Society for the Study of Economic Inequality

Working Paper Series

# Empirical approaches to inequality of opportunity: Principles, measures, and evidence

Xavier Ramos
Dirk Van de gaer

ECINEQ WP 2012 – 259

ECINEQ 2012 – 259
Juy 2012

www.ecineq.org

# Empirical approaches to inequality of opportunity: Principles, measures, and evidence*

**Xavier Ramos**†
*Universitat Autonòma de Barcelona, IZA and EQUALITAS*

**Dirk Van de gaer**
*Sherppa, IAE and CORE*

**Abstract**

We put together the different conceptual issues involved in measuring inequality of opportunity, discuss how these concepts have been translated into computable measures, and point out the problems and choices researchers face when implementing these measures. Our analysis identities and suggests several new possibilities to measure inequality of opportunity. The approaches are illustrated with a selective survey of the empirical literature on income inequality of opportunity.

**Keywords**: equality of opportunity, measurement, compensation, responsibility, effort, circumstances.
**JEL Classification**: D3, D63.

---

* We thank Erik Schokkaert, Kristof Bosmans, Jose Luis Figueroa and Stephen Jenkins for useful comments and suggestions. We gratefully acknowledge comments received on preliminary versions presented at the ECINEQ conference (Catania, Italy) and seminars at Universitat Autònoma de Barcelona and Alicante. The paper was nalized when the second author was visiting the Institut d'Anàlisi Econòmica, CSIC, Campus UAB, Bellaterra 08193, Spain. Dirk Van de gaer acknowledges financial support from the FWO-Flanders, research project 3G079112. Xavier Ramos acknowledges financial support of projects ECO2010-21668-C03-02 (Ministerio de Ciencia y Tecnología), 2009SGR-307 and XREPP (Direcció General de Recerca).

† **Contact details**: Xavie Ramos: Xavi.Ramos@uab.cat. Universitat Autònoma de Barcelona, Depart Econ Aplicada, Campus UAB, Bellaterra 08193, Spain.
Dirk Van de gaer: SHERPPA, Vakgroep Sociale Economie, F.E.B., Ghent University, Tweekerkenstraat 2, B-9000 Gent, Belgium.

# 1 Introduction

Beyond the mere concern for individual differences or disparities in outcomes, which has dominated distributive concerns for many decades, the theory of equality of opportunity (Dworkin, 1981a,b; Arneson, 1989; Cohen, 1989) puts individual responsibility in the forefront when assessing situations of economic advantage and disadvantage. It is argued that outcomes such as income level, education attainment or health status, are determined by factors or variables that are beyond individuals' responsibility (so-called circumstances) and by factors for which individuals are deemed responsible (so-called effort or responsibility variables). Inequalities that are due to circumstances are deemed ethically unacceptable while those arising from efforts are not considered offensive. That is, the 'ideal' situation or benchmark is not perfect equality *per se*, as in the measurement of inequality of outcome, but a distribution where efforts are rewarded adequately and the effect of circumstances is compensated for, so that only disparities due to efforts remain.

Both attitude survey research (see, e.g., Schokkaert and Devooght (2003) and Gaertner and Schwettmann (2007)), and experimental evidence (see Cappelen et al (2010)) provide strong evidence that, in judging income distributions, people largely distinguish between circumstances and efforts in the way suggested by equality of opportunity theories. For instance, Cappelen et al (2010) elicit information on what people hold each other responsible for, by means of a dictator game where the distribution phase is preceded by a production phase, and find that a large majority of the participants did not hold people responsible for the randomly assigned price, an impersonal factor beyond individual control, but did hold them responsible for their choice of working time.

This evidence about the social preferences people endorse should be distinguished from the influence of inequality of opportunity on preferences for redistribution, political orientation and actual behavior. A growing amount of empirical evidence shows that preferences for redistribution and political orientation are shaped by fairness concerns. For instance, Alesina and La Ferrara (2005) show for the United States that people who believe that individual economic success is related to individual effort rather than family background or luck, have lower preferences for redistribution, while Alesina and Angeletos (2005), using data from the World Value Survey, find that fairness perceptions are associated with the individuals' political orientation: when people believe that effort is the main determinant of economic advantage, redistribution and taxes are low, whereas in societies where people think of birth and connections as the main determinants of economic success, taxes and redistribution will be higher. Since the determinants of economic inequality (circumstances versus efforts) influence individual incentives, these determinants are related with aggregate economic outcomes, such as economic growth. In its World Development Report of 2006, the World Bank argues that income inequality due to circumstances may lead to suboptimal accumulation of human capital and thus to lower growth, while income inequality due to responsibility-related variables may encourage individuals to invest in human capital and exert the largest effort possible (World Bank, 2005). In line with this, Marrero and Rodríguez (2010), using data for the U.S. from the Panel Survey on Income Dynamics, find that income inequality due to effort enhances income growth, while the part of income inequality which is accounted for by circumstances correlates negatively with growth. Our concern in this paper is not which measure of inequality of opportunity is best suited to explain a particular phenomenon about reality (which depends on the phenomenon under scrutiny and the way the world works), but rather with the measures that have been proposed in the normative literature dealing with the measurement of inequality of opportunity.

In recent years, we have seen an explosion of empirical literature that tries to determine whether opportunities are equally distributed, and tries to measure the extent of inequality of opportunity or the contribution of inequality of opportunity to total income inequality –see, e.g., Almas *et al.* (2011), Björklund *et al.* (2011), Bourguignon *et al.* (2007), Checchi and Peragine (2010), Devooght (2008), Lefranc *et al.* (2008) and Pistolesi (2009). The measurement of equality of opportunity entails many methodological and empirical questions that are often difficult to resolve. Rather than addressing these issues in a systematic and coherent manner, the literature has developed very rapidly in many seemingly unrelated directions. As a result, there is often no explicit correspondence between the theoretical principles and the measures put forth and employed to empirically implement the equality of opportunity approach. In this survey we bridge this gap by presenting and discussing in a systematic manner the main conceptual issues and outline the solutions that have been proposed in the literature. Our analysis identifies and suggests several new possibilities to measuring inequality of opportunity. However, we limit ourselves in several respects. First, we discuss inequality of opportunity for income. Hence we do not address the issues related to multi-dimensional outcomes, which arise for instance naturally in the capabilities approach[1]. Other one-dimensional outcomes, such as health and education have been analyzed using similar techniques as the ones we describe here[2]. Due to the one-dimensional focus the opportunity set to which individuals have access contains only incomes, which drastically simplifies the comparison of individual's opportunity sets[3]. Second, we do not discuss the design and evaluation of policies from an equality of opportunity perspective, as this raises different important, complex and often model dependent issues [4].

The theoretical literature has pointed out that the idea of equality of opportunity embodies two basic principles. The *compensation principle*, which demands that inequalities due to circumstances be eliminated and the *reward principle*, which is concerned about how to reward efforts amongst individuals

[1]See Schokkaert (2009) for a recent discussion of the capabilities approach.

[2]See, e.g., Dias (2009) or Trannoy et al (2010) for health and Peragine and Serlenga (2008) for education.

[3]For an overview of the literature on the evaluation of more general opportunity sets, see Barbera et al (2004).

[4]Roemer (1998a) is a good starting point, and Pignataro (2011) provides a recent survey. For theoretical contributions on optimal tax policy see Fleurbaey and Maniquet (2006 and 2007), Ooghe and Luttens (2007) or Jacquet and Van de gaer (2011). For the evaluation of social programs see Van de gaer et al (2012).

with identical circumstances.

Regarding the compensation principle, an important methodological issue has to do with whether we want to take an ex-ante or rather an ex-post approach to compensation. The ex-post approach looks at each individual's actual outcome and is concerned with outcome differences amongst individuals with the same responsibility characteristics –and different circumstances. This approach is very demanding on the data since we need to observe responsibility variables. When, as it is often the case, no direct observations on responsibility variables are available, we need to impose working assumptions about the relationship between responsibility characteristics and outcomes that enable the identification of an underlying responsibility variable. The ex-ante approach, instead, focuses on prospects, so there is equality of opportunity if all individuals face the same set of opportunities (or sets that are equally valued), regardless of their circumstances. The ex-ante approach is interesting *per se* if we believe that opportunity ought to be measured by the set of opportunities that individuals face, but it is also useful when effort has not been exerted as yet. One empirical advantage of this approach is that efforts need not be identified, since outcome prospects are usually measured by some measure of centrality of the distribution of the outcome amongst individuals with identical circumstances. Using the framework in Fleurbaey and Peragine (2011), we illustrate their result that the ex-ante and ex-post approach are incompatible.

Regarding the reward principle, the focal points in the literature are liberal reward and utilitarian reward. The former says that the government should not redistribute income between those that share all circumstance characteristics, as their income differences are exclusively due to differences in efforts. The latter says that we should not be concerned with (i.e. express zero inequality aversion with respect to) income differences that are only due to differences in efforts, such that we should only be concerned with the sum of the incomes of those that only differ in terms of effort. We introduce a third reward principle, "inequality averse reward", which is motivated by either the stochastic nature of incomes and risk aversion (see Lefranc et al (2009)) or the existence of unjustified inequalities in incomes after conditioning on circumstances (see Roemer (2010)). All three reward principles are shown to be incompatible with ex-post compensation.

Several approaches to measure inequality of opportunity, based on information on outcomes, circumstances and efforts have been proposed in the literature. We distinguish direct measures that measure how much inequality remains when only inequality due to circumstances is left from indirect measures that measure how much inequality remains after opportunities are equalized. We also discuss the rationale for the stochastic dominance and norm based approaches found in the literature.

When researchers want to compute inequality of opportunity, they are confronted with several difficulties. They have to decide which outcomes to focus on, which variables are circumstances and which efforts. This is a normative issue, but sensitivity analysis with respect to the circumstances-effort split helps to establish the robustness of the results. Not all circumstances are always observed. Unobserved circumstances typically lead to an underestimation of the amount of inequality of opportunity. Efforts are often unobserved and observed efforts are correlated with circumstances. The former problem can be resolved using a non-parametric technique proposed by Roemer (1993) or parametric techniques (Björklund et al. (2011) or Salvi (2007)). The latter is typically resolved using regression analysis, as suggested by Bourguignon et al. (2007). We analyze the implications of these issues and the solutions used in the literature. To give the reader a flavor of the kinds of results that can be obtained with the various approaches, we discuss the empirical findings of some selected recent studies. As the empirical literature is booming the last few years and several new studies appear every month, no attempt is made to be exhaustive in the overview of empirical findings.

The paper is structured as follows. Section 2 first uses a simplified version of the framework recently developed by Fleurbaey and Peragine (2011) to illustrate the incompatibilities between ex-post and ex-ante compensation and between ex-post compensation and the different reward principles. The next section discusses how the insights from the ex-post versus ex-ante compensation debate and reward principles have been used to construct measures of inequality of opportunity. Section 4 discusses several data imperfections: unobserved circumstances, construction of measures of efforts, luck and econometric error terms. Section 5 illustrates the issues discussed in the previous sections by presenting the results of some empirical studies. Section 6 concludes.

# 2 Principles

In this section we introduce the major insights from the theoretical literature on the evaluation of distributions of incomes from a perspective of equality of opportunity. We assume that we only observe (or want to use) information about individuals' incomes, their circumstances and their efforts. In particular, let $N = \{1, 2, \ldots, n\}$ with $n \geq 2$ be the set of individuals. For each individual $k \in N$, we observe $y_k \in \mathbb{R}_{++}$, his income, $a_k^R \in \mathbb{R}^{d^R}$, a vector of characteristics for which individual $k$ is responsible (efforts) and $a_k^C \in \mathbb{R}^{d^C}$, a vector of characteristics for which he is not responsible (circumstances). A type is a set of individuals sharing the same circumstances: for every value of the $d^C$-dimensional vector $a_k^C$ that occurs in the population, a type is defined.[5] Let there be $m^C$ such types, indexed by $i \in \{1, \ldots, m^C\}$. Similarly, a tranche is a set of individuals sharing the same efforts: for every value of the $d^R$-dimensional vector $a_k^R$ that occurs in the population, a new tranche is defined.[6] Let there be $m^R$ such tranches, indexed by $j \in \{1, \ldots, m^R\}$.

In this section we assume that incomes are exclusively determined by efforts and circumstances such that all those having the same circumstances and efforts obtain the same outcome. Hence, the relevant data can be summarized

[5] This definition of "type" was introduced by Roemer (1993).
[6] This definition of "tranche" was introduced by Peragine (2004).

by the $m^C \times m^R-$dimensional matrix of incomes $Y = [Y_{ij}] \in \mathbb{R}_{++}^{m^C \times m^R}$, giving the income for each circumstance-effort combination occurring in the population, and the matrix $P = [P_{ij}]$, giving the frequency with which circumstance-effort combination $ij$ occurs in the population. Naturally, all $P_{ij} \geq 0$ and $\sum_{j=1}^{m^R} \sum_{i=1}^{m^C} P_{ij} = 1$. The frequency of type $i$ in the population is $P_{i.} = \sum_{j=1}^{m^R} P_{ij}$ and the frequency of tranche $j$ is $P_{.j} = \sum_{i=1}^{m^C} P_{ij}$. We only consider situations where all $P_{ij} > 0$.[7] Let $\mathcal{P} = \left\{P \in \mathbb{R}_{++}^{m^C \times m^R} \mid \sum_{j=1}^{m^R} \sum_{i=1}^{m^C} P_{ij} = 1\right\}$. Our purpose is to find an ordering $\succeq_P$ of matrices $Y$ for every $P \in \mathcal{P}$. Observe that this ordering is conditional on $P$; in this section the types and tranches, as well as the distribution of the population over types and tranches are kept fixed when we compare different income matrices.

