Vitorocco Peragine
University of Bari

ABSTRACT

Regional Disparities and Inequality of Opportunity: The Case of Italy

In this paper we provide a new methodology to measure opportunity inequality and to decompose overall inequality in an "ethically offensive" and an "ethically acceptable" part. Moreover, we provide some empirical applications of these new evaluation tools: in the first exercise, we compare the income distributions of South and North of Italy on the basis of a measure of opportunity inequality. Then, we repeat the exercise using the cognitive abilities in a sample of 15-year old students. In both circumstances we find that the less developed regions in the South are characterized by greater incidence of inequality of opportunity.

Corresponding author:

Daniele Checchi
University of Milan
Department of Economics
Via Conservatorio 7
20124 Milan
Italy
Email: daniele.checchi@unimi.it

1 Motivation

Equality of opportunity (EOp) seems to be the prevailing conception of social justice in western liberal societies (Roemer, 1998). Indeed, this idea has been defended and put forward 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) . According to the opportunity egalitarian view, what the principle of justice requires is not equality of individuals' final achievements; once the means or opportunities to reach a valuable outcome have been equally split, which particular opportunity, from those open to her, the individual chooses, 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 individual control. Actually, recent work in the field of axiomatic normative theory[1] has shown that the ideal of EOp can be decomposed into two distinct - and sometimes conflicting - ethical principles: the first, egalitarian in spirit, states that differences in individual achievements which can be unambiguously attributed to differences in factors beyond the individual responsibility, are inequitable and to be compensated by society; this is called the Principle of Compensation. On the other hand, differences of achievements which can be attributed to factors within the personal responsibility are equitable and should not to be compensated; this is the Principle of Responsibility.

Although the large consensus gained by the opportunity egalitarian view, it is a common practice among economist that of evaluating social inequities by looking at the degree of income inequality or, in alternative, at the degree of income poverty in the society.

In our view, the analysis of opportunity inequality in a society, in addition to being interesting *per se*, has also an instrumental value: studying how the opportunity inequality in a given country evolves over time can help to better understand the genesis of the income inequalities; a reduction in the opportunity inequality can indicate a social improvement, *ceteris paribus.* Moreover, studying the differential intensity of opportunity inequality across regional areas, professional categories or even income classes, can give clearer information on the priorities of a redistributive policy.

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 measure opportunity inequality and to decompose overall inequality in an "ethically offensive" and an "ethically acceptable" part. The second goal is to provide an empirical application of these new evaluation tools, and to show how they compare with standard methods of income inequality measurement and opportunity inequality measurement. We believe that our analysis is able to

---

[1]See Bossert 1995, Fleurbaey 1994,1995 and for recent surveys Fleurbaey and Maniquet 2003 and Peragine 1999.

shed some light on aspects otherwise undetected and undetectable by previous distributional analysis.

In the empirical part of the paper we compared 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 cognitive abilities among students; in both cases we evaluate the actual distributions of the two macro-regions on the basis of different notions of equality of opportunity.

2 The approach

The theory of equality of opportunity poses two different economic issues: the first is the design of a public policy intended to implement the EOp view[2]; the second is the problem of measuring the degree of opportunity inequality and ranking social states in terms of equality of opportunity[3].

The focus of the present paper is on this second issue. In a social ordering framework, the principle of compensation can be expressed by stating the *relevance of the circumstances-based inequalities*; while the responsibility principle is expressed by stating the *irrelevance of the effort-based inequalities*. Hence the aim becomes one of seeking inequality orderings which are sensitive with respect to the former, but express neutrality with regard to the latter inequalities.

The model we use is taken from Peragine (2004b). We consider a society of individuals, where each individual income or, in general, achievement, is causally determined by two classes of factors, circumstances and responsibility. We then propose two different partitions of the population. The first is a partition into types, a *type* being a subset of the total population characterized by homogeneity with respect to circumstances. The second is a partition into *tranches,* a *tranche* being the subset of people who have exercised the same level of effort. In this framework, by using a social welfare function approach, Peragine (2002, 2004b) obtains two distinct classes of distributional conditions, typically expressed in terms of Lorenz and generalized Lorenz dominance: according to the first class, one distribution is preferred to another distribution if, and only if, the former dominates the latter at any tranche; whereas the second, weaker criterion, requires dominance of the types-mean distributions. The first class of conditions are close in spirit to the welfare criterion proposed by Roemer (1993), while the second are related to the welfare criterion proposed in van de Gaer

---

[2] A recent literature has explored different aspects and applications of the opportunity egalitarian theory for the design of public policy: for an application within the optimal taxation framework, see Roemer 1998, Roemer et al. 2003, Aaberge et al. 2003; within the fair division literature, see Bossert 1995, Fleurbaey 1995, Fleurbaey and Maniquet 2003; an application to the problem of education financing is contained in Betts and Roemer, 1998.

[3] Clearly the two issues are closely related. For instance, in the same way as the theory of income tax progressivity and redistribution is based on the theory of income inequality, a metric of opportunity inequality is necessary in order to evaluate the impact of an opportunity egalitarian policy (on this see Peragine 2004a).

(1993).

While these dominance conditions have a strong normative appeal, nonetheless they generate partial rankings: often these criteria will fail to rank real world income distributions[4]. Ruiz-Castillo (2000) and Villar (2004) propose complete welfare rankings based on total or average income and Theil's inequality measure[5]. In this paper the focus is on inequality, rather than welfare rankings. Therefore, in order to obtain complete ranking of distributions, we propose an approach based on a member of the entropy family of inequality indices, well-known for its decomposability properties.

3   The analytical framework

We have a society of N 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 $w \in \Theta \subseteq \Re_+$. In this paper we consider as a primitive an ordering $\succ$ over Ω, assumed to be antisymmetric, so that, in general, $O^{i+1} \succ O^i$ for $i \in \{1, ..., n-1\}$. Hence we are able to rank individuals according to their circumstances. An example would be a ranking based on parental education or parental *status*. 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[6] is generated by a function $g : \Omega \times \Theta \rightarrow \Re_+$, 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 sostituitability 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 vector $X = \{x_1, ..., x_N\} \in \Re_+^N$.

Next, we assume that the income function g it is increasing in both circumstances and responsibility:

Assumption 1 *For any* $w \in \Theta$, $g(O^{i+1}, w) > g(O^i, w)$ $\forall i \in \{1, ..., n-1\}$.
Assumption 2 g *is monotonically increasing in* w.

---

[4] In fact, it is the case of our empirical analysis: the dominance conditions characterized in Peragine (2004b) fail to rank the distributions.

[5] See also Goux and Maurin (2003).

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

Let $D^N := \{X \in \mathbb{R}_+^N : \textit{Assumptions 1 and 2 are satisfied}\}$ and let us denote by $D := \bigcup_{N \in \mathbb{N}} D^N$ the set of admissible income distributions.

Although fairly reasonable from a theoretical viewpoint, assumptions 1 and 2 pose an important empirical question: are the empirical distributions we observe belonging to the set D^N? In other words, are assumptions 1 and 2 satisfied in real world distributions? We shall see that our sample distributions do satisfy these crucial assumptions.

We now propose two different partitions of the total population. First, given the ordering defined over Ω, we can partition any 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. Letting N_i^X be the number of people in type i of distribution X, such that $\Sigma_{i=1}^n N_i^X = N_X$, we denote by $\mathbf{x}_i = \{x_i^1, ..., x_i^{N_i^X}\} \in \Re_+^{N_i^X}$ the type i income distribution. Thus the income profile $X \in \Re_+^{N_X}$ can be written as

$$X = \{\mathbf{x}_1, ..., \mathbf{x}_i, ..., \mathbf{x}_n\} \in \Re_+^{N_X}. \tag{1}$$

The second partition is based on the responsibility w: for all degrees of responsibility $w \in \Theta$, we call *tranche w* the set of individuals whose responsibility level is equal to w. As we are considering the case of non-observability of the responsibility variable, we need to deduce the *degree* of responsibility exercised from some observable behavior. More precisely, we need a *proxy* in order to measure in an ordinal sense and to compare the effort of different individuals. The idea is the following.

