Approaches to Inequality of Opportunity: Principles, Measures, and Evidence *

Xavier Ramos †
Universitat Autònoma de Barcelona, IZA and EQUALITAS

Dirk Van de gaer ‡
Sherppa, IAE and CORE

June 5, 2015

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 identifies and suggests several new possibilities to measure inequality of opportunity. The relevance of the conceptual issues and modelling choices are illustrated with findings from 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, Stephen Jenkins and Paolo Brunori 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 finalized 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 G.0791.12. Xavier Ramos acknowledges financial support of projects ECO2010-21668-C03-02 (Ministerio de Ciencia y Tecnología), 2009SGR-307 and XREPP (Direcció General de Recerca).

†Universitat Autònoma de Barcelona, Depart Econ Aplicada, Campus UAB, Bellaterra 08193, Spain.

‡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.

Moreover, 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.

Equality of opportunity and its measurement is not only relevant from a normative point of view. First, 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. Second, 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 (2013), 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.

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 –see, e.g., Almas *et al.* (2011), Björklund *et al.* (2012), 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. Our main contribution is to bridge the gap between the theoretical and empirical literature by presenting and discussing in a systematic manner the main conceptual issues and the measurement methods that have been proposed. None of the recently published review papers has this as a primary aim. Trannoy and Roemer (2013) and Pignataro (2012) cover much more ground than the present paper, but are less focussed on the link between the conceptual and the measurement issue. In addition, our analysis suggests several new possibilities to measuring inequality of opportunity, and discusses the implications of data limitations for the interpretation of measures of inequality of opportunity.

We limit ourselves in several respects. First, we discuss inequality of opportunity for income. Hence we do not address the issues related to multidimensional outcomes, which arise for instance naturally in the capabilities approach - see, e.g., Schokkaert (2009) for a recent discussion of the capabilities approach. Inequality of opportunity for other one-dimensional outcomes, such as health and education has been analysed with techniques similar to the ones we describe here, see, e.g. Rosa Dias (2009) or Trannoy et al (2010) for health and Peragine and Serlenga (2008) or Salehi-Isfahani et al (2014) for education. 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. Roemer (1998a) is a good starting point for the evaluation of economic policies, and Pignataro (2012) provides a recent survey. For theoretical contributions on optimal tax policy see, e.g., 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 (2014) and Figueroa and Van de gaer (2014).

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 are eliminated and the *reward principle*, which is concerned about how to reward efforts amongst individuals with identical circumstances.

Regarding the compensation principle, one can either take an ex-post or an ex-ante view. The ex-post view looks at individual's actual income and is concerned with income differences amongst individuals with the same responsibility characteristics –and different circumstances. The ex-ante approach, instead, focuses on prospects, so there is inequality of opportunity if individuals face different opportunities sets (or sets of different values), because of their circumstances. We find that, if efforts and circumstances are distributed independently, then ex-post equality of opportunity is equivalent to ex-ante equality of opportunity and we confirm Fleurbaey and Peragine (2013) that ex-ante and ex-post compensation 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. We confirm again Fleurbaey and Peragine (2013) that both liberal and utilitarian reward are incompatible with ex-post compensation. We investigate a third reward principle, bounded inequality averse reward, which says that some redistribution between individuals that have the same circumstances but different levels of effort is desirable, as the market reward to efforts leads to inequalities that are excessive.

We discuss several tests for the existence of inequality of opportunity. Three approaches to measure the amount of inequality of opportunity on the basis of 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. Finally, we discuss norm based measures that compute the distance between individuals' actual incomes and a fair income distribution.

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, and what to do hinges crucially on the answer to the question what individuals are responsible for. 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 (Salvi (2007) or Björklund *et al.* (2012)). 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.

The paper is structured as follows. Section 2 first uses a variant of the framework recently developed by Fleurbaey and Peragine (2013) to illustrate the relationship between the core principles in the inequality of opportunity literature. The next section discusses how the insights from this theoretical debate have been used to construct measures of inequality of opportunity. Section 4 discusses the identification of circumstances and efforts. Section 5 illustrates the relevancy of the issues discussed in the previous sections by presenting the empirical findings of some recent 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, we have m^C different circumstances and m^R different efforts, and, for simplicity, we assume that each combination of circumstances and efforts occurs at most once, such that the relevant data can be summarized by the $m^C \times m^R-$dimensional matrix $Y = [Y_{ij}] \in D$ with $D \equiv \{-\} \cup \mathbb{R}_{++}^{m^C \times m^R}$. Entry Y_{ij} equals "-" if the corresponding combination of efforts and circumstances does not occur in the population, and Y_{ij} is the income obtained by someone with circumstance i and effort j otherwise. Following Roemer (1993) (Peragine (2004)), a type (tranche) is a set of people having the same circumstances (efforts).