## 2.1 Ex-ante versus ex-post

The first fundamental idea in the literature on equality of opportunity is that differences that are due to circumstances should be compensated. As stated by Fleurbaey and Peragine (2011), compensation can be done using an ex-post or ex-ante approach. Ex-post compensation tries to make the outcomes for those individuals having the same effort as equal as possible. Formally,

**EPC** (Ex-Post Compensation): For all $Y, Y' \in \mathbb{R}_{++}^{m^C \times m^R} : Y \succ_P Y'$ if there exists $Y'_{ij} \geq Y_{ij} \geq Y_{lj} \geq Y'_{lj}$ with either the first or the last inequality holding strict and for all $ab \notin \{ij, lj\} : Y'_{ab} = Y_{ab}$.

The condition in the axiom requires that, for effort $j$, the distribution of outcomes is more equal in matrix $Y$ than in $Y'$. Ex-ante compensation, on the other hand, prefers redistribution from a type that is unambiguously better-off to a type that is unambiguously worse-off.

**EAC** (Ex-Ante Compensation): For all $Y, Y' \in \mathbb{R}_{++}^{m^C \times m^R} : Y \succ_P Y'$ if (i) there exists $i$ and $l$ such that for all $j \in \left\{1, \ldots, m^R\right\} : Y_{ij} \geq Y_{lj}$ and (ii) there exists $j, q \in \left\{1, \ldots, m^R\right\} : Y'_{ij} > Y_{ij}, Y_{lq} > Y'_{lq}$ and for all $ab \notin \{ij, lq\} : Y'_{ab} = Y_{ab}$.

Condition (i) guarantees that in matrix $Y$ type $i$ is unambiguously better-off than type $l$, while condition (ii) implies that the inequalities between types $i$ and $l$ are larger in matrix $Y'$ than in matrix $Y$.

While both conditions look reasonable, it has been shown by Fleurbaey and Peragine that they are incompatible. To see this, consider the following outcome matrices for a situation where we have 4 types and 2 tranches:

$$Y^1 = \begin{bmatrix} 20 & 15 \\ 15 & 10 \\ 30 & 6 \\ 25 & 1 \end{bmatrix} \text{ and } Y^2 = \begin{bmatrix} 21 & 15 \\ 15 & 9 \\ 30 & 7 \\ 24 & 1 \end{bmatrix}.$$

[7] This is done to save on notation; otherwise most of the axioms listed below have to be made conditional on the relevant $P_{ij}$ occurring in the axioms being strictly positive.

Starting from $Y^1$, we observe that the first row has better opportunities than the second and the third row has better opportunities than the fourth. Increasing the inequalities between the first and second row (by increasing $Y^1_{11}$ and decreasing $Y^1_{22}$) and increasing the inequalities between the third and fourth row (by increasing $Y^1_{32}$ and decreasing $Y^1_{41}$) results in $Y^2$, such that, by EAC, we have $Y^1 \succ_P Y^2$. Now, start from $Y^2$, increase the inequalities in the first column (by decreasing $Y^2_{11}$ and increasing $Y^2_{41}$) and increase the inequalities in the second column (by increasing $Y^2_{22}$ and decreasing $Y^2_{32}$) and we get $Y^1$. Hence, by EPC, $Y^2 \succ_P Y^1$, contradicting our previous finding. We have thus illustrated the following proposition.

**Proposition 1** (Fleurbaey and Peragine (2011)): EPC and EAC are incompatible.

The existence of this incompatibility implies that, if one wants to evaluate outcome matrices from the perspective of equality of opportunity, a choice has to be made between ex-ante and ex-post compensation.

## 2.2 Reward principles

The second fundamental idea in the literature on equality of opportunity is that efforts should be adequately rewarded. Liberal reward is the first and most prominent reward principle in the axiomatic literature on fair allocations (see, e.g., Bossert (1995), Fleurbaey (1995a) and Bossert and Fleurbaey (1996)) and fair social orderings (see, e.g. Fleurbaey and Maniquet (2005, 2008)). It states that government taxes and transfers should respect differences in incomes that are due to differences in responsibility. Take two individuals belonging to the same type but that have exerted different efforts resulting in different pre-tax incomes. According to the liberal reward principle, the tax policy has to respect the income differences that are due to differences in exerted effort, which implies that these individuals should pay the same tax. Let $T_{ij}$ be tax on individuals of type $i$ and tranche $j$. The liberal reward principle can then be stated formally as follows.

**LR** (Liberal Reward): $\forall i \in \left\{1, \ldots, m^C\right\} : T_{ij} = T_{ik}$ for all $j, k \in \left\{1, \ldots, m^R\right\}$.

It is easy to see that there is a tension between LR and EPC. Consider the following income matrix $Y^3$, which gives before tax incomes in a situation with 2 types and 2 tranches:

$$Y^3 = \begin{bmatrix} 30 & 5 \\ 20 & 10 \end{bmatrix}.$$

Suppose we want to use a tax policy to compensate for ex-post inequalities. Within the first column of $Y^3$, this calls for a transfer from the first type to the second type. Within the second column, an opposite transfer is required. Consequently, ex-post compensation can go against LR. Bossert (1995) and Fleurbaey (1995a) have shown that the two principles are, in general, incompatible[8].

**Proposition 2** (Bossert (1995) and Fleurbaey (1995)): LR and EPC are incompatible.

A second reward principle which Fleurbaey (2008) calls utilitarian reward has been used more frequently in the empirical literature. The principle says that respecting the income differences that are due to differences in effort requires zero inequality aversion with respect to differences in incomes that are due to differences in efforts, hence we have to focus on the sum of the incomes of those that share the same circumstances, leading to the following axiom[9].

**UR** (Utilitarian Reward): For all $Y, Y' \in \mathbb{R}_{++}^{m^C \times m^R}$ : $Y \succ_P Y'$ if there exists $i \in \{1, \ldots, m^C\}$ such that $\sum_{j=1}^{m^R} Y_{ij}P_{ij} > \sum_{j=1}^{m^R} Y'_{ij}P_{ij}$ and for all $l \neq i$ we have $\sum_{j=1}^{m^R} Y_{lj}P_{lj} = \sum_{j=1}^{m^R} Y'_{lj}P_{lj}$.

As shown by Fleurbaey and Peragine, utilitarian reward is incompatible with ex-post compensation. To illustrate this, consider the following income matrices for a situation where we have 2 types and 2 tranches:

$$Y^3 = \begin{bmatrix} 30 & 5 \\ 20 & 10 \end{bmatrix} \text{ and } Y^4 = \begin{bmatrix} 30-\lambda & 5+\varepsilon \\ 20+\frac{P_{22}}{P_{21}}\varepsilon & 10-\varepsilon \end{bmatrix} \text{ with } \lambda > \frac{P_{12}}{P_{11}}\varepsilon.$$

Observe that, for $\varepsilon$ sufficiently small, the change from $Y^3$ to $Y^4$ represents a decrease in inequality within the first and within the second column. Hence, by EPC, $Y^4 \succ_P Y^3$. Next, in moving from $Y^3$ to $Y^4$ the change in the second row is such that $P_{21}Y_{21}^4 + P_{22}Y_{22}^4 = P_{21}Y_{21}^3 + P_{22}Y_{22}^3$, while with the restriction on $\lambda$, we have $P_{11}Y_{11}^4 + P_{12}Y_{12}^4 < P_{11}Y_{11}^3 + P_{12}Y_{12}^3$, such that by UR, $Y^3 \succ_P Y^4$, contradicting our previous finding. Hence we have the following proposition.

**Proposition 3** (Fleurbaey and Peragine (2011)): UR and EPC are incompatible.

A third reward principle explicitly rejects utilitarian reward by claiming that some compensation is due even after taking circumstances into account. A first reason (see Lefranc et al (2009)) forces us to widen the framework to include random variation as a third factor leading to differences in incomes (apart from

[8]Originally the incompatibility was shown in the literature searching for first best allocation rules. To deal with LR in the context of social orderings, a richer model of the economy is required in which the ordering uses information of the actual transfer system (see Fleurbaey and Peragine (2011)). As we are unaware of any empirical application that takes this information into account, we limit ourselves to pointing out the tension between LR and EPC.

[9]Theoretical contributions focus on utilities rather than incomes, hence the name of the axiom.

effort and circumstances): after conditioning on circumstances, incomes are stochastic and since individuals are risk averse, we should evaluate opportunity sets in a risk averse way. A second reasoning is ex-post and rejects liberal and utilitarian reward as still some compensation is due after conditioning on an incomplete list of circumstances. Roemer (2010) attacks the liberal reward principle, because liberal reward holds that no more compensation should be made than that needed to correct inequalities due to different (measured) circumstances. As such, current property rights are no more adjusted than is necessary to compensate people for disadvantageous circumstances. His point is that it is difficult to see why the current property rights should be the benchmark. Hence, he advocates not to focus on incomes, but to take an increasing concave (inequality adverse) transformation of incomes as the relevant outcome variable. This can be called an inequality-averse reward principle.

**IAR** (Inequality Averse Reward): For all $Y, Y' \in \mathbb{R}_{++}^{m^C \times m^R} : Y \succ_P Y'$ if there exists $i \in \left\{1, \ldots, m^C\right\}$ and $\delta \in \mathbb{R}_{++}$ such that $Y_{ij} = Y'_{ij} - \delta \geq Y_{ik} = Y'_{ik} + \delta$ and for all $ab \notin \{ij, ik\} : Y'_{ab} = Y_{ab}$.

Again, it is easy to see that this reward principle conflicts with EPC. Consider the following income matrices for a situation with 2 types and 2 tranches.

$$Y^5 = \begin{bmatrix} 10 & 40 \\ 20 & 30 \end{bmatrix} \text{ and } Y^6 = \begin{bmatrix} 10 & 40 \\ 19 & 31 \end{bmatrix}$$

From IAR it follows immediately that $Y^5 \succ_P Y^6$, while from EPC, $Y^6 \succ_P Y^5$, a contradiction. As a result, we have

**Proposition 4** : IAR and EPC are incompatible.

Propositions 2, 3 and 4 illustrate the difficulty to reconcile EPC with reward principles, a difficulty emphasized in Fleurbaey and Peragine (2011). No such incompatibilities arise with EAC.

## 2.3 Luck

Luck is a complex and important factor, which determines most economic outcomes. Many different factors have been put under the label “luck” by economists -see, e.g. Meade (1974). As Lefranc, Pistolesi and Trannoy (2009) neatly explain, different forms of luck deserve different treatment. We distinguish between forms of luck that require full, partial or no compensation.

The first form of luck reflects Rawl’s idea of social lottery, i.e. economic advantage that is due to factors related to the family or social origin one happens to fall into, such as family or social networks and influences. Such *social background luck* is almost universally considered as a circumstance and ought to be fully compensated for.

The second, *genetic luck*, captures Rawl’s concept of natural lottery, where constituent characteristics of the individual, such as genetically inherited factors like talent, are responsible for differential success. These constituent attributes are thought to be pre-determined and exogenous to the individual and so, *ceteris paribus*, most authors agree that they are circumstances such that outcomes should be equal regardless of them. This view is not uncontested, however. For instance, Nozick (1974)'s view of self-ownership argues that individuals deserve to benefit from their inborn traits.

The third corresponds to Dworkin's brute luck, defined as those situations where the individual cannot alter the probability that an event takes place. By definition, the individual is not responsible for such events happening and thus it seems reasonable to argue in favor of full compensation. However, since full compensation of brute luck may entail huge redistribution, cause large distortions thereby diminishing opportunities for all and implementation of compensation for brute luck requires a lot of information about individuals which is usually not available, some authors have put forward other, weaker justice requirements. For instance, Vallentyne (2002) suggests to compensate only for *initial brute luck*, that is, brute luck that occurs before individuals are deemed responsible for their choices and preferences.[10]

The last form of luck is Dworkin's *option luck* and arises when individuals deliberately take risk, which is assumed to be calculated, isolated, anticipated and avoidable. Since by definition risks of option luck are avoidable and taken deliberately, some authors argue that the resulting differences in outcomes are legitimate. Contrary to that, Fleurbaey (1995b) argues in favor of full compensation because with option luck small errors of choices may involve disproportionate penalties, which he considers unjust. In his recent book, though, Fleurbaey (2008) provides arguments for partial compensation. He sees option luck as consisting of two distinct components: on the one hand, the individual decision to choose a lottery, which is a voluntary act and thus does not deserve compensation, and on the other hand, the randomness intrinsic to any lottery, which should be at least partially compensated for.

To summarize: social background luck and brute luck are circumstances. Also genetic luck is generally considered to be a circumstance. Option luck is either an effort or a partial circumstance.

# 3 Measures

The analysis in the previous section highlights the point made in Ooghe et al (2007) that ex-post inequality of opportunity is concerned with the inequalities within each column of $Y$, while ex-ante inequality of opportunity is concerned with the inequalities between the rows of $Y$. This has an important implication: when effort is distributed independently of type [11], full equality of ex-post

[10] Vallentyne (2002) defines *later brute luck* as the brute lack that occurs after a 'canonical' moment (Arneson, 1990) where individuals become responsible for their choices and preferences. As Lefranc et al. (2009) suggest, as long as initial and later brute luck are related, compensation for the former implies at least partial compensation also for the latter.