In each type i there will be a distribution of effort; given the circumstances, which are common for all the individuals in the same type, and the function g, there will ensue a distribution of income. These distributions will differ across types; note however that the distribution function is a characteristic of the type, not of any individual. Equality of opportunity holds that individuals should not be held responsible for their circumstances, that is, for their type. In constructing an inter-type- comparable measure of effort, we must therefore take account of the fact that some individuals come from types with 'good' distributions of effort, and some from types with 'poor' distributions. Roemer therefore suggests to take the inter-type comparable measure of effort to be the quantile of the effort distribution in his type at which an individual sits; this, given the monotonicity of the income function, will correspond to the quantile in the income distribution of the type. We say that all individuals at the *pth* quantile of their income distributions, across types, have tried equally hard. Using the quantile measure of effort sterilizes out the 'good' or 'bad' nature of the distribution of effort in the type.

Thus, considering types $1, ..., n$, we define the tranche p in population N as the subset of individuals whose incomes are at the pth rank of their respective type income distributions. We have m quantiles, denoted by $p \in \{1, ..., m\}$. Working in a discrete framework, we need to assume that, for all $i \in \{1, ..., n\}$, N_i^X is divisible by m. Considering a given type i, with the relevant income vector $\mathbf{x}_i \in \Re_+^{N_i^X}$, let us denote the vector of incomes in quantile p of type i

by $\chi_{i,p} \in \Re^{\frac{N_i^X}{m}}$. Analogously to the types-partition introduced before, we can now construct a disjoint exhaustive partition of the population into tranches. If $\frac{N_i^X}{m}$ is the number of people in quantile p of type i, then $\Sigma_{i=1}^n \frac{N_i^X}{m} = \frac{N}{m}$ is the number of individuals in any tranche p. Thus, the subset of the population, identified by type, who have exercised responsibility p, is represented by the following *tranche p* vector, $\boldsymbol{\chi}_p = \{\chi_{1,p}, ..., \chi_{n,p}\} \in \Re_+^{\frac{N}{m}}$, with the relevant mean income denoted by μ_p^X. Accordingly, the income profile $X \in \Re_+^N$ can now also be written as

$$X = \left\{ \boldsymbol{\chi}_1, ..., \boldsymbol{\chi}_p, ..., \boldsymbol{\chi}_m \right\} \in \Re_+^N. \tag{2}$$

Notice that, given D^N, the set of feasible income profiles, the types populations N_i, $i \in \{1, ..., n\}$, are not predetermined; thus we consider income distributions with different types partitions.

Now compare the formulation in (1) with that given in (2). They dictate two different approaches to measure opportunity inequality. The first approach focus on the types income distributions and is based on the following definition of equality of opportunity:

Definition 1 *The types approach. There is EOp if and only if the expected value of income is the same, regardless of the type.*

Thus, the types approach puts special emphasis on ex ante inequalities between 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.

The second approach instead focus on the actual distributions across individual at the same level of effort, and is based on the following definition of equality of opportunity:

Definition 2 *The tranches approach. There is EOp if and only if all those who exerted the same degree of effort have the same chances of achieving the objective, regardless of the type.*

Thus, the tranche approach emphasizes inequalities within effort groups: an aggregation of such tranches inequalities will provide us with a measure of the overall opportunity inequality. On the other hand, differences between the tranches are interpreted as different rewards due to people autonomous choice, and are not considered as unfair.

Both approaches look as plausible evaluation strategies consistent with the EOp principle. Moreover, as shown in the example below, they can provide us with very different rankings of society in terms of EOp. Therefore we shall explore both the types and tranche approaches; within each procedure, we shall try and decompose overall inequality in an "ethically offensive" and an "ethically acceptable" part.

3.1 An example

Let us consider a society where individuals' incomes depend on their parents' educational attainments and their level of effort. Suppose the society agree to consider inequalities due to parental education as unjust, while regard inequalities due to individual effort as just. Society agrees also on a ranking of parental educational attainments. There are three levels of parental educational attainments (N: No school; P: primary, S: secondary), and two levels of effort (Low, High). Hence the society can be represented by a 3×2 matrix. Consider the following example:

Society 1

parents education (c)/effort level (e)	Low	High
No education	10	20
Primary school	20	30
Secondary school	30	40

Society 2

parents education (c)/effort level (e)	Low	High
No education	10	20
Primary school	10	40
Secondary school	10	60

The rows correspond to the types; the columns to the tranches. The two societies share a basic feature: *ceteris paribus*, it is better to have a better educated parent, and to exert a higher level of effort. However, the two societies differ in the extent to which effort and social circumstances are combined to produce income. More precisely, we assume that the effort e is a dummy variable: $e_{low} = 0; e_{high} = 1$; and that the circumstances are represented by a variable c such that: $c = 1$ if no education; $c = 2$ if primary school ; $c = 3$ if secondary school.

Society 1 is generated by the function

$$x = f_1\left(c, e\right) = 10(c + e)$$

While society 2 is generated by the function

$$x = f_2\left(c, e\right) = 10 + 10\left(e\left(c - e\right) + \left(c \times e\right)\right)$$

Hence, society 1 corresponds to the case where the income function is additively separable in circumstances and effort: the two factors are perfect substitutes. In contrast, society 2 corresponds to the case where having a well-educated parent is an advantage for the high effort individuals only.

An inspection of the tables above shows that, according to the types approach, society 1 and society 2 exhibit the same level of opportunity inequality (as they have the same types means); on the other hand, according to the

tranche approach society 1 shows more effort inequality than society 2^7. We have proved that the types and the tranche approach may generate different rankings of distributions.

4 Measuring and decomposing opportunity inequality

4.1 The tranches approach

In this section we focus on the following representation of an opportunity-responsibility-income distribution. We have the income profile
$X = \{\boldsymbol{\chi}_1, ..., \boldsymbol{\chi}_p, ..., \boldsymbol{\chi}_m\}$, where the tranche p vector, $\forall p \in \{1, ..., m\}$, is defined as: $\boldsymbol{\chi}_p = \{\chi_{1,p}, ..., \chi_{n,p}\} \in \Re_+^{\frac{N}{m}}$.

Consider the set of incomes within a given quantile p of any type i, denoted by $\chi_{i,p}$. Within $\chi_{i,p}$ there will be different income levels; however, as we are taking the quantile as a proxy for the unobservable responsibility, all individuals with income in $\chi_{i,p}$ are considered as having exercised the same degree of responsibility; no matter the slight differences among their incomes. That is to say, any income inequality within $\chi_{i,p}$ is not explained by our model, and it is considered as normatively irrelevant. Therefore, starting from an income profile $X = \{\boldsymbol{\chi}_1, ..., \boldsymbol{\chi}_p, ..., \boldsymbol{\chi}_m\} \in D^N$ we can generate an artificial distribution $X^S \in D^N$ by substituting, to each income $x \in \chi_{i,p}$, for all $i \in \{1, ..., n\}$ and for all $p \in \{1, ..., m\}$, the arithmetic mean of the vector $\chi_{i,p}$, denoted by $\mu_{i,p}^X$. Hence, with this transformation, denoting by $1_{i,m}$ the unit vector of length $\frac{N_i^X}{m}$, we obtain the new "smoothed" vector[8] $\chi_{i,p}^S = \{\mu_{i,p}^X 1_{i,m}\} \in D^{\frac{N_i^X}{m}}$. Accordingly, the "smoothed" tranche p vector, for all $p \in \{1, ..., m\}$, can now be defined as $\boldsymbol{\chi}_p^S = \{\chi_{1,p}^S, ..., \chi_{i,p}^S, ..., \chi_{n,p}^S\} \in D^{\frac{N}{m}}$ and the smoothed income profile X^S can be defined as:

$$X^S = \left(\boldsymbol{\chi}_1^S, ..., \boldsymbol{\chi}_p^S, ..., \boldsymbol{\chi}_m^S\right) \in D^N. \tag{1}$$

In this section we are interested in finding criteria to ranking distributions to which the above defined smoothing transformation has been applied. Hence, for all $X, Y \in D^N$, we denote by $\boldsymbol{\chi}_p^S, \boldsymbol{\nu}_p^S$ and by X^S, Y^S the relevant smoothed vectors. For simplicity, we will refer to these simply as the tranche and the population vectors respectively.