In the first subsection we define equality of opportunity from an ex-post and an ex-ante perspective. The two remaining subsections formulate requirements on a reflexive and transitive binary relation $\succeq$ defined on D, such that $X \succeq Y$ means "opportunities are not more unequal in matrix X than in matrix Y". As usual, $\succ$ and $\sim$ denote the corresponding asymmetric and symmetric relation.

2.1 Equality of opportunity

Ex-post inequality of opportunity has to do with the inequalities within each column of Y, while ex-ante inequality of opportunity is concerned with the inequalities between the rows of Y (see Ooghe et al (2007) and Fleurbaey and Peragine (2013)). Hence, there is ex-post equality of opportunity when, within each column of the matrix Y, entries are equal. From the ex-ante perspective, if we do not have a clear idea how the incomes in a row determine the value of the opportunities associated with it, we can only say that there is (unambiguously) ex-ante equality of opportunity if all the rows are equal. This implies that matrices for which ex-ante equal opportunities are defined have to belong to the domain $D^A = \{Y \in D \text{ such that for all } i \in \{2, \ldots, m^C\}, Y_{ik} = \text{“}-\text{”} \text{ if and only if } Y_{1k} = \text{“}-\text{”}\}$. Hence we obtain the following axioms.

EOP (Equality of Opportunity ex-Post): Given $Y \in D$, define $Y^{EP} \in D$ such that (i) if $Y_{ik} = \text{“}-\text{”}$, then $Y_{ik}^{EP} = \text{“}-\text{”}$ and (ii) if $Y_{ik} \in \mathbb{R}_{++}$, then Y_{ik}^{EP} is the average of the real numbers in column k of Y. If $Y^{EP} \neq Y$, then $Y^{EP} \succ Y$.

EOA (Equality of Opportunity ex-Ante): Given $Y \in D^A$, define $Y^{EA} \in D^A$ such that (i) if $Y_{ik} = \text{“}-\text{”}$, then $Y_{ik}^{EA} = \text{“}-\text{”}$, and (ii) if $Y_{ik} \in \mathbb{R}_{++}$, then Y_{ik}^{EA} is the average of the real numbers in column k of Y. If $Y^{EA} \neq Y$, then $Y^{EA} \succ Y$.

These axioms determine the most equal matrix, within the set of matrices that have the same structure and the same average income for each column.

Consider the following income matrix for a society of 7 individuals, 3 types (rows) and 3 tranches (columns):

$$Y^1 = \begin{bmatrix} 10 & 20 & - \\ 10 & 20 & 30 \\ - & 20 & 30 \end{bmatrix}.$$

Clearly, there are no inequalities within the columns of Y^1, and so there is equality of opportunity ex-post, but the rows are not equal. Hence, equality of opportunity ex-ante and ex-post are not the same. In many empirical applications, (see section 4.2) efforts are by construction independently distributed from circumstances. For the domain D, that means that every combination of effort and circumstances occurs exactly once or that some efforts do not occur at all. That is, the matrix has to belong to the domain D^A. We then have immediately from the definition of EOP and EOA that $E^{EP} = Y^{EA}$. Hence the following proposition.

Proposition 1 : If efforts and circumstances are distributed independently, EOP is equivalent to EOA.

2.2 Ex-ante versus ex-post compensation

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 (2013), compensation can be done ex-post or ex-ante. Ex-post compensation tries to make the incomes for those individuals having the same effort as equal as possible. Formally,

EPC (Ex-Post Compensation): For all $Y, Y' \in D$ such that $Y_{ij} \in \mathbb{R}_{++} \Leftrightarrow Y'_{ij} \in \mathbb{R}_{++}$, if there exist $j \in \{1, \ldots, m^R\}$, i and $l \in \{1, \ldots, m^C\}$, Y_{ij}, Y'_{ij}, Y_{lj} and $Y'_{lj} \in \mathbb{R}_{++}$ and $\delta \in \mathbb{R}_{++}$ such that $Y_{ij} = Y'_{ij} - \delta \geq Y_{lj} = Y'_{lj} + \delta$ and for all $ab \notin \{ij, lj\} : Y'_{ab} = Y_{ab}$, then $Y \succ Y'$.

The only difference between matrices Y and Y' is in column j. The distribution of incomes in column j in matrix Y can be obtained by a Pigou-Dalton transfer within column j in matrix Y', making the elements of the column j more equal. Ex-ante compensation prefers redistribution from a type that is unambiguously better-off to a type that is unambiguously worse-off.