[11] Formally, for all $j \in \{1, ..., m^R\}$ and $i, k \in \{1, ..., m^C\}$ with $i \neq k$ it must be that $p_{ij}/p_{i.} = p_{kj}/p_{k.}$

opportunities (absence of inequalities within columns) implies full equality of ex-ante opportunities (equal rows).

When comparing actual income distributions from the perspective of inequality of opportunity, the framework has to be adjusted to allow comparisons between income distributions with different circumstance-effort distributions i.e. with different matrices $P$. In addition, the framework should allow for unobserved and random variables. Hence individual $k$ 's income, $y_k$, is assumed to depend on his circumstances $a_k^C$, his efforts, $a_k^R$, unobserved variables $u_k$ and a random term $e_k$, such that

$$y_k = g\left(a_k^C, a_k^R, u_k, e_k\right) \quad \text{where} \quad g : \mathbb{R}^{d^C} \times \mathbb{R}^{d^R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{++}.$$

As the $u_k$ is unobserved, and the functional form $g$ is unknown, the parametric approach imposes a functional form to estimate the equation, yielding the function

$$\widehat{g}\left(a_k^C, a_k^R, e_k\right) \quad \text{where} \quad \widehat{g} : \mathbb{R}^{d^C} \times \mathbb{R}^{d^R} \times \mathbb{R} \to \mathbb{R}_{++}.$$

An estimate of $y_k$, $\widehat{y}_k$, can be obtained by setting $e_k$ equal to zero in the above equation. Observe that the effect of the unobserved variable will be taken over by the effect of observed circumstances and efforts, to the extent that these are correlated with the unobserved variables. The rest of the effect of unobservables as well as specification errors go into the estimated random variation, $\widehat{e}_k$, which is defined implicitly by the equation $y_k = \widehat{g}\left(a_k^C, a_k^R, \widehat{e}_k\right)$. For some purposes, it is convenient to estimate incomes as a function of, on the one hand, either circumstances or efforts and, on the other hand, random variation:

$$\widehat{g}^C\left(a_k^C, e_k\right) \quad \text{where} \quad \widehat{g}^C : \mathbb{R}^{d^C} \times \mathbb{R} \to \mathbb{R}_{++},$$

$$\widehat{g}^R\left(a_k^R, e_k\right) \quad \text{where} \quad \widehat{g}^R : \mathbb{R}^{d^R} \times \mathbb{R} \to \mathbb{R}_{++}.$$

These equations can be used to estimate incomes by, respectively, $\widehat{y}_k^C$ and $\widehat{y}_k^R$ by setting $e_k$ equal to zero. In the first (second) case, the effect of both omitted efforts (circumstances) and unobservables are taken over by circumstances (efforts) to the extent that these are correlated. The rest of their effect as well as specification errors go into the estimated random variation, $\widehat{e}_k^C$ $(\widehat{e}_k^R)$ which is defined implicitly by the equation $y_k = \widehat{g}^C\left(a_k^C, \widehat{e}_k^C\right)$ $(y_k = \widehat{g}^R\left(a_k^R, \widehat{e}_k^R\right))$. For future reference, let $N_{k.} = \left\{i \in N \mid a_i^C = a_k^C\right\}$ and $N_{.k} = \left\{i \in N \mid a_i^R = a_k^R\right\}$, be the sets of individuals sharing the circumstances $a_k^C$ (belong to the same type) and efforts $a_k^R$ (belong to the same tranche), respectively.

Following Pistolesi (2009), we can distinguish between a direct and indirect approach to the measurement of inequality of opportunity. Other approaches in the literature are based on stochastic dominance or deviations between actual income and norm income. We discuss these approaches in turn and conclude the section with an overview.

## 3.1 Direct measures

A first approach determines the amount of inequality of opportunity directly by estimating the inequality in a counterfactual income distribution $y^c$ in which all inequalities due to differences in effort have been eliminated, such that only the inequality that is due to differences in circumstances is left:

$$I\left(y^{c}\right). \tag{1}$$

The crucial distinction between ex-ante and ex-post approaches lies in the construction of the counterfactual $y^c$. From an ex-ante viewpoint, we should replace every individual's actual income by some evaluation of his opportunity set. Evidently, the value assigned to his opportunity set should not depend on his own effort level.

So far the ex-ante approaches proposed to implement $I\left(y^{c}\right)$ rely mostly on a non-parametric estimate of the value of an individual's opportunity set. The following proposals have been made in the literature:

$$y_k^{c1} = \frac{1}{|N_{k.}|}\sum_{i\in N_{k.}} y_i, \tag{2}$$

$$y_k^{c2} = \frac{1}{|N_{k.}|}\sum_{i\in N_{k.}} i\widetilde{y}_i, \tag{3}$$

where $\widetilde{y}_i$ is the $i-$th smallest level of income in the set $N_{k.}$. The first proposal is due to Van de gaer (1993) and measures the value of the opportunity set by the average income of those that are of the same type as the individual considered. This equals the surface under the Pen parade of the income distribution of the individual's type. The income distribution $y^{c1}$, in which every individual's income is replaced by the mean income of his type is called the "smoothed income distribution" by Checchi and Peragine (2010). As the specification implies no inequality aversion for differences in incomes that are due to responsibility, this specification is inspired by utilitarian reward. Since all income differences that are due to circumstances are morally objectionable, Van de gaer proposes an inequality index with infinite inequality aversion.[12] Checchi and Peragine (2010) and Ferreira and Gignoux (2011) point out that most standard inequality indices measuring inequality of opportunity as inequality in the smoothed income distribution satisfy, apart from utilitarian reward an additional desirable property: when a transfer is made from an individual in a richer type to an individual in a poorer type, regardless of the former individual being poorer than the latter, inequality of opportunity falls. They express a preference for the mean log deviation, because the index allows an exact decomposition of total income inequality into (1) with (2) as the value of each individual's opportunity set and a counterfactual in which all types have an opportunity set of the same value[13], since it is the only decomposable inequality measure that is path independent (Foster and Shneyerov, 2000). Recently, Aaberge et al. (2011) proposed a rank dependent measure (which includes the Gini coefficient

[12] As this ordering is only sensitive to what happens to the worst-off type, it does not satisfy axioms EAC nor UR defined in section 2. It is straightforward to formulate a leximin extension of the ordering that satisfies both axioms, however.

[13] This counterfactual is equal to (12) below.

as a special case) as inequality index to measure ex-ante inequality in the vector $y^{c1}$. This procedure shares the two standard properties mentioned above, but does not have the exact decomposition property. The second proposal, (3), was formulated by Lefranc et al (2008) and measures the value of the opportunity set by the surface under the generalized Lorenz curve of the income distribution of the individual's type. As such, it embodies the inequality averse reward principle. They propose a Gini coefficient to measure ex-ante inequality in the vector $y^{c2}$.

So far, the only parametric estimate of $y^c$ has been put forth by Ferreira and Gignoux (2011). As efforts can be correlated with circumstances (see also section 4.2.4), they propose to measure the value of an individual's opportunity set as

$$y_k^{c3} = \widehat{g}^C \left(a_k^C, 0\right), \tag{4}$$

such that everybody's opportunity set is valued by the reduced form estimate of his income, given his circumstances and with the random term equal to its expected value 0. As all within-type inequalities are eliminated in the vector $y^{c3}$, this is indeed an ex-ante approach: only between type inequalities are relevant. Ferreira and Gignoux (2011) interpret the resulting inequality measure as a parametric estimate of (1) with (2) as the value of the opportunity sets, such that, for the reasons given above they advocate the use of the mean log deviation as inequality index.

An evident parametric alternative, closer in spirit to (2) would rely on the estimate of $\widehat{g}$ rather than $\widehat{g}^C$, and use

$$y_k^{c4} = \frac{1}{|N_{k.}|} \sum_{i \in N_{k.}} \widehat{g}\left(a_i^C, a_i^R, 0\right). \tag{5}$$

Compared to the Ferreira and Gignoux measure described above, using $y^{c4}$ in a measure of inequality of opportunity (1), has the advantage that it deals with the covariance between $a^C$ and $a^R$ in a more flexible type-dependent way. The disadvantage is that, contrary to $y^{c3}$ the estimation of $\widehat{g}$, requires observations on $a^R$. We are unaware of any application of $y^{c4}$.

From an ex-post point of view, to eliminate all inequalities that are due to efforts, we replace every individual's income by the income he could have obtained if he would have put in a reference level of effort. Roemer (1993) was the first to propose such an ex-post approach to compute (1) and used a non-parametric procedure. He fixes a reference value for the responsibility variable $\overline{a}^R$ and defines set $N_{k.}^{\overline{a}^R} = \left\{i \in N_{k.} \mid a_i^R = \overline{a}^R\right\}$, which contains all individuals that have the same circumstances as individual $k$ and have the reference value for the responsibility vector. Next define

$$y_k^{c5}\left(\overline{a}^R\right) = \frac{1}{\left|N_{k.}^{\overline{a}^R}\right|} \sum_{i \in N_{k.}^{\overline{a}^R}} y_i, \tag{6}$$

the average income of those that are of the same type as individual $k$ and have the reference value for the responsibility characteristic.[14] Applying (1) results in an inequality measure whose value depends on the reference value $\overline{a}^R$, which we denote by

$$I\left(y^{c5}\left(\overline{a}^R\right)\right).$$

Roemer argues that the choice of reference value $\overline{a}^R$ is arbitrary, and proposes therefore the following averaged inequality measure:

$$\frac{1}{n}\sum_{l=1}^{n} I\left(y^{c5}\left(a_l^R\right)\right). \tag{7}$$

As all inequalities that are due to differences in circumstances are morally objectionable, Roemer proposes to apply an infinite inequality aversion to compute $I\left(y^{c5}\left(a_l^R\right)\right)$ in (7) and puts $I\left(y^{c5}\left(a_l^R\right)\right)$ equal to the lowest value of the vector $y^{c5}\left(a_l^R\right)$ divided by mean income.[15] In a recent paper, Aaberge et al. (2011) propose to use a rank dependent measure to compute $I\left(y^{c5}(a_l^R)\right)$ in (7).

The ex-post approach to implement (1) semi-parametrically was proposed by Pistolesi (2009)[16] and it is obtained by setting a reference value for the responsibility variable, $\overline{a}^R$ in the estimate of the function $g\left(a_k^C, a_k^R, e_k\right)$:

$$y_k^{c6}\left(\overline{a}^R\right) = \widehat{g}\left(a_k^C, \overline{a}^R, e_k\right). \tag{8}$$

Compared to the non-parametric methodology, the parametric methodology has the advantage that it always yields meaningful estimates for $y_k^{c6}$, even when the combination $\left(a_k^C, \overline{a}^R\right)$ does not occur in the sample. Pistolesi experiments with different inequality measures: he uses the Theil index, the mean log deviation, the half squared coefficient of variation and the standard deviation of logs. In the computation of $y^{c6}$, $e_k$ can be set equal to zero, or to its estimated value $\widehat{e}_k$. The former amounts to treating $e_k$ as an effort variable with reference value zero, the latter to treating it as a circumstance. Most authors take the mean value for effort in the sample as the reference value $\overline{a}^R$. Following Roemer, one can use an averaged inequality measure similar to (7), where $y^{c6}\left(a_l^R\right)$ replaces $y^{c5}\left(a_l^R\right)$. We are unaware of any application of such a direct parametric averaged inequality of opportunity measure. There exist some theoretical results on the consequences of taking different reference values in the context of particular models -see, e.g., Luttens and Van de gaer (2007), but the choice of reference value remains an unsettled issue.

## 3.2 Indirect measures

A second approach determines the amount of inequality of opportunity indirectly by comparing the inequality in the actual distribution of income, $I(y)$,

[14] In case $N_{k.}^{\overline{a}^R} = \emptyset$ this procedure runs into difficulties. If Roemer's identification axiom is used to identify effort (see section 4.2.1) this does not occur.

[15] As it is only sensitive to the lowest income for each effort, the index does not satisfy EPC as defined in section 2. It is easy to formulate a leximin extension that satisfies it, however.

[16] Schokkaert el al. (1998) already applied a similar procedure in a health context.

to the inequality in a counterfactual income distribution where there is no inequality of opportunity $I\left(y^{E0}\right)$. This results in the measure

$$\Theta_I\left(y, y^{EO}\right) = I\left(y\right) - I\left(y^{EO}\right). \tag{9}$$

Almost all applications of indirect measures to inequality of opportunity construct a counterfactual income distribution that eliminates all inequality between individuals having the same effort. As such, they are measures of ex-post inequality of opportunity, but, remember that when effort is distributed independently of type, absence of inequality of opportunity ex-post implies equality of opportunity ex-ante. We show that for each of the counterfactuals listed in the previous subsection, there exists a dual counterfactual in the indirect approach that implies ex-post equality of opportunity.