---

[7]To see this, notice the following: (*i*) the income vector corresponding to column Low in society 1 shows the same level of inequality as the column High in society 2, for any scale invariant measure of inequality; (*ii*) the column High in society 1 show more inequality than the column Low in society 2 - which in fact does not shows any inequality at all. Hence, for any scale invariant, additively decomposable inequality meaure, and for any weithing scheme used to compute the within tranche inequality term, society 2 will be declared more opportunity equal than society 1 according to the tranche approach.

[8]Smoothing transformations analogous to the one introduce here could be formulated by using any other "representative income", such as the geometric or harmonic mean or the equally distributed equivalent income (on this see Foster and Shneyerov 2000). Here we use the arithmetic mean because we want to preserve the same total income.

Notice that with this transformation all the *unexplained* inequality in our model is canceled out: all the inequality we observe is only attributable to the circumstances O_i or to the effort level w. Clearly, an empirical question here arises: how important is the transformation $X \rightarrow X^S$? As we'll see in the empirical part of the paper, this smoothing transformation has a fairly acceptable impact over the original distribution.

We want to distinguish the overall inequality observed in the income vector $X^S \in D^N$ into: inequality due to opportunity inequality and inequality due to individual responsibility. Now, according to the assumptions introduced in section 2, we can say that the within-tranches inequality is to be interpreted as income inequality due to opportunity inequality; on the other hand, the between tranches inequality surely reflects inequality in the exercise of individual responsibility.

Consider the three following reference vectors:

$$(a) \quad X^S \quad = \quad \left(\boldsymbol{\chi}_1^S, ..., \boldsymbol{\chi}_p^S, ..., \boldsymbol{\chi}_m^S \right) \in D^N$$

$$(b) \quad X_B^S \quad = \quad \left(\mu_{\boldsymbol{\chi}_1^S} 1_{\frac{N}{m}}, ..., \mu_{\boldsymbol{\chi}_p^S} 1_{\frac{N}{m}}, ..., \mu_{\boldsymbol{\chi}_m^S} 1_{\frac{N}{m}} \right) \in D^N$$

$$(c) \quad X_W^S \quad = \quad \left(\tilde{\mathbf{x}}_1^S, ..., \tilde{\mathbf{x}}_p^S, ..., \tilde{\mathbf{x}}_m^S \right) \in D^N$$

where $\mu_{\boldsymbol{\chi}_p^S}$ is the mean of the tranche p income vector, $1_{\frac{N}{m}}$ is the unit vector of length $\frac{N}{m}$, and $\tilde{\mathbf{x}}_p, \forall p \in \{1, ..., m\}$ is obtained by rescaling each income $\mu_{i,p}^X$ in the following way:

$$\forall i \in \{1, ..., n\}, \forall p \in \{1, ..., m\}, \mu_{i,p}^X \rightarrow \frac{\mu_X}{\mu_{\boldsymbol{\chi}_p^S}} \mu_{i,p}^X.$$

The distribution X^S is the overall income vector; X_B^S is a hypothetical smoothed distribution in which each person's income is replaced with the mean income of the tranche to which he or she belongs. This smoothing process removes all inequality within the tranches; X_W^S is a standardized distribution obtained by proportionally scaling each tranche distribution until it has the same mean as the overall distribution. Standardization suppresses between-tranche inequality while leaving tranche inequality levels unaltered.

The interpretation in the current context is as follows. The artificial vector X_B^S is the distribution obtained by eliminating opportunity inequality. An inequality index applied to this distribution captures only and fully the inequality due to individual responsibility. On the other hand, by rescaling all tranches distributions until all tranches have the same mean income, we are left with an income vector $\left(X_W^S \right)$ in which the only inequality present is the within-tranches inequality: an inequality index applied to this distribution captures only and fully the income inequality due to opportunity inequality.

Therefore, considering any two income distributions $X, Y \in D^N$, and a given measure of inequality $I : D^N \rightarrow \Re$, we can say that *distribution X exhibits a lower degree of opportunity inequality than distribution Y if and only if*

$$I \left(X_W^S \right) < I \left(Y_W^S \right).$$

Moreover, we can use a decomposable measure of inequality[9] and have a decomposition as follows:

$$I\left(X^S\right) = I\left(X_B^S\right) + I\left(X_W^S\right)$$

By expressing $I\left(X_W^S\right)$ as a residual, we have the following decomposition:

$$I\left(X_W^S\right) = I\left(X^S\right) - I\left(X_B^S\right)$$

which is to be interpreted as: Opportunity inequality = Total inequality - Effort inequality.

4.2   The types approach

In this section we present an analysis similar to the one presented in the previous section, but focusing now on the types approach.

Consider now the following reference vectors:

$$
\begin{aligned}
(a') \ X &= \left(\mathbf{x}_1, ..., \mathbf{x}_i, ..., \mathbf{x}_n\right) \in D^N \\
(b') \ X_B &= \left(\mu_{\mathbf{x}} 1_{N_1}, ..., \mu_{\mathbf{x}_i} 1_{N_i}, ..., \mu_{\mathbf{x}_n} 1_{N_n}\right) \in D^N \\
(c') \ X_W &= \left(\tilde{\mathbf{x}}_1, ..., \tilde{\mathbf{x}}_i, ..., \tilde{\mathbf{x}}_n\right) \in D^N
\end{aligned}
$$

where we recall that $\mu_{\mathbf{x}_i}$ is the mean of the type i income vector, and $\tilde{\mathbf{x}}_i, \forall i \in \{1, ..., n\}$ is obtained by rescaling each type i income in the following way:

$$\forall i \in \{1, ..., n\}, \forall h \in \{1, ..., N_i\}, x_i^h \rightarrow \frac{\mu_X}{\mu_{\mathbf{x}_i}} x_i^h$$

In this case, (a') is the overall income vector, (b') eliminates within-types inequality, and (c') eliminates between-types inequality.

The interpretation is as follows. By measuring the inequality in the artificial vector X_B, obtained by replacing each income with its type mean income $\mu_{\mathbf{x}_{i,.}}$, we capture only and fully the between-types inequality, which, in turn, in the types approach reflects the opportunity inequality. On the other hand, by rescaling all types distributions until all types have the same mean income, we are left with an income vector (X_W) in which the only inequality present is the within-types inequality, to be interpreted as inequality due to individual

---

[9] To obtain a decomposition as the one proposed in the text - which holds for a general class of representative incomes, not only the arithmetic mean - one needs to use a "path independent" inequality measure as defined and characterized by Foster and Shneyrov (2000). In the empirical application we'll use the *mean logarithmic deviation (MLD)*, which is the only index which has a path-independent decomposition using the arithmetic mean as the representative income. For a distribution X, with mean μ_X and size N, the MLD is defined as:

$$MLD\left(X\right) = \frac{1}{N} \sum_{i=1}^{N} \ln \frac{\mu_X}{x_i}$$

responsibility. Therefore, considering any two income distributions $X, Y \in D^N$, and a given measure of inequality I, we say that *distribution X exhibits a lower degree of opportunity inequality than distribution Y if and only if*

$$I\left(X_B\right) < I\left(Y_B\right).$$

Just as in the previous section, we can use a "path independent" measure of inequality I, and we obtain the following decomposition:

$$I\left(X_B\right) = I\left(X\right) - I\left(X_W\right)$$

which is interpreted as: Opportunity inequality = Total income inequality - Responsibility inequality.

5   The empirical analysis

Any empirical application of the theory described in the previous sections requires the identification of the concepts of individual *objective* and of the relevant list of *circumstances*. In this paper, we present two applications of the proposed decomposition of inequality, which differ in terms of the individual objective: the first application deals with actual earnings, and is directly connected to the human capital approach; the second one considers the distribution of cognitive abilities among students, and therefore refers to inequalities existing before the access to the market. However, in both applications we are concerned with the same set of circumstances, represented by the family background, which is in turn measured by the level of parents' education.

5.1   The first application: the distribution of labour earnings

In the first application we choose income as the selected objective, and we study the effect of the family background on individual earnings. Now, we could identify the following channels through which parents affect the income earning capacity of their children (see Dardanoni *et al. 2004* - check in reference the year is 1993):

a) provision of social connections which are relevant in the labour market;

b) formation of beliefs and skills in children, through family culture and investment;

c) genetic transmission of native ability;

d) instillation of preferences and aspirations.