EAC (Ex-Ante Compensation): For all $Y, Y' \in D$ such that $Y_{ij} \in \mathbb{R}_{++} \Leftrightarrow Y'_{ij} \in \mathbb{R}_{++}$, if there exist i and $l \in \{1, \ldots, m^C\}$ such that (i) for all $j \in \{1, \ldots, m^R\}, Y_{ij}$ and $Y_{lj} \in \mathbb{R}_{++}$: $Y_{ij} \geq Y_{lj}$ with at least one inequality holding strict and (ii) there exist j and $q \in \{1, \ldots, m^R\}$, Y_{ij}, Y'_{ij}, Y_{lq} and

$Y'_{lq} \in \mathbb{R}_{++}$ and $\delta \in \mathbb{R}_{++}$ such that $Y_{ij} = Y'_{ij} - \delta$ and $Y_{lq} = Y'_{lq} + \delta$ and for all $ab \notin \{ij, lq\} : Y'_{ab} = Y_{ab}$, then $Y \succ Y'$.

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:

$$Y^2 = \begin{bmatrix} 20 & 15 \\ 15 & 10 \\ 30 & 6 \\ 25 & 1 \end{bmatrix} \text{ and } Y^3 = \begin{bmatrix} 21 & 15 \\ 15 & 9 \\ 30 & 7 \\ 24 & 1 \end{bmatrix}.$$

Starting from Y^2, we observe that the first row has better opportunities than the second and the third has better opportunities than the fourth. Increasing the inequalities between the first and second row (by increasing Y^2_{11} and decreasing Y^2_{22} with 1 unit) and increasing the inequalities between the third and fourth row (by increasing Y^2_{32} and decreasing Y^2_{41} with 1 unit) results in Y^3, such that, by EAC, we have $Y^2 \succ Y^3$. Now, start from Y^3, increase the inequalities in the first column (by decreasing Y^3_{11} and increasing Y^3_{41} with 1 unit) and increase the inequalities in the second column (by increasing Y^3_{22} and decreasing Y^3_{32} with 1 unit) and we get Y^2. Hence, by EPC, $Y^3 \succ Y^2$, contradicting our previous finding. We have thus illustrated the following proposition.

Proposition 2 (Fleurbaey and Peragine (2013)): EPC and EAC are incompatible.

The existence of this incompatibility implies that, if one wants to use a compensation principle to compare income matrices from the perspective of inequality of opportunity, a choice has to be made between ex-ante and ex-post compensation.

2.3 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, 2011)). It states that government taxes and transfers should respect differences in incomes that are due to differences in responsibility. Hence, to incorporate the idea of liberal reward, the inequality of opportunity ordering must be defined on both the income matrices and net transfers. The latter can be summarized by a $m^C \times m^R-$ dimensional matrix $R = [R_{ij}] \in D$ with $D \equiv \{-\} \cup \mathbb{R}^{m^C \times m^R}$. Entry

R_{ij} equals "-" if the corresponding combination of efforts and circumstances does not occur in the population, and R_{ij} is the net transfers received by someone with circumstance i and effort j otherwise. Our formulation of liberal reward says that, if two individuals that have the same circumstances receive different net transfers, then decreasing the inequality between their transfers decreases inequality of opportunity.

LR (Liberal Reward): For all $Y, Y' \in D$ such that $Y_{ij} \in \mathbb{R}_{++} \Leftrightarrow Y'_{ij} \in \mathbb{R}_{++}$, if for the corresponding matrices R and R' there exists $i \in \{1, \ldots, m^C\}$, j and $k \in \{1, \ldots, m^R\}$ such that $R_{ij} = R'_{ij} - \delta \geq R_{ik} = R'_{ik} + \delta$ and for all $ab \notin \{ij, ik\} : R'_{ab} = R_{ab}$, then $Y \succ Y'$.

It is well-known that LR and EPC are incompatible. Suppose that incomes are generated by the following function: $Y_{ij} = R_{ij} + 10\,|i - j|$ and consider the following matrices for a situation where we have 2 types and 2 tranches:

$$(a) \quad R^4 = \begin{bmatrix} 40 & 30 \\ 30 & 40 \end{bmatrix} \text{ and } Y^4 = \begin{bmatrix} 40 & 40 \\ 40 & 40 \end{bmatrix}.$$

$$(b) \quad R^5 = \begin{bmatrix} 39 & 31 \\ 31 & 39 \end{bmatrix} \text{ and } Y^5 = \begin{bmatrix} 39 & 41 \\ 41 & 39 \end{bmatrix}.$$

Clearly, by LR we have that (b) is better than (a), but according to EPC (a) is better than (b). Moreover, according to both EOP and EOA, there is equality of opportunity in (a), such that LR also conflicts with EOP and EOA. Hence we get the following proposition.