Consider first the dual counterfactuals associated with ex-ante approaches in section 3.1. The dual counterfactual to (2) was proposed by Checchi and Peragine (2010): they construct the counterfactual

$$y_k^{EO1} = \frac{1}{|N_{.k}|} \sum_{i \in N_{.k}} y_i, \tag{10}$$

which replaces every income by the average income of those sharing the same efforts and compute (9), using the mean log deviation as inequality measure. Evidently, (10) expresses the idea of utilitarian reward. It is straightforward to provide an alternative, based on inequality averse reward, by defining the dual to (3):

$$y_k^{EO2} = \frac{1}{|N_{.k}|} \sum_{i \in N_{.k}} i\widetilde{y}_i.$$

Also the duals to the parametric ex-ante approaches can be used to define counterfactuals implying ex-post equality of opportunity: the first inspired by (4), the second by (5):

$$y_k^{EO3} = \widehat{g}^R\left(a_k^R, 0\right),$$
$$y_k^{EO4} = \frac{1}{|N_{.k}|} \sum_{i \in N_{.k}} \widehat{g}\left(a_i^C, a_i^R, 0\right).$$

To estimate the relevant equations in both cases we need observations on $a^R$. The second alternative has the advantage of dealing with the correlation between circumstances and efforts in a more flexible tranche-dependent way.

Next, consider the duals based on the counterfactuals in the direct ex-post approach. To obtain the dual to Roemer's direct ex-post non-parametric approach (6) and (7), fix a reference value for circumstances, $\overline{a}^C$. Next define

$$y_k^{EO5}\left(\overline{a}^C\right) = \frac{1}{\left|N_{.k}^{\overline{a}^C}\right|} \sum_{i \in N_{.k}^{\overline{a}^C}} y_i,$$

the average income of those that have the same responsibility vector as individual $k$ and have the reference value for circumstances. In the vector $y^{EO5}$ everybody with the same responsibility vector has the same income, such that there is full ex-post equality of opportunity. To eliminate the dependence of the resulting measure of inequality of opportunity on the choice of reference circumstances, $I\left(y^{EO}\right)$ in (9) can be replaced by the averaged inequality index

$$\frac{1}{n}\sum_{l=1}^{n} I\left(y^{EO5}\left(a_l^C\right)\right).$$

We have not yet seen anyone suggesting this approach.

The dual to (8) is due to Bourguignon et al. (2007): fix a reference value for the circumstance variable, $\overline{a}^C$ to obtain

$$y_k^{EO6}\left(\overline{a}^C\right) = \widehat{g}\left(\overline{a}^C, a_k^R, e_k\right). \tag{11}$$

They use the Theil index as inequality measure, Pistolesi (2009) uses, in addition, the mean log deviation, the half squared coefficient of variation and the standard deviation of logs. Also here $e_k$ can be set equal to zero or to its estimated value $\widehat{e}_k$. The former treats it as a circumstance with reference value zero, the latter as an effort. Most authors take the mean value for circumstances in the sample as the reference value $\overline{a}^C$. Again this choice can be criticized for being arbitrary. This can be overcome by replacing $I\left(y^{EO}\right)$ in (9) with the averaged inequality index

$$\frac{1}{n}\sum_{l=1}^{n} I\left(y^{EO6}\left(a_l^C\right)\right).$$

We are unaware of any application of such a parametric aggregate indirect inequality of opportunity measure.

All the above approaches rely on counterfactuals ensuring ex-post equality of opportunity and thereby entail ex-ante equality of opportunity when efforts are distributed independently of type. But even then, however, ex-post equality of opportunity is not necessary for ex-ante equality of opportunity. In the literature there is only one proposal that assigns to individuals opportunity sets of equal value, without imposing full ex-post equality of opportunity. This proposal is the non-parametric proposal by Checchi and Peragine (2011), which evaluates individual's opportunity sets by (2) and constructs the counterfactual

$$y_k^{EO7} = y_k \frac{\mu\left(y\right)}{y_k^{c1}}, \tag{12}$$

where $\mu\left(y\right)$ is mean income of vector $y$ such that everybody's income is scaled up or down by the ratio of average income and the value of his opportunity set as measured by (2). Observe that $\frac{1}{|N_{k.}|}\sum_{i\in N_{k.}} y_i^{EO7} = \mu\left(y\right)$, such that, when opportunity sets are measured as in (2), in distribution $y^{EO7}$ everybody has indeed an opportunity set of the same value. They use the mean log deviation as inequality measure. Evidently, this procedure can be applied when opportunity sets are valued differently, like, e.g., when they are valued according to (2). The corresponding counterfactual becomes

$$y_k^{EO8} = y_k \frac{\mu(y)}{y_k^{c2}}.$$

## 3.3 Stochastic dominance

The stochastic dominance approach to the measurement of inequality of opportunity originates from the ex-ante framework.[17] In our discussion in section 3.1, we have seen two non-parametric measures of the value of an opportunity set, (2) and (3), the first being inspired by utilitarian reward, the second by inequality averse reward. In both cases, the value of an individual's opportunity set is an increasing function of the outcomes obtained by those that belong to his type. This is an uncontroversial starting point for an ex-ante approach and suggests that ex-ante inequality of opportunity can be established as soon as some type's cumulative distribution function of income first order stochastically dominates another type's cumulative distribution function. Hence the absence of first order stochastic dominance between type's cumulative distribution functions can be seen as a test for ex-ante equal opportunities. Formally, let, for all $i \in \left\{1, \ldots, m^C\right\}$, $F_i(y)$ denote the cumulative distribution function of income of type $i$. A weak test of ex-ante equality of opportunity tests the following condition.

**AFOSD** (Absence of First Order Stochastic Dominance): there does not exist $i, l \in \left\{1, \ldots, m^C\right\}$, such that, for some $y \in \mathbb{R}_+ : F_i(y) < F_l(y)$ and for all $y \in \mathbb{R}_+ : F_i(y) \leq F_l(y)$.

If one adheres to an inequality averse reward principle, one can go further. In that case, as advocated persuasively by Lefranc et al. (2009), absence of first order stochastic dominance can be strengthened to the requirement of absence of second order stochastic dominance between types' cumulative distribution functions.

**ASOSD** (Absence of Second Order Stochastic Dominance): there does not exist $i, l \in \left\{1, \ldots, m^C\right\}$, such that, for some $y \in \mathbb{R}_+ : \int_0^y F_i(\widetilde{y})\, d\widetilde{y} < \int_0^y F_l(\widetilde{y})\, d\widetilde{y}$ and for all $y \in \mathbb{R}_+ : \int_0^y F_i(\widetilde{y})\, d\widetilde{y} \leq \int_0^y F_l(\widetilde{y})\, d\widetilde{y}$.

## 3.4 Norm based measures

We know from proposition 2 that liberal reward and ex-post compensation are incompatible. The axiomatic literature on (opportunity) fair allocations pro-

[17] In section 4.2.1, we will see that under a frequently made hypothesis in empirical work (Roemer's identification axiom) the tests developed here become also relevant from an ex-post perspective.

ceeded by characterizing first best redistribution mechanisms that satisfy weakened versions of the principles -see, Fleurbaey (2008) for an overview. Such redistribution mechanisms assign to every individual, as a function of his circumstances and efforts, an income in such a way that both liberal reward and ex-post compensation are to some extent satisfied. As shown by Devooght (2008) and Almas et al (2011), these (partial) solutions to the liberal reward / ex-post compensation dilemma can be incorporated in a measure of equality of opportunity or, in their language, a measure of offensive or unfair income inequality, respectively. The idea is to treat the level of income that these rules assign to a particular individual as the norm that he should get, and measure offensive inequality by the distance between the actual income vector $y$ and the norm income vector $y^n$. Formally, one computes

$$I\left(y, y^n\right), \tag{13}$$

where the function $I(\cdot,\cdot)$ has to satisfy at least two requirements. First, since it matters how far each individual is from *his* norm income, the measure must satisfy partial symmetry (i.e. be invariant to permutations of $(y_k, y_k^n)$ pairs), but not full symmetry (where different permutations can be applied to the vectors $y$ and $y^n$). Second, due to the heterogeneity of the population in terms of compensation and responsibility characteristics, the usual transfer principle does not apply. These arguments induce Devooght (2008) to propose Cowell's (1985) measure of distributional change, a special case of which is the generalized entropy class. Measures of distributional change have the property that a transfer from a rich to a poor person decreases the value of the measure if and only if the ratio of the actual income of the rich and poor person is larger than the ratio of their norm incomes. Almas *et al.* (2011) define unfair treatment of each individual as the absolute value of the difference between his actual income and norm income and propose an unfairness Gini to aggregate these differences. Here, a transfer from a person who is less unfairly treated to a person who is more unfairly treated diminishes the value of the index.

Devooght takes the egalitarian equivalent allocation, first suggested in the equality of opportunity context by Bossert and Fleurbaey (1996), as the norm. Almas et al. take in the main part of their analysis the generalized proportionality allocation, first proposed by Bossert (1995), as the norm.[18] As a final remark, the computation of the norm incomes proposed by Devooght and Almas requires estimation of the outcome function, $\widehat{g}\left(a_k^C, a_k^R, e_k\right)$. To compute the norm, in both papers, the $e_k$ is replaced by its estimated value $\widehat{e}_k$.

Other first-best redistribution mechanisms exist that do not require the estimation of $\widehat{g}\left(a_k^C, a_k^R, e_k\right)$ and can be computed non-parametrically -see, e.g., the

[18] They do sensitivity analysis and report results for two versions of the egalitarian equivalent norm –which requires the choice of reference circumstances– and also for the conditional equality norm –which requires the choice of reference efforts. In both cases, without much argument, the reference is set equal to its average value in the sample. Their empirical results appear insensitive to the choice of norm distribution, but it is unclear whether the choice of the reference value matters or not.

observable average conditional egalitarian and the observable average egalitarian mechanism proposed in Bossert et al (1999). They have not yet been used in the norm based approach and can be combined with any inequality measure that satisfies partial symmetry and does not satisfy the usual transfer principle (like the unfairness Gini, the generalized entropy or the divergence measures discussed by Magdelou and Nock (2011)) to obtain valid non-parametric alternatives for the norm based approach.

### 3.5 Overview

Table 1 summarizes our survey of approaches to the measurement of inequality of opportunity. Six observations follow from our survey.

A first observation is that we propose several new measures. New indirect ex-post measures ($y^{EO2}$, $y^{EO3}$, $y^{EO5}$) are generated by constructing counterfactuals with complete ex-post equality on the basis of the counterfactuals used in the direct approach. A new parametric measure of direct ex-ante ($y^{c4}$) and its dual indirect ex-post measure ($y^{EO4}$) combines features of the non-parametric approach ($y^{c1}$ and $y^{EO1}$, respectively) and the parametric approach ($y^{c3}$ and $y^{EO3}$, respectively). We also showed how Checchi and Peragine (2010)'s indirect ex-ante approach can be adjusted to deal with inequality averse reward in $y^{EO8}$. We argued that for the approaches that require the choice of a reference value for either efforts ($y^{c5}$ and $y^{c6}$) or circumstances ($y^{EO5}$ and $y^{EO6}$) should receive more attention. Roemer's averaged inequality of opportunity measure may overcome the arbitrariness of the choice of reference value to some extent. Finally, we pointed out that the norm income approach can be applied non-parametrically by using the observable average egalitarian equivalent or the observable average conditional egalitarian allocation mechanisms, proposed by Bossert et al (1999).

A second observation is that many different inequality measures have been used, often without much justification. The only exceptions are in the norm based and in the direct measurement approach. In the former, an inequality measure that replaces the standard transfer principle by a more suited transfer principle and satisfies partial symmetry is necessary. In the latter, an infinite inequality aversion was motivated from the normative point of view that all inequalities that are due to differences in circumstances are unacceptable. We believe that this argument is a powerful one for welfare measurement, but is less convincing for measuring inequality of opportunity as it ignores most inequalities. The indirect approach is often used to answer the question to which extent income inequality is due to inequality of opportunity. This is a meaningful question for any plausible measure of income inequality, but for true opportunity egalitarians, those concerned with equality of opportunity rather than equality of outcome, the aswer to the question is irrelevant. Sometimes additional arguments can be used to single out a particular measure or sets of measures. For instance, Checchi and Peragine (2010) and Ferreira and Gignoux (2011) motivate the use of the mean log deviation by pointing out that it is the only decomposable inequality measure that is path independent (Foster

Table 1: Approaches to the measurement of inequality of opportunity

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | Direct and Indirect Measures | | | |
| | | | Direct Ex-Ante | | | Indirect Ex-post |
| NP | $y^{c1}$ | $\infty$ IA | Van de gaer (1993) | $y^{EO1}$ | MLD | Peragine and Checchi (2010) |
| | | RDM | Aaberge et al. (2011) | | | |
| | $y^{c2}$ | Gini | Lefranc et al. (2008) | $y^{EO2}$ | | |
| P | $y^{c3}$ | MLD | Ferreira and Gignoux (2011) | $y^{EO3}$ | | |
| | $y^{c4}$ | | | $y^{EO4}$ | | |
| | | | Direct Ex-Post | | | |
| NP | $y^{c5}$ | $\infty$ IA | Roemer (1993) | $y^{EO5}$ | | |
| | | RDM | Aaberge et al. (2011) | | | |
| P | $y^{c6}$ | Set1 | Pistolesi (2009) | $y^{EO6}$ | Theil | Bourguignon et al. (2007) |
| | | | | | Set1 | Pistolesi (2009) |
| | | | | | | Indirect Ex-Ante |
| NP | | | | $y^{EO7}$ | MLD | Peragine and Checchi (2010) |
| | | | | $y^{EO8}$ | | |
| | | | Stochastic Dominance (Ex-ante) | | | |
| | | | First Order | | | Second Order |
| NP | O'Neill et al. (2001) | | | Lefranc et al. (2008, 2009) | | |
| | Lefranc et al. (2008, 2009) | | | | | |
| | | | Norm Based (Ex-post) | | | |
| NP | $y^n$ | Set2 | Observable average egalitarian equivalent allocation | | | |
| | $y^n$ | Set2 | Observable average conditional egalitarian allocation | | | |
| P | $y^n$ | DC | Egalitarian equivalent allocation | | | Devooght (2008) |
| | $y^n$ | Gini | Generalized proportional allocation | | | Almas et al. (2011) |

Note 1: NP=non-parametric; P=Parametric

Note 2: $\infty$ IA: Infinite Inequality Aversion; RDM: Rank Dependent Mean. MLD: Mean Logarithmic Deviation; Set1: MLD, Theil, half squared coefficient of variation and standard deviation of Log of income; Set2: any inequality measure satisfying partial symmetry and a weak but not strong transfer principle and Shneyerov, 2000), which implies that non-parametric direct and indirect approaches yield the same results.[19] Pistolesi (2009) uses for the direct measurement approach a whole set of inequality measures, as his main concern is to compare direct and indirect parametric methodologies.