Clearly, various notions of equality of opportunity correspond to different choices of which of these channels are to be regarded as circumstances. Now, also on the basis of the data available, we declare that only factors a) and b) count as circumstances. Factors a) and b) are proxied, in our analysis, by the level of parents' education. This amounts to say that any other factors, as native

ability, race, luck, and so on, are implicitly classified as within the sphere of individual responsibility. If, even under this extremely conservative view of what constitutes responsibility, our society exhibits a certain degree of inequality of opportunity for income, then we are legitimate to conclude that a "minimal" compensatory policy should be predicated on family characteristics of the individuals. Put it differently, our analysis is able to identify only the lower bound of opportunity inequality. Consequently, the extent of the compensatory intervention dictated by such a minimalist view of individual responsibility should be viewed only as a lower bound for a global redistributive program.

We draw data on individual annual earnings and family background from the Survey on the Income and Wealth of Italian Households (SHIW), waves 1993, 1995, 1998 and 2000. Conducted any other year by the Bank of Italy, the survey collects data on representative samples of approximately ten thousands of Italian households each wave. Respondents provide information on parent's education and occupation, their own educational achievement and other demographic variables[10]. We have restricted the sample to observations with positive earnings from dependent employment (given the low reliability of self-decleared incomes from self-employed). The survey asks for net earnings; based on existing fiscal laws and information about family composition, we have reconstructed the gross earnings.

Family backgrounds is measured by the highest educational attainment in the couple of parents. Local labour markets are taken into account by splitting the sample into Northern regions and Central-Southern regions. Overall we consider 18024 observations[11] (see table 1). In accordance with the literature on determinants of earnings (see the review in Card 1999), since individual earnings varies according to many other observable characteristics that we want to ignore (like gender or age), we have regressed actual gross earnings on some controls (gender, experience and experienced squared[12], survey years) and we have taken the residuals from this regression. This variable represents earnings obtained by an average person ("average" gender and experience) measured in Italian liras at 2000 prices. Descriptive statistics are reported in table 2. It is easy to note that earnings are increasing in the level of parental background, and they are also higher in the North than in the South by an average of 12% (but the gap is decreasing with the parental background). By "clearing" the data from individual heterogeneity based on gender and experience we reduce measured inequality, as it can be appreciated by comparing the first and the second columns of table 3, reporting alternative measures of inequality.

But data are still "contaminated" by unobservable components (like ability

---

[10] We could not use surveys collected before 1993, because information about parental background was absent. The English version of the questionnaire, data and survey documentation can be downloaded from the Bank of Italy web site: http://www.bancaditalia.it/statistiche/ibf.

[11] The Bank of Italy survey contains a panel component of approximately one-third of the interviewed. We have retained this component in order to save observations. This implies that the same individual may appear more than once in our sample.

[12] Due to the lack of information about actual experience in the survey, we have calculated potential experience as $(age - years\ of\ education - 6)$.

or luck) that may confound the analysis of inequality. Under the maintained assumption that individuals at the same percentile of earning distribution have exerted the same degree of effort, we have partitioned the earnings distribution (conditional on parental background and macroregion of residence) into 20 quantiles, and we have replaced individual (predicted) income with the average income of each cell (20 quantiles $\times$ 5 types of background $\times$ 2 macroregions). As it can be judged by comparing the second and third column of table 3, the reduction in measured earnings inequality is rather limited.

We are now in the condition of analyzing two earnings distributions, in the North and in the Centre-South, according to two characteristics, parental education and individual effort, having dispensed for individual heterogeneity and unobservable components.

The data satisfy our initial assumptions: rank ordering within each quantile is respected (implicating that effectively the income function is strictly increasing in opportunity) for both regions (see table 4). This is rather unsurprising, given the high impact of parental education on attained education. However, if all the effect of parental background would pass onto children through parental education, we should observe an insignificant impact once we control for attained education. This is not the case, when we look at table 7. In the first two columns we have regressed individual (log)earnings on gender, work experience and parental education. In third and fourth columns we introduce the educational attainment of the interviewed, and the impact of parental education declines but remains significant, especially in Southern regions.

The main results of our analysis are summarized in tables 5 and 6 and in figure 1. According to the tranche approach (see table 5), hence when inequality of income attributable to inequality of opportunity is calculated by percentile over types and macroregion (within tranche inequality), we obtain that inequality of opportunity is equal to 0.0079 in Center-South and 0.0038 in North. Therefore the inequality of opportunity is double in Southern regions; moreover, the gap is concentrated in lowest percentiles (with reversed situation in upper percentiles - see figure 1). When we consider the inequality of income attributable to inequality of responsibility/effort, we calculate the inequality index between tranches for each region, obtaining respectively a measure of 0.073 in Center-South and 0.061 in North. Since effort inequality is in the same order of magnitude between regions, whereas inequality of opportunity was double in the Center-South, as a consequence inequality of opportunity accounts for 1/10 of earnings inequality in Southern regions and just 1/20 in Northern ones. This is consistent with finding less mobility in South than in the North of Italy, even if it exhibit a converging trend.[13] As long as parental education reduces its impact onto educational attainment of younger cohorts, the benefits in terms of inequality reduction will accrue more to the South than to the North. Analogous size effects are found when we follow a "type" approach (see table 6): inequality of opportunity is almost double in Southern regions, and its incidence is lower in Northern regions.

---

[13] See Checchi and Dardanoni (2002).

13

Not surprisingly, our results are consistent with standard regression analysis (as reported in table 7). When parental education is measured by dummy variables, we find that they bear a coefficient which is almost double in the South than in the North when we exclude the educational attainment of the individual (first two columns), whereas it goes up to three-four times when the educational attainment is included (third and fourth column). This regression helps us to understand why the inequality of opportunity is higher in the South: while most of the parental background exert its effect through favouring the educational attainment of the children in the North, it keeps on playing a role independently from education in the South. This could represent the impact that family networking play in finding good jobs. But it could also be related to the social capital: since parental background is correlated with the average educational attainment of the environment, the positive effect of parental education (once we control for educational attainment of the child - see fourth column in table 7) could also represent a sort of "peer effect" during the educational career, that later on manifests itself in higher earnings[14].

We have also performed some robustness checks, by considering the role of migration in shaping this outcome. Despite not having a direct measure on whether an interviewed is a migrant, we can observe the region of birth and the region of actual residence. By defining migrant a person born in one macroregion and living in the other (thus excluding local migration), we have 2.0% in the South and 16.9% in the North. This category is a mixed one, since includes individual who migrated following their parents (which therefore contributed to the inequality of opportunity) and individuals who migrated by their own (thus contributing to the inequality of effort). By excluding these individuals from the analysis, we observe a decline of inequality mainly in the North, more on the inequality of opportunity than on the inequality of effort: the former passes from 0.00389 to 0.00287, while the latter declines from 0.0610 to 0.0548.

5.2   The second application: the distribution of cognitive abilities

In a similar way to earnings, we have analyzed the effect of family background on the distribution of cognitive abilities. The analysis is based on data from the program PISA (Program for International Student Assessment), a program sponsored by OECD and conducted every three years in more than 30 countries. This survey was conducted for the first year in 2000 to assess reading ability of 15-year-old students. Students are required to sit a 3-hour test, which is meant to assess *reading literacy*[15]; the test score is standardized to an international mean of 500 and a standard deviation of 100. Among the most striking results of this survey there is the wide school heterogeneity within Italy. In 2000, the country average score in a sample of 4946 students was 491. But there were

---

[14] See the analysis in the next section.