Proposition 3 (Fleurbaey and Peragine (2013)): LR is incompatible with EPC, EOP and EOA.

Clearly, in the present (and most common) formulation of the measurement of inequality of opportunity on the basis of incomes only, it is impossible to take liberal reward into account, as it requires information on net transfers. However, section 3.3 deals with an attractive measurement approach that allows to take LR (to some extent) into account.

A second reward principle is utilitarian reward (Roemer (1993) and Van de gaer (1993)). 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 and we get the following axiom[1].

[1] Theoretical contributions focus on utilities rather than incomes. This is why Fleurbaey (2008) calls it utilitarian reward.

UR (Utilitarian Reward): For all $Y, Y' \in D$ such that $Y_{ij} \in \mathbb{R}_{++} \Leftrightarrow Y'_{ij} \in \mathbb{R}_{++}$, if for all $i \in \{1, \ldots, m^C\}$ it is such that $\sum_{j=1, Y_{ij} \in \mathbb{R}_{++}}^{m^R} Y_{ij} = \sum_{j=1, Y'_{ij} \in \mathbb{R}_{++}}^{m^R} Y'_{ij}$, then $Y \sim Y'$.

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:

$$Y^6 = \begin{bmatrix} 30 & 5 \\ 20 & 10 \end{bmatrix} \text{ and } Y^7 = \begin{bmatrix} 29 & 6 \\ 21 & 9 \end{bmatrix}.$$

By EPC, $Y^6 \succ Y^7$, while by UR, $Y^6 \sim Y^7$, a contradiction. We therefore have the following proposition.

Proposition 4 (Fleurbaey and Peragine (2013)): UR and EPC are incompatible.

A third reward principle explicitly rejects utilitarian reward by claiming that some inequality aversion is due even after taking circumstances into account, as the market reward to effort leads to excessive income inequalities. Alternative motivations rely on the presence of stochastic elements (Lefranc et al (2009)) or unobserved circumstances (Roemer (2012)). We formalize a bounded inequality averse reward principle, which says that rich to poor transfers between individuals having the same circumstances is desirable, provided the poorer individual receives at least a fraction $b \leq 1$ of the money taken from the richer individual.

b-BIAR (b-Bounded Inequality Averse Reward): For all $Y, Y' \in D$ such that $Y_{ij} \in \mathbb{R}_{++} \Leftrightarrow Y'_{ij} \in \mathbb{R}_{++}$, if there exist $i \in \{1, \ldots, m^C\}$, k and $q \in \{1, \ldots, m^R\}$ and δ_P and $\delta_R \in \mathbb{R}_{++}$, such that $Y'_{ik} = Y_{ik} + \delta_P \leq Y_{iq} - \delta_R = Y'_{iq}$, with $b \leq 1$, $0 < b \leq \frac{\delta_P}{\delta_R}$ and for all $ac \notin \{ik, iq\} : Y'_{ac} = Y_{ac}$, then $Y' \succ Y$.

We take δ_R away from a richer individual and give δ_P to a poorer individual. This transfer is desirable if $b \leq \frac{\delta_P}{\delta_R}$; b is the minimal fraction of the money taken from the richer individual that has to be recieved by the poorer individual to establish an improvement. Put differently, $1 - b$ is the maximum leak allowed. It is obviously incompatible with UR. It also conflicts with LR, as the following example with 2 values for effort shows. Assume that $Y_{ij} = R_{ij} + R_{ij}(\frac{1}{b} - 1)(j - 1) + (j - 1)$. We only change row i, R_i^8, in the matrix. The starting situation is (a), and a transfer of ϵ is performed to arrive at (b).

$$(a) \quad R_i^8 = \begin{bmatrix} 20 & 20 - \gamma \end{bmatrix} \text{ and } Y_i^8 = \begin{bmatrix} 20 & 1 + \frac{20-\gamma}{b} \end{bmatrix}.$$

$$(b) \quad R_i^9 = \begin{bmatrix} 20 - \epsilon & 20 - \gamma + \epsilon \end{bmatrix} \text{ and } Y_i^9 = \begin{bmatrix} 20 - \epsilon & 1 + \frac{20-\gamma+\epsilon}{b} \end{bmatrix}.$$

The values of the stricly positive real numbers ϵ and γ can always be chosen such that $\gamma \leq (20 - 19b)/(1 - b)$ and $\epsilon \leq \frac{\gamma}{2}$. As long as $\epsilon \leq \frac{\gamma}{2}$, by LR, situation