A third observation is that the stochastic dominance approach is by its very nature non-parametric. We motivated it so far from an ex-ante point of view, but in section 4.2.1, proposition 5, we will argue that, if Roemer's identification axiom is assumed, rejection of the absence of first or second order stochastic dominance implies ex-post inequality of opportunity.

A fourth observation is that norm based approaches have only been applied using the income allocations from the axiomatic literature concerned with ex-post inequality of opportunity as the norm distribution. The counterfactuals used in the indirect approach can also be used as the norm income distribution. Using either $y^{EO7}$ or $y^{EO8}$ yields a norm based on ex-ante equality of opportunity without requiring ex-post equality of opportunity when efforts are distributed independently of type.

Fifth, it is important to realize that the indirect approach cannot be interpreted as a norm based approach. In the norm based approach it crucially matters who gets what, while in the indirect approach this is not the case, as different permutations can be applied to $y$ and $y^{EO}$ in (9). This makes the indirect approach unattractive as a normative measure of inequality of opportunity. It can be used to decompose income inequality into inequality that is due to circumstances and efforts, but only if a suitable inequality measure is chosen, such that the inequality that is due to circumstances is a meaningful direct estimate of inequality of opportunity -see the discussion following equation (3).

Finally, as especially the previous observations make clear, the theoretical basis for many of the inequality measures that are used or can be thought of, remains rather weak. A lot of work remains to be done to sort out the attractive from the unattractive ones.

# 4 Data imperfections

In this section we confront some important problems facing the application of the framework in the previous section: how to choose and measure circumstances, how to measure efforts and the consequences of imperfectly measuring circumstances or efforts.

## 4.1 Circumstances

Measured inequality of opportunity crucially depends on the set of circumstances chosen. The larger the set of circumstances, the larger the inequality of opportunity.[20] Thus, a proper selection of the circumstances is paramount.

[19] From the discussion in the paragraph below equation (3), it follows that for the mean log deviation, $\theta_I\left(y, y^{EO7}\right) = I\left(y\right) - I\left(y^{EO7}\right) = I\left(y^{c1}\right)$.

[20] Notice, thought, that if there is a negative correlation between the 'new' circumstances, which were previously classified as effort, and the 'old' circumstances, which were already

Often researchers are limited by the scarcity of data on circumstances beyond basic individual characteristics and family background, such that most empirical studies are confined to a small set of basic circumstances. We discuss this issue in section 4.1.2. In principle, the set of circumstances that should be included follows from the answer to the question what should individuals be held responsible for. This is taken up next.

### 4.1.1 Selection of circumstances

Three prominent views can be found in the political philosophy literature on the difficult and unsettled question what people are responsible for.

A first view argues that individuals ought to be held responsible only for what lies within their control –as defended, *inter alia*, by Arneson (1989), Cohen (1989), and Roemer (1993, 1998a). Control is related to the recognition of free will, the existence of which is sometimes disputed. Those who deny the existence of free will, such as the hard determinists, take an extreme position and include nearly all observables in the circumstance set and consider almost all inequalities as unfair. Most empirical studies, however, adopt a possibilist criterion, which is consistent with the existence of free will, and classifies as circumstance family background variables, such as parental education or occupation, individual characteristics, such as gender, ethnicity or age, and innate characteristics, such as IQ. Under this view, contextual variables such as access to basic services, e.g. clean water, sanitation, electricity or transportation, should also be included in the circumstance set.

A second approach contends that individuals ought to be held responsible for their preferences and the ensuing choices –as advocated, *intera alia*, by Rawls (1971), Dworkin (1981a, 1981b), Van Parijs (1995) and Fleurbaey (2008). Under this view, the set of circumstances gets reduced to a minimal set of variables including innate characteristics or traits such as talent or beauty.[21] In contrast, variables such as gender or ethnicity, which are typically classified as circumstances in empirical analyses, should belong to the realm of responsibility if the differential effect they bring about reflects exclusively differences in preferences, i.e. are not the result of discriminatory treatment.[22]

---

included in the set, opportunity inequality may decrease, and not increase (Cappelen and Tungodden, 2006).

[21]Since Hamermesh and Biddle (1994) we know that physically attractive workers obtain sizable rents from their looks. More recently, Mobius and Rosenblat (2006) identified three transmission channels for such beauty premium.

[22]As Fleurbaey (2008) persuasively explains, under the believe that free will exists, the control approach comes very close to the preference approach to responsibility, as genuine control is "typically defined in terms of choices reflecting authentic preferences" (p. 250). In addition, the preference approach may be extended to hold people responsible for any preference or characteristic which they endorse, i.e. which they would have chosen were they in control. Notwithstanding all this, he goes on to argue, the two approaches may yield substantively different conclusions when advantage results from preferences, which have not been chosen in any sense and are not endorsed by the individual. Since control, choice and endorsement are very hard to observe, it is very difficult to test empirically whether the control and the preference approach are close or far from each other.

In line with Nozick (1974)'s self-ownership argument,[23] a third view considers that individuals are entitled to the products of all personal characteristics, including genetic ones such as innate talent. This leads to the other extreme position where the set of circumstances is empty, and all inequalities are legitimate. There is no room for equality of opportunity in this view.

#### 4.1.2 Unobserved circumstances

As soon as we agree that there are circumstances for which people should be compensated, we enter the realm of equality of opportunity and need to measure these circumstances: application of equal opportunity theories without observing any circumstances is impossible. In practice, measuring circumstances is easier than measuring efforts and different datasets can be combined to obtain a more comprehensive set of circumstances, as in Ferreira et al (2011). Even then an exhaustive list of circumstances is typically not available, however. Assume that we have directly observed the relevant efforts, but did not observe all relevant circumstances. In that case, the partitioning of the population in true types is a finer partitioning than the one on the basis of observed types and the outcomes of the observed types is a weighted average of outcomes conditioned on true types with weights determined by the population frequency of the true types in the observed types. This leads to a downward bias in ex-post inequality of opportunity, as the inequality within the columns based on observed types is smaller than in the columns based on the true types. Similarly, as the rows associated with the observed types are weighted averages of the rows associated with the true types, unobserved circumstances lead to an underestimation of ex-ante inequality of opportunity. Ferreira and Gignoux (2011) in particular stress that estimates of inequality of opportunity based on an incomplete list of circummstances should be interpreted as a lower bound of true inequality of opportunity.

#### 4.1.3 Contribution of different circumstances to inequality of opportunity

Consider the indirect measurement approach (see section 3.2), which determines the amount of income inequality that remains when there is no inequality of opportunity left. The Bourguignon et al. (2007) approach determines this counterfactual income distribution as the one that results when everyone has the same reference circumstances -see (11). By not equalising all circumstances at once Bourguignon et al. (2007), show that it is possible to estimate the partial effect of one (or a set) of circumstance variables $J$, controlling for the others ($j \neq J$). Following their specification of the function $g\left(a_k^C, a_k^R, e_k\right)$, let

$$\ln y_k = \beta^C a_k^C + \beta^R a_k^R + u_k,$$

[23] The self-ownership argument states that individuals own themselves and thus have a legitimate claim over the products of their talents and abilities.

and construct alternative counterfactual distributions

$$y_k^{EO(J)} = \exp\left[\widehat{\beta}^J \overline{a^{CJ}}_k + \widehat{\beta}^{j \neq J} a_k^{Cj \neq J} + \widehat{\beta^R} a_k^R + \widehat{e}_k\right],$$

where $\overline{a^{CJ}}_k$ is the vector of reference values of the circumstances in set $J$ and $a_k^{Cj \neq J}$ the vector of actual circumstances of individual $k$ of the circumstances in the complement of the set $J$. This allows to compute inequality of opportunity due to a given (set of) circumstance(s), $J$ in spirit of the indirect ex-ante parametric approach by replacing $y^{EO}$ in (9) by $y^{EO(J)}$ defined above. To compute each circumstance's contribution to overall inequality one can use the Shapley decomposition (Shorrocks, 1999), which avoids the path dependency problem whereby results are sensitive to the ordering in which circumstances are put at their reference value in the analysis. This approach has become quite popular recently (see, e.g. Björklund et al. (2011)). Usually, the mean value of the circumstance characteristic is taken as the reference value.

## 4.2 Constructing measures of effort

To apply ex-post compensation, we need to identify individuals' efforts in a normatively attractive way. Effort variables are shaped by circumstances. Preferences and tastes, for instance, are partly shaped by family background. Whether we should correct for this is closely related to the answer to the question what people are responsible for (see subsection 4.1.1). Those defending responsibility for preferences (and the resulting choices) will typically argue that it does not matter where these preferences come from, as long as people identify with them. Those defending responsibility by control (like Roemer (1993, 1998a and 1998b) argue that, as people do not control their circumstances, raw effort variables should be cleaned to obtain normatively relevant efforts. This view is dominant in most empirical applications to date. We discuss four different procedures used in the literature to construct normatively relevant effort(s).

### 4.2.1 Unobservable effort, non-parametric identification

If no effort variables are observed, the lack of data can only be overcome with some auxiliary hypotheses. The most elegant and frequently used comes from John Roemer (1993), and is stated as follows.

**RIA** (Roemer's Identification Assumption): those that are at the same percentile of the distribution of income conditional on their type have exercised the same degree of effort.

This assumption allows us to take the percentile within the income distribution of an individual's type as the normatively relevant measure of his effort. By construction effort is distributed uniformly over $[0, 1]$ for all types and consequently independently distributed of type.

RIA can be derived from more fundamental hypothesis about the income generating process and the distribution of circumstances and effort. More in particular, as pointed out by Fleurbaey (1998, p.221), RIA assumes that (A1) the multi-dimensional effort variables $a_i^R$ can be aggregated into a scalar measure of responsibility $a_i^r$ in such a way that with every value for $a_i^R$ corresponds exactly one value for $a_i^r$ and that income is a strictly increasing function of $a_i^r$ and (A2) $a_i^r$ is distributed independently of $a_i^C$. As argued by Roemer, while (A2) is, within the responsibility by control view, a natural assumption for normatively relevant effort, assumption (A1) is very strong.

The assumption is very powerful, for, thanks to RIA, the equality of opportunity framework becomes operational even when effort is unobservable: one only needs to compute the cumulative distribution of incomes conditional on types, and equate the percentile corresponding to each individual's income within the cumulative distribution of his type to his level of effort. This allows the construction of a new matrix $\widetilde{Y}^R$, to which one can apply all the ideas mentioned in section 2 and all the non-parametric procedures from table 1. Observe that, by construction, all elements in the same row of $\widetilde{Y}^R$ contain the same number of individuals.

The plot of the inverse of the cumulative income distributions conditional on types gives for each percentile the corresponding income level. If the plots for two types differ at some percentile, we have ex-post inequality of opportunity: for the same degree of responsibility, two individuals of different types receive a different level of income. Fixing the percentile value and reading the corresponding income values for all types is like looking at a particular column in the matrix $Y$ of the previous section; it amounts to taking an ex-post perspective. Alternatively, looking at the cumulative distribution function for each type is very much like looking at a row in the matrix $Y$, with a continuous effort variable. Hence the cumulative distribution function also provides the information necessary for an ex-ante perspective.

The above insights provide the basis for analyzing conditional distribution functions from a perspective of equality of opportunity. At the one hand, with unobservable effort (and RIA), ex-post equality of opportunity is satisfied if and only if the following property holds true.

**ECDF** (Equal Cumulative Distribution Functions): for all $i, l \in \left\{1, \ldots, m^C\right\}$ and for all $y \in \mathbb{R}_+ : F_i\left(y\right) = F_l\left(y\right)$.

At the other hand, ex-ante equality of opportunity directly leads to the requirement of the absence of first order stochastic dominance between types' cumulative distribution functions, i.e. condition AFOSD in section 3.3. We now have recovered the result that since effort is distributed independently of type, full equality of ex-post opportunities implies full equality of ex-ante opportunities, stated in the first paragraph of section 3.

**Proposition 5**: Accepting RIA, ex-post equality of opportunity implies ex-ante equality of opportunity.

When RIA is imposed, testing whether (AFOSD) can be rejected can thus be interpreted as a weak test of ex-post equality of opportunity (ECDF).