[15] "Reading literacy – performing different kinds of reading tasks, such as forming a broad general understanding retrieving specific information, developing an interpretation or reflecting on the content or form of the text." (OECD-PISA 2000, p.13).

large differences across five macro-regions, even conditioning by school types: the average score in high schools in North-Western regions was 572 against 503 in the South; the corresponding values for vocational schools were 473 and 398, respectively. But the region of living and mostly the type of secondary school attended are out of the control of a 15-year-old youngster, and could reasonably be considered as affecting the opportunity set.[16] As before, we have summarized family background through the educational attainment of parents, which is slightly more detailed than in the SHIW survey, distinguishing between different types of secondary school attended by parents.[17] Looking at table 8, we notice that students' abilities are increasing in parental education, and are constantly higher in Northern regions. In order to reduce the heterogeneity of the test scores, we have controlled for gender and age, obtaining the presumed ability of a representative student (see table 9). The replacement of actual ability with theoretically predicted one leaves the distribution practically unaltered, as it can easily checked in table 10. Even replacing individual ability with the mean ability of a representative individual with the same parental background, living in the same macro-area and exerting the same amount of effort, does not reduce measured inequality of abilities.

Once more, we introduce the assumption that individuals at the same percentile of ability distribution have exerted the same degree of effort. This assumption may be questioned in the context of student surveys. One could obtain a very high test score because either she is a natural genius, or because she was lucky enough to meet extraordinary teachers in the course of her life, or even because she is the daughter of school professors. In all cases her performance has nothing to do with her effort. However, had not she exerted some effort, other things constant her performance would have been lower. Under the present assumption about effort distribution, we now look at the ability distributions in the two regions (see table 11). Our analysis suffer from the limited number of observations in many cells, which has required the reduction of the number of quantiles from 20 to 10, and also the grouping of illiterate parents with parents with primary education. This explain why we observe three cases of rank reversal of the average ability defined over type/decile/region.

As before, the main result of our proposed decomposition is reported in

---

[16] We are implicitly assuming that the type of school attended is independent of the effort exerted by the student. Two justifications are in order. The first one refers to the institutional nature of the Italian educational system, where the type of upper secondary school is made at the age of 13, by the family after some counselling from the school professors. The second one is that when we model the choice of the secondary school attended in the PISA dataset, we find that parental education and past school performance (whether been retarded in the past) are both correlated with the current school attendance. For both reasons we feel justified in considering the type of school as the result of a direct choice of parents. One could object that a lazy student would have obtained a scarce performance in the previous school level (lower secondary), would have been oriented to a "light" secondary school, and therefore the choice of the secondary school would reflect *effort* instead of *background*.

[17] Due to missing information on parental education, the actual sample reduces to 4769 students. Remember that this sample is non representative of the corresponding age cohort, since upper secondary school is not compulsory in Italy, and typically students from worst background tend not to enrol.

tables 12 (tranche approach) and 13 (type approach). The inequality of opportunities is almost the double in the Southern region when compared to Northern one, and their incidence on total inequality is higher (respectively 7.2% versus 5.3%). This is true whatever is the way in which we measure the inequality of opportunity, either the "tranches" or the "types" approach. The inequality of opportunity affect disproportionately the students in the bottom tail of the score distribution, as shown by figure 2.

Why are students in Southern schools more unequal in terms of cognitive abilities ? Regression analysis, reported in table 14, help us to shed some light on this issue. As expected, the measured ability of student is correlated with parental education (column 1 and 2), but the impact is $3 - 5$ percentage points higher in the South.[18] If the ability acquired after 10 years of formal schooling is the outcome of an educational production function, where parental education, school resources and environmental effects all play a role, and if family resources and school resources are substitutes in this production function[19], then a higher impact of parental education is a signal of *lower* school resources invested in this region.[20] When we introduce the type of secondary school attended (column 3 and 4), we observe that the impact of parental education almost disappears in the Northern region, while is remains significant in the South. This could indicate that, despite students being sorted according to their family backgrounds different tracks, within each track the family of origin still matter wherever the quality of education provided by the local school remains limited.

6 Concluding remarks

The philosophy of equality of opportunity suggests that social and economic inequalities due to factors beyond the individual responsibility are inequitable and to be compensated by society; whereas inequalities due to personal responsibility are equitable and not to be compensated. Therefore, according to the opportunity egalitarian conception, in order to assess the equitability of a state of affairs one has to distinguish, in a given distribution of outcomes, the inequalities due to personal responsibility as opposed to the inequalities due to non responsible factors or opportunities. In this paper we have provided a theoretically sound methodology to measure opportunity inequality and to decompose overall inequality in an "ethically offensive" and an "ethically acceptable" part. Moreover, we have provided an empirical application of these new evaluation tools, and shown how they compare with standard methods of income inequality measurement and opportunity inequality measurement.

---

[18] For comparability reasons with table 7, we have chosen to use the log of the test score instead of the actual level.

[19] See Checchi 2004 for an estimation of such a production function for Italy. See Brunello and Checchi 2003 for the discussion of the substitutability/complementarity issue in the case of Italy.

[20] This is consistent with what is found in the literature: see Aspis 2003, where a non representative sample of Italian students is analyzed, finding that on average a student in Southern regions receive less resources (in the order of 500 euros per student per year).

We have compared two Italian macro-regions, South and Centre-North, according to equality of opportunity. In the first application we have focused on individual earnings, while in the second we have considered the distribution of cognitive abilities among students; in both cases we have studied how these different individual achievements vary according to the family background, as measured by the level of parents' education. Our main findings are that, in both applications, the less developed regions in the South are characterized by greater disparities at the global level. In addition they also suffer of greater incidence of inequality of opportunity. According to our results, parents' education play a great role in the level of individual achievements, and this effect is stronger in the South than in the North.

Common to many other less developed regions, Southern Italian regions experience the worse of possible worlds: lower per-capita income accompanied by greater overall income inequality, whose a larger fraction is ethically inequitable. This situation seems pervasive, since we observe it at the earlier stage of schooling careers (when people is still in school) as well as later on in the labour market. One wonders whether these territorial disparities may be addressed by appropriate policies. To begin with, our results point to the role of educational policies. Given the greater inequality of opportunity, it is harder to excel in school in the South than in the Centre-North of Italy: at comparable level of effort, a culturally poorer background is more limiting in the situation where parental background is more effective. As long as family resources and school resources are substitutes in the (underlying) educational production function, a higher impact of parental education is a signal of lower school resources invested in these regions. Here we derive our first policy indication: more resources should be invested in the schools of Southern regions in order to reduce the disparities in terms of inequality of opportunity.

The same type of difficulty seems to emerge in the labour market. Gifted individuals are at a greater disadvantage in the South than in the North when coming from lower social origins. This could represent the impact of family networking in finding good jobs, as well as a reduced availability of good jobs in less technologically advanced areas. This greater obstacles and/or lack of adequate incentives in local labour markets can be linked to existing evidence of internal migration flows (see Viesti 2005), which speaks of a sort of "brain drain", that is strong migration of high skilled workers from the South towards the Northern regions. While part of this migration is certainly explained by the different unemployment rates, it is plausible - and it will be indeed interesting to verify empirically - that the choice to migrate is specially concentrated among individuals with poor family background.

If greater inequality of opportunities in the labour market originates from the opaque working of the labour market, there are no easy solutions. Favouring external migration reduces the inequality of opportunities as measured ex post, but not ex ante. In addition, it depresses the incentives to emerge, given the higher obstacles attributable to factors beyond individual controls. Fairer competition in accessing rationed jobs would constitute the most appropriate policy, and this can be achieved at some extent in the allocation of public jobs.

In the private sector, more transparent intermediation could help in compensating the disadvantage created by differential backgrounds. But these are rather ephemeral suggestions, for a country where more than 50% of the working population declares to have obtained the current job through recommendations of relatives or friends. The final objective of a more fluid society is still a long way off.

References

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

[2] Aspis (2003), *La spesa pubblica per istruzione e cultura in Italia: i principali indicatori*, Angeli

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

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

[5] Brunello G. and D.Checchi (2003) School quality and family background in Italy. forthcoming in *Economics of Education Review*

[6] Bossert W. (1995) Redistribution mechanisms based on individual characteristics. *Math Soc Sciences* 29, 1-17.

[7] Card, D. (1999). The causal effect of education on earnings. in O.Ashenfelter-D.Card (eds). *Handbook of Labor Economics* – vol.3. North Holland, New York.

[8] Checchi D. (2004) Da dove vengono le competenze scolastiche ? L'indagine PISA 2000 in Italia *Stato e Mercato*, 72: 413-453

[9] Checchi D. and Dardanoni V. (2002) Mobility Comparisons: Does using different measures matter ? *Research on Economic Inequality* vol.9: 113-145.