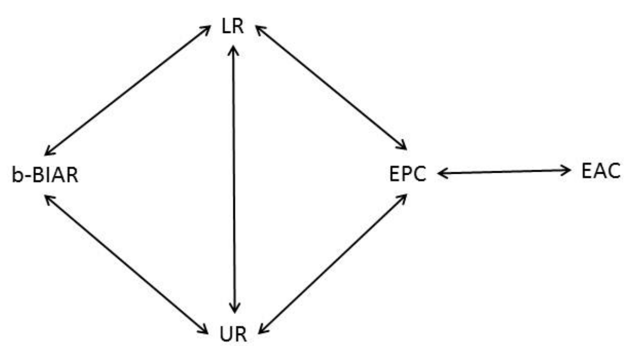
(b) is preferred to (a). With $\gamma \leq (20 - 19b)/(1 - b)$ the individual in the second column is the richest in situation (a). Consider the move from (b) to (a). The poorest (the individual in the first column) gains ϵ, the richest looses $\frac{\epsilon}{b}$, such that we have that $\delta_P/\delta_R = b$, and by b-BIAR, situation (a) has to be preferred to (b).

We conclude that the three reward principles are inconsistent.

Proposition 5 : LR, UR and b-BIAR are incompatible.

The following figure summarizes the results from propositions 2-5. The compensation principles are incompatible, the reward principles are incompatible and ex-post compensation is incompatible with liberal and utilitarian reward.

Figure 1: Incompatibilities between compensation and reward principles.



3 Measures

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. In addition, the framework should allow for unobserved variables, as usually available datasets do not include all relevant circumstances and efforts, and random variables, to capture the effects of (brute) luck –see Section 4.1.1 for a discussion on the types and the role of luck. Hence individual k 's income, y_k, is assumed to depend on his observed circumstances a_k^C, his observed efforts, a_k^R, unobserved variables u_k and a random term ς_k, such that

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

As 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

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

where e_k captures the effects of both u_k and ς_k. An estimate of y_k can be obtained by setting e_k equal to zero in the above equation. However, in the presence of omitted variables this estimate is likely to be biased and coefficients cannot be interpreted as causal. Indeed, the effect of the unobserved variables will be taken over by the effect of observed circumstances and efforts, to the extent that these are correlated with the unobserved variables. When studying labour income, for instance, u_k usually includes cognitive and non-cognitive ability. As long as ability is correlated with circumstances –such as parental background– and effort –such as own education attainment or hours worked–, the coefficient estimates of these circumstances and efforts will be biased, which in turn will over or under estimate the counterfactual incomes used by the parametric measures of inequality of opportunity outlined below. Omitted variable bias has been largely ignored by the empirical literature, which has instead turned its attention to developing lower bound estimates –see section 4.1.2 for a discussion of the effect of unobserved circumstances on measured inequality of opportunity. The rest of the effect of unobservables as well as specification errors goes into the estimated random variation, $\widehat{e}_k$, which is defined implicitly by the equation $y_k = \widehat{g}\left(a_k^C, a_k^R, \widehat{e}_k\right)$ –see section 4.3 for a discussion on error terms.

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:

$$\widehat{g}^C\left(a_k^C, e_k\right) \quad \text{where} \quad \widehat{g}^C: \mathbb{R}^{d^C} \times \mathbb{R} \to \mathbb{R}_{++}, \tag{1}$$

$$\widehat{g}^R\left(a_k^R, e_k\right) \quad \text{where} \quad \widehat{g}^R: \mathbb{R}^{d^R} \times \mathbb{R} \to \mathbb{R}_{++}. \tag{2}$$

As above, these equations can be used to estimate incomes by setting e_k equal to zero, but these estimates are likely to suffer from omitted variable biased. In the first (second) equation, the effect of omitted efforts (circumstances) 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 $y_k = \widehat{g}^C\left(a_k^C, \widehat{e}_k^C\right)$ ($y_k = \widehat{g}^R\left(a_k^R, \widehat{e}_k^R\right)$).

Non-parametric procedures typically rely on averaging procedures. 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. Hence, y_k^{c1} and y_k^{EO1}, the average income of those having the same circumstances, respectively efforts, as individual k, are defined as

$$y_k^{c1} = \frac{1}{|N_{k.}|}\sum_{i\in N_{k.}} y_i, \tag{3}$$

$$y_k^{EO1} = \frac{1}{|N_{.k}|}\sum_{i\in N_{.k}} y_i. \tag{4}$$

Clearly, (3) and (4) are non-parametric estimates of (1), and (2), respectively. Issues similar to the ones described concerning the effects of omitted efforts (circumstance) and unobservables for the parametric case arise here. The non-parametric approach treats individual observations in a transparent way, making it easier to verify whether inequality measures satisfy the principles of the previous section. For measures based on the parametric approach, this is much harder. Van de gaer and Ramos (2014) show that this depends, a.o., on the specification and the estimation technique.