Finally, suppose that we apply RIA and use the percentile within type as a measure of effort to construct the matrix $\widehat{Y}$, but we only condition the cumulative distribution functions on the vector of observable circumstances $a^{CO}$, not on the entire vector of circumstances $a^C = \left[a^{CO}, a^{CU}\right]$, where $a^{CU}$ are unobserved circumstances. Strong assumptions are necessary to relate the results obtained on the basis of the constructed matrix $\widehat{Y}$ to the true matrix $Y$. To see this, consider the simple case where $a^{CO}$ and $a^{CU}$ are one-dimensional. Moreover, assume that $a^{CU}$ is either $\underline{a}^{CU}$ or $\overline{a}^{CU}$. In that case,

$$F\left(y \mid a_i^{CO}\right) = F\left(y \mid a_i^{CO}, \underline{a}^{CU}\right) p_i\left(\underline{a}^{CU}\right) + F\left(y \mid a_i^{CO}, \overline{a}^{CU}\right) p_i\left(\overline{a}^{CU}\right),$$

where $p_i\left(\underline{a}^{CU}\right)$ and $p_i\left(\overline{a}^{CU}\right)$ are the fraction of the observations with $a^{CO} = a_i^{CO}$ that have $a^{CU} = \underline{a}^{CU}$ and $\overline{a}^{CU}$, respectively. The cumulative distribution function $F\left(y \mid a_i^{CO}\right)$ serves as the basis to identify effort for observable type $a_i^{CO}$, and is a weighted average of the cumulative distribution functions of true types, $F\left(y \mid a_i^{CO}, \underline{a}^{CU}\right)$ and $F\left(y \mid a_i^{CO}, \overline{a}^{CU}\right)$. The only case in which the percentile of $F\left(y \mid a_i^{CO}\right)$ provides correct information on the percentiles of the true types is when $F\left(y \mid a_i^{CO}, \underline{a}^{CU}\right) = F\left(y \mid a_i^{CO}, \overline{a}^{CU}\right)$, meaning that, after conditioning on observed circumstances, the unobserved circumstance does not affect outcomes. In all other cases, effort will be wrongly identified and, the larger the effect of the omitted circumstance on the true conditional cumulative distribution functions, the less representative identified effort is for true effort. We summarize this point in the following proposition.

**Proposition 6**: Accepting RIA, omitted circumstances induce wrong identification of effort unless the unobserved circumstances, after conditioning on observed circumstances, no longer affect income.

### 4.2.2 Unobservable effort, parametric identification

The non-parametric methodology of the previous subsection allows to identify each individual's (normatively relevant) effort as the percentile within the income distribution of his type. Clearly, this approximation works well only if every type contains a substantial number of individuals. If not, a parametric methodology can be a better alternative.

With unobservable effort, Björklund et al. (2011) allow the distribution of effort conditional on type to have different variances, as initially suggested by Roemer (1998a). They assume that effort has two components: a type specific component, $\eta_k^i$, whose variance $(\sigma_i^2)$ differs across types $i$ and which captures the part of effort that is correlated with circumstances, and a second component, $\omega_k$, with a homogeneous variance, $\sigma^2$. The latter is defined as a standardization of the former, $\omega_k = \eta_k^i/(\sigma_i^2/\sigma^2)$, so that the income generating process can be written as:

$$\ln y_k = \beta^C a_k^C + \eta_k^i = \beta^C a_k^C + \widetilde{\eta}_k^i + \omega_k, \tag{14}$$

where $\widetilde{\eta}_k^i = \left(\eta_k^i - \omega_k\right)$ measures the influence of circumstances on the conditional variation of the outcome around the expected value for each type, $i$. The term

$\widetilde{\eta}_k^i$, then, captures the indirect effect of circumstances, while $\omega_k$ is assumed to capture 'pure' effort.

Notice that the econometric error terms are lumped together with efforts, implying that everything that traditionally enters the error term (specification error, omitted variable bias) determines measured effort. Roemer's approach is non-parametric, such that it does not suffer from specification errors (unless the assumed one-dimensionality of effort is considered to be a specification error), but omitted variable bias (circumstances) is also for the Roemer approach an issue (see proposition 6).

#### 4.2.3 Unobservable effort, panel data and parametric identification

Consider the case where no efforts are observed, but the researcher has access to panel data. In this case, Salvi (2007) shows that detailed modeling of the income generating process can be helpful. She exploits the longitudinal features of panel data to distinguish between time-varying and time-invariant circumstances and efforts. Efforts are assumed unobservable and divided into individual traits that do not change over time ($a_k^R$), such as skills, preferences, aspirations or individual talents, and the exertion of effort, which is time-varying ($a_{kt}^R$). Individual traits, $a_k^R$, are modeled as unobservable time-invariant individual effects, while the exertion of effort, $a_{kt}^R$, cannot be distinguished from the idiosyncratic error term, $\upsilon_{kt}$. Circumstance variables are also broken down into time-varying ($a_{kt}^C$) and time-invariant ($a_k^C$), and are assumed observable. Thus, the income variable is modeled as:

$$\ln y_{kt} = \underbrace{\alpha_1 a_{kt}^C}_{\text{t-v circ.}} + \underbrace{\alpha_2 a_k^C}_{\text{t-inv circ.}} + \underbrace{a_k^R}_{\text{ind. traits}} + \overbrace{\underbrace{a_{kt}^R}_{\text{exertion}} + \underbrace{\upsilon_{kt}}_{\text{brute luck+white noise}}}^{\text{error term, } \varepsilon_{it}}. \tag{15}$$

Individual traits, $a_k^R$, are allowed to be correlated with circumstances, i.e. circumstances may affect the individual preferences and aspirations but not her level of effort exertion, which is supposed to be orthogonal to circumstances.

Using the estimates $\left(\widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{a}_k^R, \widehat{\varepsilon}_{kt}\right)$ of equation (15), Salvi proceeds to compute a counterfactual distribution similar to (11) by setting (all) circumstances at the sample mean value $\overline{a}_{kt}^C$ and $\overline{a}_k^C$:

$$y_{kt}^{EO} = \exp\left[\widehat{\alpha}_1 \overline{a}_{kt}^C + \alpha_2 \overline{a}_k^C + \widehat{a}_k^R + \widehat{\varepsilon}_{kt}\right].$$

She estimates inequality of opportunity by means of (9); her approach is indirect ex-post parametric. Notice that, as in Björklund et al. (2011), the econometric error terms are also lumped together with efforts.

#### 4.2.4 Observable effort correlated with circumstances

A final case considered in the literature occurs when we observe (all) effort variables, but they are correlated with circumstances. Roemer (1993 and 1998b) gives the example of peoples' occupation (circumstance) which affects both the number of cigarettes smoked (observable effort) and health (outcome). As white collar workers smoke on average less than blue collar workers, we cannot equate the number of cigarettes smoked to normatively relevant effort because that would implicitly hold workers responsible for their circumstance (occupation). Hence, to obtain normatively relevant effort, we have to clean the number of cigarettes smoked from the impact of occupation. Roemer suggests the technique described in section 4.2.1, to determine an individual's responsibility as his percentile in his type's distribution of number of cigarettes smoked. There exist evident alternatives. As proposed by Schokkaert el al. (2004) and Bourguignon et al. (2007), a variety of econometric techniques, like regression analysis, can be used to obtain cleaned normatively relevant effort variables. The normatively relevant effort level then becomes the disturbance term in a regression of occupation on number of cigarettes smoked.

Bourguignon et al. (2007), develop this idea and model earnings, $y_k$, as function of effort $\left(a_k^R\right)$ and circumstance $\left(a_k^C\right)$ variables,

$$\ln y_k = \beta^C a_k^C + \beta^R a_k^R + u_k, \tag{16}$$

and let endogenous effort depend on circumstances:

$$a_k^R = H a_k^C + v_k, \tag{17}$$

where $\beta^C$ and $\beta^R$ are parameter vectors, $H$ is a matrix of parameters relating circumstances and efforts, and $v$ and $u$ denote pure random factors. Equation (17) allows correlation between $a_k^C$ and $a_k^R$, and models the indirect effect of circumstances on the outcome through their influence on efforts. This allows the estimation of direct and indirect effects of circumstances on earnings. The counterfactual distribution $y^{EODT}$, which includes both direct and indirect effect of circumstances, may be obtained by using the parameter estimates $\left(\widehat{\beta}^C, \widehat{\beta}^R, \widehat{H}\right)$ and setting the vector of circumstances at the sample mean $\overline{a}^C$ in both equations (16) and (17). To do so we can estimate a reduced form of (16) and (17):

$$\ln y_k^a = \psi a_k^C + \varepsilon_k, \tag{18}$$

where $\psi = \left[\beta^C + \beta^R H\right]$ and $\varepsilon_k = \beta^R v_k + u_k$ resulting in estimates $\left(\widehat{\psi}, \widehat{\varepsilon}_k\right)$ and put

$$y_k^{EODT} = \exp\left[\widehat{\psi}\overline{a}^C + \widehat{\varepsilon}_k\right].$$

An estimate of the counterfactual distribution where only the direct effects of circumstances are equalized, $y^{EOD}$, may also be obtained by estimating (16) and computing

$$y_k^{EOD} = \exp\left[\widehat{\beta}^C\overline{a}^C + \widehat{\beta}^R a_k^R + \widehat{u}_k\right]. \tag{19}$$

Inequality of opportunity indices are then obtained by comparing actual inequality $I(y)$ to the inequality in the simulated earnings distributions, $I\left(y^{EODT}\right)$, i.e.

by computing (9), with an appropriate measure of inequality. The importance of direct and indirect effects can be computed by means of the (simulated) distribution of earnings $y^{EOD}$, where circumstances are equal for the direct effect only.

Using econometric techniques has the advantage that it allows the researcher to take more than one effort variable into account; contrary to Roemer's approach effort can be truly multi-dimensional. As compared to Roemer's non-parametric methodology, the methodology in this subsection relies on functional form assumptions to obtain inequality of opportunity estimates (but see Pistolesi (2009), discussed in section 5.2, for a semi parametric density estimation technique to obtain the relevant counterfactual distributions). Three reasons may justify such cost. First, controlling for circumstances in a multivariate regression framework uses data more efficiently, and allows for finer categories. As the vector of observed circumstances becomes larger (and the number of categories within each variable increases) the number of types and tranches grow exponentially, which leads to type-tranche combinations with very few (possibly zero) observations, such that sampling variances are very large, and estimates become unreliably imprecise[24]. Second, the above problem is even more severe when (some) circumstances are continuous variables. Clearly, there exist non-parametric techniques like kernel density estimation that have already been used and allow one to deal with continuous circumstances (see section 5.3), but these techniques require large data sets to yield reliable estimates. As a result, in case one only has small datasets, parametric approaches become an attractive alternative. Third, as explained in section 4.1.3, the parametric methodology permits the estimation of the partial effect of one (or a set) of the circumstance variables $J$, controlling for the others ($j \neq J$), such that we can compute inequality of opportunity due to a given (set of) circumstance(s), $J$. However, also in this approach, the econometric error terms are lumped together with efforts, implying that everything that traditionally enters the error term (specification error, omitted variable bias) determines measured effort. In the next section, we reflect on how to deal with error terms.

## 4.3 Luck and error terms

In section 3 we introduced omitted variables $u_k$ and random variables $e_k$ in the analysis. In practice, $u_k$ captures the effects of omitted circumstances and efforts, while specification errors heavily affect $e_k$. Given the diversity of the samples and econometric techniques used, it becomes difficult in general to say much about the importance of error terms. Nevertheless, the following observations can be made.

Due to data limitations most empirical studies include a limited set of circumstances in their list. Virtually all studies include a measure of social background luck (parental income, parental education). Very few surveys have ob-

[24] In such cases, the proposal made in Donni et al. (2012), to use a latent class technique that endogenously determines the types (and number of types) provides a way out for non-parametric methodologies.

servations on genetic luck. An exception is Björklund et al. (2011): they find IQ, even though measured at the age of 18 considered to be a prominent exponent of genetic luck, to be the most influential factor behind inequality of opportunity in Sweden. This suggests that, if genetic luck is not included in the list of circumstances, genetic luck can be an important contributor to the error term. We are unaware of forms of brute luck or option luck being included in the list of circumstances such that they always enter the error terms. As it is usually claimed (see section 2.3) that genetic luck should be fully compensated, some compensation is due for brute luck, and we cannot know what part of the error term should be included as a circumstance, the argument seems to call for some compensation for the effects of luck such that the principle of utilitarian reward (using a full list of circumstances) has to be replaced by inequality averse reward (since one is typically using only a limited list of circumstances).

# 5 Empirical applications

In the light of the discussions in sections 3 and 4, this section reviews a selected sample of studies that apply the concepts and techniques explained in these two previous sections. Our emphasis is not on the results obtained per se but in the methodological choices taken and their bearing on inequality of opportunity comparisons.

## 5.1 Direct measures

Cogneau and Mesplé-Somps (2008) compute (1), to compare ex-ante and ex-post inequality of opportunity in five African countries. The outcome variable is household consumption per head and circumstances are based on fathers' social origins (farmers, non farmers with at most primary education and non farmers with more than primary education) and region of birth. They measure ex-post inequality, identifying effort assuming RIA and using the minimum income relative to the mean as inequality index in (7). This is the inequality index following from Roemer's work. Ex-ante individuals' opportunity sets are valued by average type income (2) and ex-ante inequality is measured by the lowest average type income divided by mean income in the country, which corresponds to the proposal by Van de gaer (1993). As the cumulative distribution functions of different types do not cross, they find that the inequality of opportunity ranking for the five countries does not depend on which of both measures is taken [25]. Taking only fathers' social origins as circumstances, inequality of opportunity is largest in Madagascar, followed by Ghana in 98, Guinea, Uganda, Ivory Coast and Ghana in 88.