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

[11] Dardanoni V., Fields G., Roemer J.E., Sanchez Puerta M.L. (2004) How demanding should equality of opportunity be, and how much have we achieved?, *mimeo.*

[12] Dworkin R. (1981a) What is equality? Part1: Equality of welfare. *Philos Public Affairs* 10, 185- 246.

[13] Dworkin R. (1981b) What is equality? Part2: Equality of resources. *Philos Public Affairs* 10, 283-345.

[14] Fleurbaey M. (1994) On fair compensation. *Theory Decision* 36, 277-307.

[15] Fleurbaey M. (1995) Three solutions for the compensation problem. *J Econ Theory* 65, 505- 521.

[16] Fleurbaey M. and Maniquet F. Compensation and responsibility, in Arrow K., Sen A. and Suzumura K. (Eds) *Handbook of Social Choice and Welfare*. Elsevier, New York.

[17] Foster J.E. and Shneyerov A.A. Path Independent Inequality Measures. *J. Econ. Theory*, 91, (2000), 199-222.

[18] Goux D. and Maurin E. (2003) On the evaluation of equality of opportunity for income: axioms and evidence, mimeo, INSEE.

[19] Lambert P.J. (2001). *The Distribution and Redistribution of Income* (3rd edition). Manchester, Manchester University Press.

[20] OECD-PISA 2000. *Measuring student knowledge and skills. The Pisa 2000 assessment of reading, mathematical and scientific literacy.* Paris

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

[22] Peragine V. *(2002)* Opportunity egalitarianism and income inequality: the rank-dependent approach. *Math. Soc. Sciences„* 44, pp.45-64.

[23] Peragine V. *(*2004a) Measuring and implementing equality of opportunity for income. *Soc. Choice Welfare,* Vol 22, pp. 1-24.

[24] Peragine V. (2004b) Ranking income distributions according to equality of opportunity, in *J. Econ. Inequality*, Vol. 2.

[25] Rawls J. (1971) *A Theory of Justice.* Cambridge: Harvard University Press.

[26] Roemer J.E. (1993) A pragmatic theory of responsibility for the egalitarian planner. *Philosophy and Public Affairs* 22, 146-166.

[27] Roemer JE (1998) *Equality of Opportunity*, Cambridge, MA: Harvard University Press.

[28] Roemer JE et al. (2003) To What Extent Do Fiscal Regimes Equalize Opportunities for Income Acquisition among Citizens?, *Journal of Public Economics,* 87, 3-4, 539-565.

[29] Ruiz Castillo, J. (2003), The Measurement of Inequality of Opportunities, in Bishop, J. & Y. Amiel (eds.), *Research in Economic Inequality*, 9: 1-34.

[30] Van de gaer D. (1993) Equality of opportunity and investment in human capital. Ph.D. Dissertation, Catholic University of Louvain.

[31] Viesti G. (2005) Recenti flussi migratori in Italia. in "Il Mulino" 4/2005.

[32] Villar A (2004) On the Welfare Evaluation of Income and Opportunity, mimeo, University of Alicante.

Table 1 – Descriptive statistics – Gross earnings – Italy (SHIW) 1993–2000
first row: mean – second row: standard deviation – third row: observations

Highest educational attainment among parents	North	Centre–South	Total
no formal education	27912.91	22898.22	24427.48
	12702.33	12810.93	12982.52
	806	1,837	2,643
primary school	30706.63	27800.04	29199.57
	16786.42	16545.32	16724.15
	4,582	4,934	9,516
lower secondary	33799.69	30012.6	32055.78
	19434.26	17057.58	18471.92
	1,700	1,451	3,151
upper secondary	37359.2	33757.34	35628.98
	24230.18	19316.67	22075.49
	1,072	991	2,063
bachelor	45221.3	42126.49	43539.26
	3055.52	34437.09	32675.73
	308	343	652
Total	32431.77	28325.72	30254.82
	19096.52	17727.43	18496.72
	8,468	9,556	18,024

Note: North includes Piemonte, Val d'Aosta, Liguria, Lombardia, Veneto, Friuli Venezia Giulia, Trentino Alto Adige and Emilia Romagna

Table 2 – Predicted earnings (controlling for gender, experience and survey year)
first row: mean – second row: standard deviation – third row: observations

Highest educational attainment among parents	North	Centre–South	Total
no formal education	37015.56	31120.24	32918.06
	11601.48	12362.02	12432.93
	806	1,837	2,643
primary school	39717.91	35942.89	37760.58
	15468.26	15936.08	15824.56
	4,582	4,934	9,516
lower secondary	43354.4	39110.05	41399.93
	17973.21	15818.8	17143.47
	1,700	1,451	3,151
upper secondary	47790.06	44429.84	46175.91
	21950.38	17917.79	20179.6
	1,072	991	2,063
bachelor	55420.49	52241	53676.79
	28118.97	32048.77	30299.22
	308	343	652
Total	41783.77	36961.85	39227.27
	17691.43	17042.3	17515.93
	8,468	9,556	18,024

Note: the table reports the residuals from the following estimate

$$earnings = \underset{(38.01)}{33133} - \underset{(19.8)}{7463} \cdot d_{1993} - \underset{(15.3)}{5737} \cdot d_{1995} - \underset{(7.6)}{2962} \cdot d_{1998} - \underset{(32.3)}{8676} \cdot female + \underset{(20.9)}{1299} \cdot \exp - \underset{(20.8)}{26} \cdot \exp^2$$

Since the residuals are centered around zero, they have been added a constant (39225) in order to match minimum and maximum of the actual series.

Table 3 – Descriptive statistics: inequality measures

Inequality measures	Actual gross earnings	Predicted gross earnings (net of gender, experience and survey year)	Mean predicted gross earnings (by region, types and quantiles) - X^s
Relative mean deviation	0.18956	0.14041	0.14036
Coefficient of variation	0.61136	0.44652	0.41004
Standard deviation of logs	0.66238	0.42004	0.39308
Gini coefficient	0.28406	0.20799	0.20703
Theil index (GE(α),α=1)	0.15257	0.08110	0.07550
Mean Log Deviation (GE(α),α=0)	0.17512	0.08043	0.07570
Entropy index(GE(α), α=−1)	0.34345	1.1e+02	0.08516
Half(Coeff.Var.squared)(GE(α),α=2)	0.18687	0.09969	0.08406

Table 4 – Mean earnings by "types" and "effort" and macro–regions

first row: mean – second row: observations

North

quantiles→ / types↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	16689.05	23225.94	25887.26	27791.33	29690.2	31072.9	32209.29	33609.23	34704.79	35654.81	36759.52	37758.36	38744.99	39721.1	41408.04	42736.3	44715.19	46942.58	51868.32	69727.2
	41	40	40	41	40	40	41	40	40	40	41	40	40	41	40	40	41	40	40	40
2	17457.04	23997.71	26990.55	28898.56	30497.23	31946.78	33416.19	34709.4	35894.13	37077.61	38298.5	39578.91	40735.97	42085.24	43631.52	45409.3	47730.52	51260.34	58022.27	86840.9
	230	229	229	229	229	229	229	229	229	229	232	227	230	228	229	229	229	229	229	229
3	17939.15	25372.05	28778.47	30920.51	32855.19	34568.98	36108.72	37404.05	38755.2	40137.84	41356.97	42775.41	44101.49	45692.34	47684.48	49905.5	52295.08	57240.89	67262.64	95956.4
	85	85	85	85	85	85	85	85	85	85	85	85	85	86	84	85	85	85	85	85
4	18580.32	26189.05	30061.17	32854.7	35343.59	37448.55	38973.71	40368.27	41670.64	43083.45	44560.82	45754.83	47389.14	49345.66	51310.64	54125.05	57764.03	64591.25	79386.67	117883.1
	54	54	53	54	53	54	54	53	54	53	54	54	53	54	53	54	54	53	54	53
5	19249.2	28521.93	32854.07	35387.42	36581.45	38662.61	41143.27	43649.44	45486.85	47144.99	48745.69	50588.95	52705.31	56045.44	61247.45	67494.63	75958.3	87124.28	102173.8	141914.4
	16	15	16	15	15	16	15	16	15	15	16	15	16	15	15	16	16	15	15	15