Parametric methods rely on functional form assumptions, and, contrary to non-parametric methods may suffer from specification errors. Three reasons may justify such cost. First, controlling for circumstances in a multivariate regression framework uses data more efficiently. 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. In such cases, the proposal made in Li Donni et al. (2014), to use a latent class technique that endogenously determines the types (and number of types) can provide a way out for non-parametric methodologies. 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, e.g., O'Neill et al. (2000) or Nilsson (2005)), but these techniques require large data sets to yield reliable estimates. Moreover, inference tools may be used for finite sample sizes in parametric approaches but are only justified asymptotically with non-parametric approaches, which begs for large sample sizes. 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, controlling for the others, such that we can compute inequality of opportunity due to a given (set of) circumstance(s).

The first subsection discusses tests for equality of opportunity. Next, following Pistolesi (2009), we distinguish between a direct and indirect approach to the measurement of inequality of opportunity, and extend the well-known duality between some non-parametric counterfactuals of both approaches (Foster and Shneyerov (2000) and Checchi and Peragine (2010)) to the parametric approach. Finally, we discuss an approach based on deviations between actual income and norm income. We conclude the section with an overview.

3.1 Testing for equality of opportunity using stochastic dominance

The stochastic dominance approach to the measurement of inequality of opportunity originates from the ex-ante framework. To assume that the value of an individual's opportunity set is an increasing function of the outcomes obtained by those that have the same circumstances as he is an uncontroversial starting point for an ex-ante approach. Indeed, the EAC axiom 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 $i \in \{1, \ldots, m^C\}$, $F_i(y)$ 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 $i, l \in \{1, \ldots, m^C\}$, 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 the b-BIAR principle with $b = 1$, one can go further. In that case, as advocated 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 $i, l \in \{1, \ldots, m^C\}$, 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 $y \in \mathbb{R}_+ : \int_0^y F_i(\widetilde{y})\, d\widetilde{y} \leq \int_0^y F_l(\widetilde{y})\, d\widetilde{y}$.

If one adheres to UR one can also go further than testing AFOSD. In that case the absence of equal mean income between types becomes a test for equality of opportunity.

AETMI (Absence of Equal Type Mean Incomes): there does not exist $i, l \in \{1, \ldots, m^C\}$, such that, $\int_0^\infty y dF_i(\widetilde{y}) < \int_0^\infty y dF_l(\widetilde{y})$.

Proposition 1 states that, if effort is distributed independently from circumstances, then ex-post equality of opportunity is equivalent to ex-ante equality of opportunity. As a consequence, rejecting ex-ante equality of opportunity is equivalent to rejecting ex-post equality of opportunity such that, if AFOSD, ASOSD or AETMI is rejected, we must also reject ex-post equality of opportunity.

3.2 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:

$$I(y^c). \tag{5}$$

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.

So far the ex-ante approaches proposed to implement $I(y^c)$ rely mostly on a non-parametric estimate of the value of each individual's opportunity set. A first proposal, inspired by UR, is due to Van de gaer (1993) and measures the value of individual k's opportunity set by y_k^{c1}, the average income of those that are of his type. The corresponding income distribution y^{c1}, is called the "smoothed income distribution" by Foster and Shneyerov (2000) and Checchi and Peragine (2010). The latter and Ferreira and Gignoux (2011) point out that most standard inequality indices can be used to measure inequality in the smoothed income distribution in (5). A second proposal, formulated by Lefranc et al (2008) uses as the value of individual k 's opportunity set

$$y_k^{c2} = \frac{2}{|N_{k.}|\,|N_{k.}+1|} \sum_{i \in N_{k.}} i\widetilde{y}_i, \tag{6}$$

where $\widetilde{y}_i$ is the $i-$th largest level of income in the set $N_{k.}$. Hence the value of the opportunity set equals the surface under the generalized Lorenz curve of the income distribution of the individual's type. As such, it embodies b-BIAR with $b = 1$.

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 by

$$y_k^{c3} = \widehat{g}^C\left(a_k^C, 0\right), \tag{7}$$

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 0, its expected value.

From an ex-post point of view, the counterfactual has to eliminate all inequalities that are due to efforts. A non-parametric ex-post proposal is due to Checchi and Peragine (2010). They construct the counterfactual

$$y_k^{c4} = y_k \frac{\mu(y)}{y_k^{EO1}}, \tag{8}$$

where $\mu(y)$ is mean income of vector y such that everybody's income is scaled up or down by the ratio of average income and y_k^{EO1}. This counterfactual preserves the inequalities between those with the same efforts (but different circumstances), but eliminates inequalities between the average incomes of those with different efforts.