[25] This is a well known property of these specific measures. As the cumulative distribution functions do not cross, for a given percentile (level of effort), it will always be the same type that has the lowest income. But then, the average over these lowest incomes has to coincide with the average value of the opportunity set of the worst-off type.

Lefranc, Pistolesi and Trannoy (2008) provide a direct measurement of ex-ante income inequality by computing the Gini opportunity index for (1) with (3) measuring the value of each individual's opportunity set. They compare nine Western countries from the perspective of opportunity equality by comparing the distributions of the pre-tax as well as the net disposable income in these countries for male-headed households aged 25-40, conditional on social background, measured by 3 levels of father's education (professional group for France and Great Britain). Italy has the highest inequality of opportunity, followed by the U.S., Belgium, France, Great Britain, the Netherlands, Norway, Sweden and West Germany. The value of the Gini opportunity inequality index is shown to be decomposable in a return (differences in the means of the conditional distribution) and a risk component (differences in the spread of the conditional distributions). In countries with the lowest value of opportunity inequality, the risk contribution is negative, meaning that inequality of risk mitigates inequality of opportunity: those coming from disadvantaged backgrounds have less risky lotteries than those coming from advantaged backgrounds. In countries with a high inequality of opportunity, risk exacerbates inequality of opportunity. Finally, using the value for their Gini opportunity index, the authors notice that there is a positive correlation (0.67) between inequality of opportunity and inequality of outcomes (measured by the standard Gini coefficient), and that there is no trade-off in the sample between per-capita income and equality of opportunity.

Ferreira and Gignoux (2011) are the only ones that employ the direct ex-ante parametric approach based on (4). They compute direct inequality of opportunity for household income per capita in six Latin American countries, using father's and mother's education, father's occupation, ethnicity and region of birth as circumstances. Ranking inequality of opportunity from high to low, Brazil is followed by Guatemala, Panama, Peru, Ecuador and Columbia. They do the same analysis using the direct non-parametric ex-ante measure based on (2), and find very similar results.

## 5.2 Indirect measures

Checchi and Peragine (2010) compute ex-ante and ex-post inequality of opportunity in Italy using a non-parametric methodology for the indirect approach (9). They apply this framework to gross annual earnings, take family background (measured by highest educational attainment of the parents) as the circumstance variable and compare inequality of opportunity of different subgroups in the population that share similar degrees of labor market conditions and female labor market participation (compare men versus women and Northern versus Central and Southern Italy). In the ex-ante approach average type income (2) measures the value of the opportunity set; the counterfactual is given by (12). In the ex-post approach, effort is identified assuming RIA and inequality of opportunity is thus measured as the difference between the inequality in the actual distribution of income and the inequality in a counterfactual distribution, where individuals are assigned the mean income of their tranche as in (10). The mean log deviation is used as inequality index. Ex-ante inequality of opportunity accounts for about 15 % of total income inequality whereas ex-post inequality of opportunity accounts for 20 %. They also find that inequality of opportunity is highest among women in southern Italy.

Björklund *et al.* (2011) use the parametric approach outlined in section 4.2.2 to estimate ex-ante inequality of opportunity for total pre-tax income in Sweden indirectly and put together the most comprehensive set of circumstances seen so far by employing information from four different datasets for Sweden, which combine survey and register information. In particular, they have information on parental long-run income, parental education, family structure, own IQ and own body mass index, which allows them to split the sample into 1152 types. They estimate the empirical type-specific variances from the OLS residuals of (14). With this large set of circumstances, they find that effort still accounts for 70 percent of income inequality in Sweden. Amongst the circumstances, IQ is the most influential variable, followed by the indirect effect of circumstances as measured by the heterogeneous type-specific variances and parental income.

Salvi (2007) uses the panel data structure outlined in section 4.2.3 to identify effort estimating (16) and computes ex-post inequality of opportunity in Nepal for 1995 and 2003. Unlike other empirical applications, she uses local infrastructure variables (presence of bus service and electric power in the village, and number of secondary schools) as circumstances and finds that family background has a small effect while infrastructure and ethnicity have a large effect. Rather surprisingly, when the MLD (instead of the Theil index) is used to compute (9) estimates are negative. That is, consumption inequality is larger when circumstances are equalized than when they are not, and thus equalizing circumstances would increase consumption inequality.

Bourguignon et al. (2007) developed a model to deal with the situation where efforts are correlated with circumstances –see section 4.2.4– and compute ex-post inequality of opportunity indirectly, using (11) and (9). They use data from urban male hourly earnings in Brazil to quantify the effect on earnings of five circumstances (father's and mother's education, father's occupation, race, and region of birth). The indirect effect of circumstances works through three effort variables (own education, migration out of hometown, and labor market status).[26] Their results show that, first, the group of five circumstances accounts for about a quarter of total earnings inequality. The direct effect explains 60% of the reduction in inequality when circumstances are equalized, the remaining 40% is due to the indirect effect. Second, family background is by far the most important circumstance determining a person's opportunities.

Pistolesi (2009) uses a semi-parametric density estimation approach to perform an analysis similar to that by Bourguignon et al. (2007), where direct and indirect effects of circumstances are separated out. He analyses the evolution of opportunity inequality in the U.S. between 1968 and 2001, using PSID data. The income variable is averaged earnings (over 5 years), circumstances

[26]Labor market status: indicates whether the worker is a formal employee, an employer, an informal employee or self-employed.

are age, years of education of both parents, father's occupation, ethnicity and region of birth. Effort is measured by the variation in schooling attainment and annual working hours unaccounted for by circumstances. Apart from the semi-parametric methodology, which makes it possible for explanatory variables to have different effects at different points in the distribution, this paper makes another important contribution. He computes independently both inequality due to effort (i.e. when circumstances have been equalized) and inequality due to circumstances (i.e. where effort has been equalized at the mean of the estimated conditional distribution of effort). The former allows him to compute (11) and (9), the latter (8) and (1). He finds that the direct and indirect ex-post approaches provide very similar estimates of inequality of opportunity[27]. Finally, it is noteworthy that the econometric modeling shows that circumstances have a major impact on human capital accumulation decisions, but their impact on labor supply decisions is limited.

### 5.3 Stochastic dominance

An application of the use of stochastic dominance is Lefranc, Pistolesi and Trannoy (2008), who compare nine Western countries from the perspective of opportunity equality by comparing the pre-tax and net disposable household income distributions in these countries for male-headed households aged 25-40, conditional on social background (see also the second paragraph of section 5.1). They compare pairwise the cumulative conditional distributions within each country by means of first and second order stochastic dominance and are the first to use rigorous statistical test for stochastic dominance, using the non-parametric stochastic dominance tests developed by Davidson and Duclos (2000). Sweden is the only country for which equality of the conditional cumulative distribution functions cannot be rejected. Then comes West Germany, followed by a group of 3 countries consisting out of Great Britain, Belgium and Norway. In France, Italy, the Netherlands and the U.S., they find second order stochastic dominance relations between all conditional cumulative distribution functions, indicating unequal opportunities between all social background groups. Observe that the stochastic dominance ordering of countries deviates somewhat from the ordering obtained from computation of the Gini opportunity index discussed in section 5.1.

Things are only slightly trickier when (some) circumstances are continuous variables. In this case, one can estimate the conditional distribution using density estimation techniques. O'Neill et al (2000) propose to use a Kernel density estimator to estimate the distribution of income conditional on parental income and use this procedure to depict the incomes of children, conditional on the income percentile of their parents in the US, i.e. the opportunity set of a child whose parent was at a particular percentile in the income distribution of his generation. Individuals whose parents are at the 25th percentile of the income

[27] This holds true for the Theil index, the Gini index, the mean log deviation and the standard deviation of logs. For the half squared coefficient of variation differences are more pronounced.

distribution have to be ranked 70-th (out of 100) to obtain average labor income, while the ranking is only 40-th when parents are at the 75th percentile. The income gap between individuals at the 50-th percentile with parents at the 25-th and 75-the percentile is 56 %. One limitation of this procedure is that it takes only one circumstance variable into account. Nilsson (2005) mends this by using a semi-parametric approach, in which all conditioning variables other than parental income (such as stability of parental relationship, whether the parents were foreign born, characteristic of the parish where the family lived) enter linearly in the children's' labor income generating equation. The resulting opportunity sets can then be drawn conditional on parental income and a specific value for the other circumstances. In his empirical application to Sweden, he finds that male individuals whose parents are at the 25th percentile of the income distribution have to be ranked 52.3 (out of 100) to obtain average labor income, while the ranking is only 42.6 when parents are at the 75th percentile. The income gap between two individuals at the 50th percentiles of their conditional distributions is nearly 10%. Compared to the results from O'Neill et al, the differences in opportunities seem to be smaller in Sweden than in the US.

## 5.4 Norm based measures

Devooght (2008) computes norm based inequality of opportunity taking the egalitarian equivalent solution as the norm and Cowell's measures of distributional change as inequality index. He uses households' pre-tax labor income in a sample of Belgian individuals in 1998. Income is estimated by means of specification (16), and the least favorable value of each circumstance characteristic is taken as reference value in the computation of the egalitarian equivalent norm. The author experiments with different sets of circumstances and different values for the reference level of circumstance characteristics. He concludes that, depending on the choices made in these respects "responsibility-sensitive inequality measurement considers about 90-97.5% of traditionally measured income inequality as offensive" (p. 290), which is much larger than the inequality of opportunity found with *non*-norm based approaches.

Almas *et al.* (2011) compute norm based inequality of opportunity taking the generalized proportionality principle as the norm and a Gini index defined over deviations from the norm as inequality index. The empirical application is based on a large sample of Norwegian citizens. Households' annual labor earnings are estimated as a function of effort and circumstance characteristics, the specification is again of the form (16), and a model enables the imputation of post-tax incomes. They experiment with different sets of circumstance and effort variables. For post tax incomes, when the set of responsibility variables is empty the unfairness Gini (which coincides with the standard Gini) equals 0.205. As the set of responsibility characteristics is extended (to include, successively hours worked, years of education, working in the public sector, county of residence, field of education and age), the unfairness Gini drops (though not monotonically) to 0.152, which means that even with this extensive set of responsibility variables, unfair inequality is about 75 % of total inequality. In addition, they find that, while the standard Gini index of inequality (which takes complete equality as the norm distribution) was lower in 2005 than in 1986, the unfairness Gini index has increased, irrespective of the sets of circumstance and responsibility variables used.

Both Devooght (2008) and Almas *et al* (2011) treat the error terms as a compensation variable, but acknowledge that the error term might contain unobserved effort variables, and so the choice to treat them as compensation variables can be questioned.[28] In another paper, Almas (2008) experiments with the role of the error terms from the estimated equation, treating it as a circumstance variable (leading to an upper bound of unfairness) or as a responsibility variable (leading to a lower bound of unfairness) in the computation of the norm. From a comparison of the US and Germany, using data on pre- and post-tax incomes in the year 2000 they find that for the upper bound of unfairness, Germany is considered to be less unfair than the US whenever the set of compensation variables is not empty, but for the lower bound, the US is less unfair than Germany.

# 6 Conclusion

We have seen that inequality of opportunity theories attempt to combine a compensation principle with a reward principle. Compensation can be done from an ex-ante or an ex-post perspective. Proposition 1 shows that the two are incompatible. The two most common reward principles are utilitarian reward and liberal reward. Proposition 2 shows that the former is incompatible with ex-post compensation and proposition 3 shows that the latter is also incompatible with ex-post compensation. We proposed a third reward principle, inequality-averse reward which can be motivated if, after listing observed circumstances, there remain factors for which compensation is due, like unobserved circumstances or arbitrary property rights or if randomness of incomes within each type lead to the use of a risk-averse evaluation of the possible incomes. Proposition 4 shows that also this reward principle is incompatible with ex-post compensation. As a result, when measuring inequality of opportunity, important choices have to be made: is the perspective ex-ante or ex-post compensation and what kind of reward principle should be used: utilitarian reward, liberal reward or an inequality-averse reward principle.

Empirical approaches fall into 3 categories. A first approach computes (differences in) standard inequality measures. Direct measures calculate the inequality in a counterfactual distribution where all inequalities due to differences in efforts have been eliminated. Indirect measures look at the difference between inequality in the actual income distribution and inequality in a counterfactual without inequality of opportunity. We stressed the duality between the counterfactuals used in both approaches and used it to formulate new indirect measures of inequality of opportunity. Moreover, all indirect measures except one use a

[28] The contrast with the unobservable effort approach is striking, since there the error term is *de facto* treated as an effort variable.

counterfactual of ex-post equality of opportunity, which implies ex-ante equality of opportunity if effort is distributed independently of type. The second approach looks for stochastic dominance between the cumulative distribution functions of different types. This approach is easiest to motivate from an ex-ante perspective, but if Roemer's Identification Axiom is accepted, existence of stochastic dominance also implies ex-post inequality of opportunity. A third approach relies on the difference between the actual income vector and a norm income vector that (imperfectly) incorporates liberal reward and ex-post compensation. The applications in the literature rely on parametric estimation of the income function to compute the norm, but we pointed out that there exist non-parametric alternatives such as the observable average egalitarian equivalent and observable average conditional egalitarian mechanisms. In addition, one can consider all the equality of opportunity counterfactuals used in the indirect measurement approach as possible norm income vectors. This is a natural approach, which raises questions about the use of indirect measures of inequality of opportunity. We feel that indirect measures should be considered as an instrument to decompose income inequality into inequality that is due to circumstances at the one hand and and efforts at the other hand, but this question is of secondary importance only, as our main concern is with inequality of opportunity itself, not with inequality of incomes, which is the concern in the decomposition exercise. To measure inequality of opportunity, the direct measures and the norm based approach are more suited. The indirect measurement and the norm based approaches as they have been used so far often rely on econometric techniques to determine the counterfactual or norm income distribution, and therefore error terms due to random error and missing covariates enter the picture. If the error terms are purely random, they should probably be treated as a compensation variable, as it is unacceptable to hold people accountable for brute luck. However, when there are missing covariates things become more blurred, as the effect of the missing covariates is partly taken over by the present covariates, and the remainder creeps in the error term. The error term then contains random error, part of missing circumstances and part of missing effort variables, in proportions that are unknown.