Centre-South

quantiles→ / types↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	9098.535	14390.67	17327.28	19614.06	21890.45	24012.46	25706.07	27521.68	29015.58	30341.4	31635.74	33002.87	34335.96	35712.88	37481.62	39159.33	41234.07	43721.98	47328.5	60180.67
	92	92	92	92	92	92	91	92	92	92	92	92	92	91	92	92	92	92	92	91
2	11967.98	18450.87	22028.43	24844.38	26943.44	28683.65	30121.33	31542.24	33001.11	34352.72	35768.52	37113.32	38458.97	39966.67	41603.79	43356.39	45329.57	47891.2	52592.81	74983.5
	247	247	247	246	247	247	246	247	247	247	246	247	247	246	247	247	246	247	247	246
3	14135.1	21144.26	24531.75	27022.88	29183.45	30859.54	32519	34068.32	35562.17	37080.75	38415.05	39741.94	41224.82	42541.11	44223.14	46140.02	48773.7	52155.82	58780.32	84613.44
	73	73	72	73	72	73	72	73	72	73	73	72	73	72	73	73	72	72	73	72
4	15520.87	23725.36	28265.38	30720.21	33265.77	35572.9	37256.86	38536.82	39903.12	41695.02	43178.81	44664.1	46381.57	47914.57	49661.29	51964.18	55175.92	60161.25	69025.85	96813.71
	50	50	49	50	49	50	49	50	49	50	50	49	50	49	50	50	49	50	50	49
5	16208.08	25559.64	29627.96	33104.68	36095.77	38758.73	40621.6	42153.24	43604.57	45100.73	46227.48	47508.41	49340.98	51512.6	53073.52	56294.4	62292.33	76748.95	96349.72	157482.7
	18	17	17	17	17	17	18	17	17	17	17	17	17	18	17	17	17	17	17	17

Figure 1 – Inequality of opportunity by macro regions – Italy (SHIW) 1993-2000

Table 5 – Inequality decomposition, by macroegions – mean log deviation – "tranche" approach

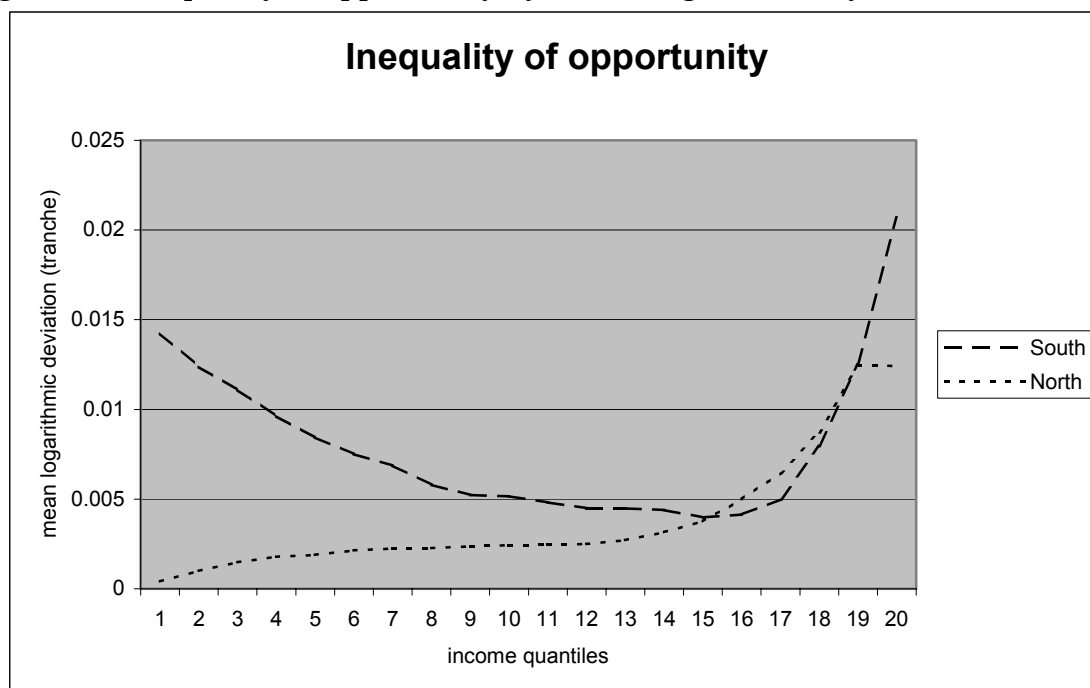|	opportunity inequality	effort inequality	Total inequality (mean predicted earnings X^s)	Total inequality (predicted earnings)	total inequality (actual earnings)
North	0.00389	0.06102	0.06491	0.06840	0.15212
Center-South	0.00793	0.07378	0.08172	0.08754	0.19119

Table 6 – Inequality decomposition, by macroegions – mean log deviation – "types" approach

	opportunity inequality	effort inequality	Total inequality (predicted earnings)
North	0.00429	0.06411	0.06840
Center-South	0.00742	0.08011	0.08754

Table 7 – Determinants of earnings – Italy (SHIW) 1993–2000

OLS – robust t–statistics in brackets – * significant at 10%; ** significant at 5%; *** significant at 1%

	1 north	2 centre– south	3 north	4 centre– south
female	-0.369	-0.352	-0.385	-0.404
	[30.47]***	[23.38]***	[33.58]***	[28.32]***
potential experience	0.057	0.065	0.057	0.061
	[17.31]***	[18.28]***	[17.31]***	[17.29]***
potential experience squared	-0.001	-0.001	-0.001	-0.001
	[15.81]***	[15.98]***	[12.25]***	[11.21]***
parent primary (isced 1)	0.099	0.285	0.022	0.164
	[5.02]***	[14.56]***	[1.17]	[9.15]***
parent lower secondary (isced 2)	0.235	0.433	0.079	0.203
	[10.23]***	[17.61]***	[3.53]***	[8.77]***
parent upper secondary (isced 3)	0.336	0.626	0.086	0.261
	[12.89]***	[22.79]***	[3.27]***	[9.61]***
parent tertiary (isced 4-5-6)	0.473	0.761	0.088	0.312
	[11.39]***	[18.63]***	[2.13]**	[7.88]***
completed years of education			0.062	0.075
			[32.55]***	[36.47]***
Observations	8468	9556	8468	9556
R²	0.20	0.16	0.29	0.27

Note: constant and survey year dummies included.

Dependent variable is the log of gross labour earnings for dependent employees.

Table 8 – Descriptive statistics – Cognitive ability – Italy (PISA) 2000
first row: mean – second row: standard deviation – third row: observations

Highest educational attainment among parents	North	Centre–South	Total
no formal education or primary school	487.00	429.88	448.67
	85.18	79.69	85.68
	75	153	228
lower secondary (isced 2)	499.60	449.95	467.58
	78.37	87.75	87.78
	435	790	1,225
upper secondary vocational (isced 3 b–c)	517.12	458.00	484.54
	82.25	83.22	87.80
	268	329	597
upper secondary academic (isced 3a)	527.96	488.66	504.72
	81.99	84.95	85.93
	705	1,020	1,725
tertiary (isced 5–6)	545.13	496.43	516.22
	83.41	86.65	88.60
	404	590	994
Total	521.93	473.02	492.37
	83.35	88.26	89.59
	1,887	2,882	4,769

Note: North includes Piemonte, Val d'Aosta, Liguria, Lombardia, Veneto, Friuli Venezia Giulia, Trentino Alto Adige and Emilia Romagna

Table 9 – Predicted ability (controlling for gender and age)
first row: mean – second row: standard deviation – third row: observations

Highest educational attainment among parents	North	Centre–South	Total
no formal education or primary school	466.94 85.09 75	408.47 79.88 153	427.70 85.97 228
lower secondary (isced 2)	477.34 76.28 435	430.51 86.20 790	447.14 85.77 1,225
upper secondary vocational (isced 3 b–c)	494.75 80.53 268	440.69 81.29 329	464.96 85.24 597
upper secondary academic (isced 3a)	508.93 81.15 705	469.81 83.86 1,020	485.80 84.95 1,725
tertiary (isced 5–6)	525.98 82.70 404	478.03 84.10 590	497.52 86.75 994
Total	501.62 82.45 1,887	454.14 86.86 2,882	472.92 88.25 4,769

Note: the table reports the residuals from the following estimate

$$ability = 213.47 + 29.46 \cdot female + 1.39 \cdot \exp erience$$
$$\quad\quad\quad (2.95) \quad\quad (11.5) \quad\quad\quad\quad (3.65)$$

Since the residuals are centered around zero, they have been added a constant (472.92) in order to match minimum and maximum of the actual series.