The ex-post approach to implement (5) parametrically was proposed by Pistolesi (2009) (and by Schokkaert el al. (1998) in a health context), and is obtained by setting a reference value for the responsibility variable, $\overline{a}^R$ in the estimate of the function $g\left(a_k^C, a_k^R, e_k\right)$:

$$y_k^{c5}\left(\overline{a}^R\right) = \widehat{g}\left(a_k^C, \overline{a}^R, e_k\right). \tag{9}$$

In the computation of y^{c5}, 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 $\overline{a}^R$. An alternative is to use an averaging procedure that consists out of the following steps. First, compute the counterfactual income distributions associated with each individual's a_k^R. Second, compute, for each of these n counterfactuals the value of the inequality measure. Finally, take the average of these inequality measures as a measure of inequality of opportunity.

3.3 Indirect measures

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

$$\Theta_I\left(y, y^{EO}\right) = I(y) - I\left(y^{EO}\right). \tag{10}$$

Almost all applications of indirect measures 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 proposition 1: when effort is distributed independently of circumstances, equality of opportunity ex-post is equivalent to equality of opportunity ex-ante such that these counterfactuals also imply ex-ante equality of opportunity. We show that for each of the counterfactuals listed in the previous subsection, there exists a dual counterfactual in the indirect approach.

Consider first the dual counterfactuals associated with ex-ante approaches in section 3.2. The dual counterfactual to (3) was proposed by Checchi and Peragine (2010): they construct the counterfactual y^{EO1}, where y_k^{EO1} is defined by (4), the average income of those sharing the same efforts. It is straightforward to provide an alternative, by defining the dual to (6):

$$y_k^{EO2} = \frac{2}{|N_{.k}|\,|N_{.k}+1|} \sum_{i \in N_{.k}} i \check{y}_i,$$

where $\check{y}_i$ is the $i-$th smallest level of income in the set $N_{.k}$.

Also the dual to the parametric ex-ante approach can be used to define counterfactuals implying ex-post equality of opportunity: the dual to (7) yields

$$y_k^{EO3} = \widehat{g}^R\left(a_k^R, 0\right).$$

Next, the dual to Checchi and Peragine's (8) is their non-parametric proposal which evaluates individual's opportunity sets by (3) and constructs the counterfactual

$$y_k^{EO4} = y_k \frac{\mu(y)}{y_k^{c1}}, \tag{11}$$

where $\mu(y)$ 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 (3). When opportunity sets are measured as in (3), in distribution y^{EO4} everybody has an opportunity set of the same value. Evidently, this procedure can also be applied when opportunity sets are valued differently, e.g., when they are valued according to (6). The corresponding counterfactual then becomes

$$y_k^{EO5} = y_k \frac{\mu(y)}{y_k^{c2}}. \tag{12}$$

Observe that (11) and (12) are different from the other counterfactuals because they assign to individuals opportunity sets of equal value without imposing full ex-post equality of opportunity.

The dual to (9) is due to Bourguignon et al. (2007): fix a reference value for the circumstance variable, $\overline{a}^C$ to obtain

$$y_k^{EO6}\left(\overline{a}^C\right) = \widehat{g}\left(\overline{a}^C, a_k^R, e_k\right). \tag{13}$$

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 using an averaging procedure similar to the one discussed at the end of section 3.2.

3.4 Norm based measures

The axiomatic literature has shown that liberal reward and ex-post compensation are incompatible (Bossert (19955), Fleurbaey (1995a)). The literature on (opportunity) fair allocations proceeded 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 an income, as a function of his circumstances and efforts, 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

$$I\left(y, y^N\right), \tag{14}$$

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 $\left(y_k, y_k^N\right)$ 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. 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, 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 observable average egalitarian equivalent and the observable average conditional 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 the approaches to the measurement of inequality of opportunity. Six observations follow from our discussion.

A first observation is that we propose several new measures. New indirect ex-post measures (y^{EO2}, y^{EO3}) are generated by constructing counterfactuals with ex-post equality on the basis of the counterfactuals used in the direct approach. We showed how Checchi and Peragine (2010)'s indirect ex-ante approach can be adjusted to deal with bounded inequality averse reward in y^{EO5}. We argued that the choice of a reference value for either efforts (y^{c5}) or circumstances (y^{EO6}) should receive more attention. An averaging procedure 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 mechanisms, proposed by Bossert et al (1999).