Although there are only few studies comparing the performance of different approaches and methods, some tentative conclusions may be drawn from the reviewed empirical literature. First, taking an ex-ante or an ex-post perspective is an important choice which can affect the results, as in Checchi and Peragine (2010). Aaberge et al. (2011), however, find similar results when using ex-ante and ex-post approaches, which suggests that whether both approaches yield similar results or not might depend on the samples used or the exact way of measuring the corresponding inequalities. Second, computing inequality of opportunity as the difference between inequality in the actual distribution of income and inequality in a counterfactual distribution where circumstances are equalized, as defined in (9), or as the inequality in a counterfactual distribution where there are no differences in effort, as defined in (1), yield similar results (Pistolesi (2009), Ferreira and Gignoux (2011)). Third, norm based approaches yield substantially different results than non-normed based methods as the share of unfair inequality is much higher than in the non-norm based approaches. Fourth, while it can be insightful to model the direct and indirect effects of circumstances (as the latter are found to account for a substantial part of overall opportunity inequality by Bourguignon et al. (2007) and Björklund et al.(2011)), if all one wants to do is assessing the extent of inequality of opportunity from a responsibility as control approach, such that both direct and indirect effects of circumstances should be taken into account, a reduced form estimate, regressing only circumstances on incomes, is enough. Fifth, when taking a parametric approach, treating error terms as circumstance or as effort may make a whole difference, as Almas (2008) shows. Hence, the robustness of the results with respect to this choice should always be checked. Sixth, there is little consensus about the most important circumstance variable: different circumstances account for the largest share of income or consumption inequality in regions with different economic conditions and degree of economic development. Björklund et al. (2011), using the largest set of circumstances of all studies to date, find IQ to be the most influential circumstance for Sweden. Bourguignon et al. (2007), however, find parental education to be the most influential circumstance for Brazil, whereas, for Nepal, Salvi (2007) concludes that family background has little effect and instead infrastructure and ethnicity are the most influential circumstances.

We can conclude that a lot of work has been done so far, but also that a lot remains to be done. Inequality of opportunity can be computed in many ways. The theoretical basis of many measures needs further scrutiny. At the present stage, especially the direct measurement and the norm based measures have attractive features, but more thought on the choice of reference values is necessary. It would also be interesting to know how sensitive the ranking of different countries is to the measure chosen, and whether differences in rankings are due to conceptual differences between the measures. This requires that the same data set is used to compute all measures.

## References


Aaberge, R., M. Mogstad and V. Peragine (2011), Measuring long-term inequality of opportunity, Journal of Public Economics 95, 193-204.

Alesina, A. and G.M. Angeletos (2005), Fairness and redistribution, American Economic Review 95, 960-980.

Alesina, A. and E. La Ferrara (2005), Preferences for redistribution in the land of opportunities, Journal of Public Economics 89, 897-931.

Almas, I. (2008), Equalizing income versus equalizing opportunity: A comparison of the United States and Germany, Research on Economic Inequality 16, 129-156.

Almas, I., A.W. Cappelen, J.T. Lind, E.O. Sorensen and B. Tungodden (2011), Measuring unfair (in)equality, Journal of Public Economics 95, 488-499.

Arneson, R. (1989), Equality of opportunity for welfare, Philosophical Studies 56, 77-93.

Barbera, S., W. Bossert ans P.K. Pattanaik(2004), Ranking sets of objects, in S. Barbera, P.J. Hammond and C. Seidl (eds.), Handbook of Utility Theory Vol. II, 893-977. Kluwer academic publishers, Dordrecht.

Björklund, A., M. Jäntti and J.E. Roemer (2011), Equality of opportunity and the distribution of long-run income in Sweden, IZA DP 5466.

Bourguignon, F., F.H.G. Ferreira and M. Menéndez (2007), Inequality of opportunity in Brazil, Review of Income and Wealth 53, 585-618.

Bossert, W. (1995), Redistribution mechanisms based on individual characteristics, Mathematical Social Sciences 29, 1-17.

Bossert, W. and M. Fleurbaey (1996), Redistribution and compensation, Social Choice and Welfare 13, 343-355.

Bossert, W., Fleurbaey, M. and D. Van de gaer (1999), Responsibility, talent, and compensation: a second-best analysis, Review of Economic Design 4, 35-55.

Cappelen, A.W. and B. Tungodden (2006), Relocating the responsibility cut: Should more responsibility imply less redistribution?, Politics, Philosophy and Economics 5, 353-362.

Cappelen, A.W., Sorenson, E.O. and B. Tungodden (2010), Responsibility for what? Fairness and individual responsibility, European Economic Review 54, 429-441.

Checchi, D. and V. Peragine (2010), Inequality of opportunity in Italy, Journal of Economic Inequality 8, 429-450.

Cogneau, D. and S. Mesplé-Somps (2008), Inequality of opportunity for income in five countries of Africa, Research on Economic Inequality 16, 99-128.

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

Cowell, F. (1985), The measurement of distributional change: An axiomatic approach, Review of Economic Studies 52, 135-151.

Davidson, R. and J-Y. Duclos (2000), Statistical inference for stochastic dominance and for the measurement of poverty and inequality, Econometrica 6, 1435-1464.

Devooght, K. (2008), To each the same and to each his own: a proposal to measure responsibility-sensitive income inequality, Economica 75, 280-295.

Dias, P.R. (2009). Inequality of opportunity in health: evidence from a UK cohort study. Health Economics 18 (9), 1057-1074.

Donni, P.L., Rodriguez, J.G. and P.R. Dias (2012), Empirical definition of social types in the analysis of inequality of opportunity: a latent classes approach, unpublished manuscript.

Dworkin, R. (1981a): What is equality? Part 1: Equality of Welfare, Philosophy & Public Affairs 10, 185-246.

Dworkin, R. (1981b): What is equality? Part 2: Equality of Resources, Philosophy & Public Affairs 10, 283-345.

Ferreira, F.H.G. and J. Gignoux (2011), The measurement of inequality of opportunity: Theory and an application to Latin America, Review of Income and Wealth, 622-657.

Ferreira, F.H.G, J. Gignoux and M. Aran (2011), Measuring inequality of opportunity with imperfect data: the case of Turkey, Journal of Economic Inaquality 9, 651-680.

Fleurbaey, M. (1995a), The requisites of equal opportunity, in W.A. Barnett, H. Moulin and N.J. Schofields (eds.), Social Choice, Welfare and Ethics, 37-53. Cambridge University Press, Cambridge.

Fleurbaey (1995b), Equal opportunity or equal social outcome?, Economics and Philosophy 11, 25-55.

Fleurbaey, M. (1998), Equality among responsible individuals, in J.-F. Laslier, M. Fleurbaey, N. Gravel and A. Trannoy (eds.), Freedom in Economics: New Perspectives in Normative Economics, 206-234. Routledge, London.

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

Fleurbaey, M. and F. Maniquet (2005), Fair social orderings when agents have unequal production skills, Social Choice and Welfare 24, 93-127.

Fleurbaey, M. and F. Maniquet (2006), Fair income tax, Review of Economic studies 73, 55-83.

Fleurbaey, M. and F. Maniquet (2007), Help the low skilled or let the hard working thrive: a study of fairness in optimal income taxation, Journal of Public Economic Theory 9, 467-500.

Fleurbaey, M. and F. Maniquet (2008), Fair social orderings, Economic Theory 34, 25-45.

Fleurbaey, M. and V. Peragine (2011), Ex ante versus ex post equality of opportunity, Economica, forthcoming.

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

Gaertner, W. and L. Schwettmann (2007), Equity, responsibility and the cultural dimension, Economica 74, 627-649.

Lefranc, A., N. Pistolesi and A. Trannoy (2008), Inequality of opportunities vs. inequality of outcomes: Are western societies all alike?, Review of Income and Wealth 54, 513-546.

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

Luttens, R.I. and D. Van de gaer (2007), Lorenz dominance and non-welfaristic redistribution, Social Choice and Welfare 28, 209-230.

Hamermesh, D.S. and J.E. Biddle (1994), Beauty and the labor market, American Economic Review 84(5), 1174-94.

Jacquet, L. and D. Van de gaer (2011), A comparison of optimal tax policies when compensation or responsibility matter, Journal of Public Economics 95, 1248-1262.

Magdalou, B. and R. Nock (2011), Income distributions and decomposable divergence measures, Journal of Economic Theory 146, 440-2454.

Marrero, A.J. and J.G Rodríguez (2010), Inequality of opportunity and growth, Working Papers ECINEQ, 154.

Meade, J.E. (1974), The Inheritance of Inequalities: Some Biological, Demographic, Social and Economic Factors. Oxford University Press, Oxford.

Mobius, M.M. and T.S. Rosenblat (2006), Why beauty matters, American Economic Review 96(1), 222-235.

Nilsson, W. (2005), Opportunities, preferences and incomes, Umea Economic Studies 649.

Nozick, R. (1974), Anarchy, State and Utopia. Basic Books, New York.

O'Neill, D., O. Sweetman and D. Van de gaer (2000), Equality of opportunity and kernel density estimation, in T.B. Fomby and R. Carter Hill (eds.), Advances in Econometrics 14, 259-274. JAI Press, Stamford.

Ooghe, E. and R.I. Luttens (2007), Is it fair to make work pay?, Economica 74, 599-626.

Ooghe, E., E. Schokkaert and D. Van de gaer (2007), Equality of opportunity versus equality of opportunity sets, Social Choice and Welfare 28, 209-230.

Peragine, V. (2004), Ranking income distributions according to equality of opportunity, Journal of Economic Inequality 2, 11-30.

Peragine, V. and L. Serlenga (2008), Higher education and equality of opportunity in Italy, Research on Economic Inequality 16, 67-97.

Pignataro, G. (2011), Equality of opportunity: Policy and measurement paradigms, Journal of Economic Surveys, doi: 10.1111/j.1467-6419.2011.00679.x

Pistolesi, N. (2009), Inequality of opportunity in the land of opportunities, 1968-2001, Journal of Economic Inequality 7, 411-433.

Rawls, J. (1971), A Theory of Justice. Oxford University Press, Oxford.

Roemer, J. E. (1993), A pragmatic theory of responsibility for the egalitarian planner, Philosophy and public affairs 22, 146-166.

Roemer, J.E. (1998a), Equality of Opportunity. Harvard University Press, Cambridge.

Roemer, J.E. (1998b), Equal opportunity for health, in J.-F. Laslier, M. Fleurbaey, N. Gravel and A. Trannoy (eds.), Freedom in Economics: New Perspectives in Normative Economics, 241-251. Routledge, London.

Roemer, J.E. (2010), On responsibility-sensitive egalitarian allocation rules, Harvard University, unpublished manuscript.

Salvi, A. (2007), An empirical approach to the measurement of equality of opportunity, Università degli Studi di Milano, unpublished manuscript.

Schokkaert, E. (2009), The capabilities approach, in P. Anand, P.K. Pattanaik and C. Puppe (eds.), The Handbook of Rational and Social Choice, 542-566. Oxford University Press, Oxford.

Schokkaert, E. and K. Devooght (2003), Responsibility-sensitive fair compensation in different cultures, Social Choice and Welfare 21, 207-242.

Schokkaert, E., G. Dhaene and C. Van De Voorde (1998), Risk adjustment and the trade-off between efficiency and risk selection: an application of the theory of fair compensation, Health Economics 7, 465-480.

Schokkaert, E., D. Van de gaer, F. Vandenbroucke and R.I. Luttens (2004), Responsibility sensitive egalitarianism and optimal linear income taxation, Mathematical Social Sciences 48, 151-182.

Shorrocks, A. F. (1999), Decomposition procedures for distributional analysis: A unified framework based on the Shapley value, University of Essex, unpublished manuscript.

Trannoy, A., T. Sandy, F. Jusot, and M. Devaux (2010). Inequality of opportunities in health in France: a first pass. Health Economics 19 (8), 921938.

Vallentyne, P. 2002), Brute luck, option luck, and equality of initial opportunities, Ethics 112, 529-557.

Van de gaer, D. (1993), Equality of Opportunity and Investment in Human Capital, Ph.D. dissertation. KULeuven, Leuven.

Van de gaer, D., J. Vandenbossche and J.L. Figueroa (2011), Children's health opportunities and project evaluation: Mexico's oportunidades program, CORE DP 2012/15.

Van Parijs, P. (1995), Real Freedom for All. Oxford University Press, Oxford.

World Bank (2005), World Development Report 2006: Equity and Development. Washington, DC: World Bank.