Table 10 – Descriptive statistics: inequality measures

Inequality measures	Actual ability	Predicted ability (net of gender and age)	Mean predicted ability (by region, types and quantiles) - X^s
Relative mean deviation	0.07218	0.07293	0.07278
Coefficient of variation	0.18423	0.18660	0.18206
Standard deviation of logs	0.20568	0.21026	0.19317
Gini coefficient	0.10290	0.10414	0.10309
Theil index (GE(α),α=1)	0.15688	0.15850	0.15671
Mean Log Deviation (GE(α),α=0)	0.07592	0.07696	0.07628
Entropy index(GE(α), α=−1)	0.01078	0.01105	0.01035
Half(Coeff.Var.squared)(GE(α),α=2)	0.01775	0.01823	0.01704

Table 11 – Mean ability by "types" and "effort" and macro–regions
first row: mean – second row: observations

North

quantiles→ types↓	1	2	3	4	5	6	7	8	9	10
1	312.54	395.83	417.12	442.52	461.72	480.82	502.81	524.90	552.15	591.67
	8	7	8	7	8	7	8	7	8	7
2	334.38	404.59	431.40	453.01	470.29	487.58	506.43	527.31	553.19	608.36
	44	43	44	43	45	42	44	44	43	43
3	357.65	408.25	441.23	465.65	483.09	503.16	525.88	544.32	580.61	642.66
	27	27	27	27	26	27	27	27	27	26
4	356.09	431.05	459.60	483.36	502.28	519.09	538.87	566.31	593.95	640.32
	71	70	71	70	72	69	71	70	71	70
5	372.26	443.97	475.33	499.01	519.25	538.19	560.78	580.94	609.79	664.67
	41	40	41	40	41	40	41	40	40	40

Centre-South

quantiles→ types↓	1	2	3	4	5	6	7	8	9	10
1	266.02	328.34	354.17	376.87	398.67	425.66	441.78	464.86	493.67	544.03
	16	15	15	16	15	15	16	15	15	15
2	265.13	342.45	377.68	404.66	428.26	445.83	464.58	488.76	518.19	569.57
	79	79	79	79	79	79	79	79	79	79
3	289.94	353.73	389.76	414.86	436.61	454.06	477.48	494.56	521.47	578.59
	33	33	33	33	33	33	33	33	33	32
4	313.41	384.12	418.93	442.86	465.10	483.74	502.41	523.30	553.19	610.98
	102	102	102	102	102	102	102	102	102	102
5	319.06	388.59	425.89	452.92	473.65	492.87	515.89	534.82	561.64	615.00
	59	59	59	59	59	59	59	59	59	59

Figure 2 – Inequality of opportunity by macro regions – Italy (PISA) 2000
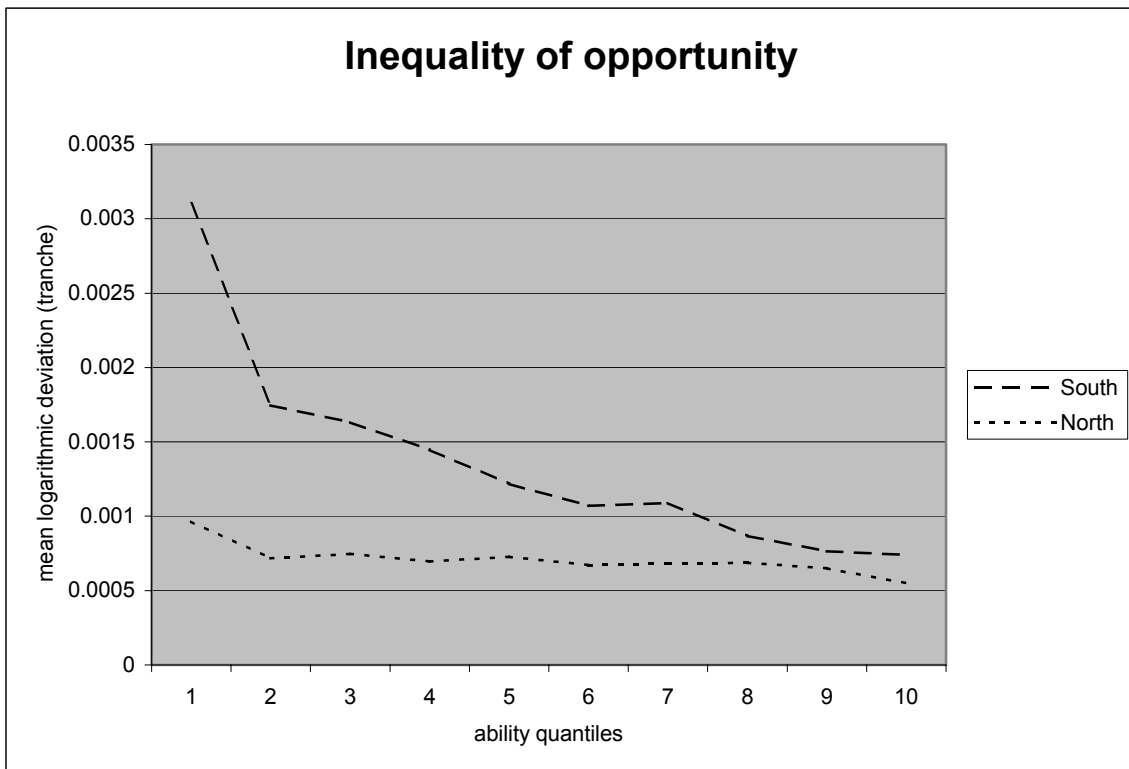
Table 12 – Inequality decomposition, by macroegions – mean log deviation – "tranche" approach – Ability

	opportunity inequality	effort inequality	Total inequality (mean predicted ability X^s)	Total inequality (predicted ability)	total inequality (actual ability)
North	0.000709	0.0126289	0.01334	0.01532	0.01446
Center-South	0.0013677	0.0174451	0.01881	0.02095	0.02032

Table 13 – Inequality decomposition, by macroegions – mean log deviation – "types" approach – Ability

	opportunity inequality	effort inequality	Total inequality (predicted ability)
North	0.0006801	0.0146399	0.01532
Center-South	0.0012013	0.0197487	0.02095

Table 14 – Determinants of cognitive ability – Italy – PISA 2000
OLS – robust t–statistics in brackets – * significant at 10%; ** significant at 5%; *** significant at 1%

	1 north	2 centre–south	3 north	4 centre–south
female	0.049	0.073	0.032	0.039
	[6.18]***	[9.79]***	[4.45]***	[5.62]***
Student age in months	0.004	0.003	0.003	0.002
	[3.51]***	[2.46]**	[2.22]**	[1.62]
parent lower secondary (isced 2)	0.031	0.047	0.033	0.028
	[1.06]	[2.64]***	[1.15]	[1.76]*
parent secondary vocational (isced 3b-c)	0.066	0.074	0.053	0.042
	[2.19]**	[3.89]***	[1.86]*	[2.42]**
parent secondary academic (isced 3a)	0.092	0.135	0.049	0.063
	[3.17]***	[7.91]***	[1.76]*	[3.97]***
parent tertiary (isced 5-6)	0.124	0.153	0.05	0.04
	[4.24]***	[8.69]***	[1.76]*	[2.47]**
attending vocational school (*profess*)			0.252	0.347
			[2.25]**	[6.03]***
attending technical track (*itis*)			0.364	0.449
			[3.26]***	[7.85]***
attending academic track (*licei*)			0.451	0.584
			[4.03]***	[10.29]***
Observations	1887	2882	1887	2882
R²	0.06	0.08	0.24	0.29

Note: constant included. Dependent variable is the log of the test score achieved in reading ability