A second observation is that many different inequality measures have been

Table 1: Approaches to the measurement of inequality of opportunity

	Testing for inequality of opportunity (Stochastic Dominance)					
	First Order			Second Order		
NP	O'Neill et al. (2001)			Lefranc et al. (2008, 2009)		
	Lefranc et al. (2008, 2009)					
	Direct and Indirect Measures					
	Direct Ex-Ante			Indirect Ex-post		
NP	y^{c1}	∞ IA	Van de gaer (1993)	y^{EO1}	MLD	Checchi and Peragine (2010)
		MLD	Checchi and Peragine (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}		
	Direct Ex-Post					
NP	y^{c4}	MLD	Checchi and Peragine (2010)			
P	y^{c5}	Set1	Pistolesi (2009)	y^{EO6}	Theil	Bourguignon et al. (2007)
					Set1	Pistolesi (2009)
				Indirect Ex-Ante		
NP				y^{EO4}	MLD	Checchi and Peragine (2010)
				y^{EO5}		
	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: ∞ 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; DC: Distributional Change.

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 has been 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. 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 and Shneyerov, 2000), which implies that the non-parametric direct approach with (3) as counterfactual and the indirect approach with (11) as counterfactual yield the same results. 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 started out by motivating it from an ex-ante point of view, but if efforts and circumstances are distributed independently, rejection of ex-ante equality of opportunity is equivalent to rejection of ex-post equality 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 norm income distributions. Using either y^{EO4} or y^{EO5} 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 (10). This makes the indirect approach unattractive as a normative measure of inequality of opportunity. 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 answer to the question is irrelevant.

Finally, as especially the previous observations make clear, the theoretical basis for many of the inequality measures that are used remains rather weak.

4 Identification of circumstances and efforts

In this section we confront some important problems facing the application of the framework described 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

Except for the indirect method with counterfactuals y^{EO1}, y^{EO2} and y^{EO3}, which require observations on efforts only, measured inequality of opportunity crucially depends on the set of circumstances chosen. Often researchers are limited by the scarcity of data on circumstances beyond basic individual characteristics and family background. 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

Incomes are determined by many factors. Many of these factors have been put under the label "luck". *Social background luck* refers to factors related to the family or social origin one happens to fall into, such as family or social networks. *Genetic luck* refers to constituent characteristics of the individual, such as genetically inherited factors like talent or sex. *Brute luck* (Dworkin 1981b) refers to situations where the individual cannot alter the probability that an event takes place. *Option luck* (Dworkin 1981b) arises when individuals deliberately take risk, which is assumed to be calculated, isolated, anticipated and avoidable. Whether individuals should be compensated for the effects of different forms of luck has been extensively discussed in the literature (see the references given below). Two prominent views can be found.

A first view argues that individuals ought to be held responsible only for what lies within their control – 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 social background luck, genetic luck and brute luck as circumstances. On the basis of this view, one can argue that also age, and contextual variables such as access to basic services, e.g. clean water, sanitation, electricity or transportation, should be included in the circumstance set.

A second view contends that individuals ought to be held responsible for their preferences and the ensuing choices –advocated, *inter alia*, by Rawls (1971), Dworkin (1981a, 1981b), Van Parijs (1995) and Fleurbaey (2008). Social background luck and genetic luck belong to the realm of responsibility if the differential effect they bring about reflects exclusively differences in preferences. All other effects of social background and genetic luck should be compensated. Brute luck is a circumstance as the individual is not responsible for such events happening. Some authors have argued, however, that full compensation for brute luck may entail huge redistribution, cause large distortions thereby diminishing opportunities for all and that implementation of compensation for brute luck requires a lot of information about individuals which is usually not available. Therefore they 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, but not for later brute luck, that is the brute luck 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 for the latter. Since risks of option luck are avoidable and taken deliberately, most proponents of the responsibility for preferences view argue that the resulting differences in outcomes are legitimate, but see Fleurbaey (1995b) for a defense of full compensation and Fleurbaey (2008) for a defense of partial compensation.

The big divide between opportunity egalitarians is between those advocating responsibility for control and those advocating responsibility for preferences. As Fleurbaey (2008) persuasively explains, under the belief 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 to or far from each other.

4.1.2 Unobserved circumstances

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. Consider the following income matrices for a society of 4 individuals and 2 effort levels:

$$Y^{10} = \begin{bmatrix} 5 & 15 \\ 10 & 20 \end{bmatrix}, Y^{11} = \begin{bmatrix} 5 & - \\ 10 & - \\ - & 15 \\ - & 20 \end{bmatrix} \text{ and } Y^{12} = \begin{bmatrix} 7.5 & 17.5 \\ 7.5 & 17.5 \end{bmatrix}.$$

In matrices Y^{10} and Y^{11} everything is observed. In the former, circumstances and efforts are uncorrelated, in the latterr they are correlated. In both cases, there is inequality of opportunity both ex-post and ex-ante. Now suppose circumstances are not observed. In that case, for each level of effort, we observe two different income levels (income 5 and 10 for the first effort level and 15 and 20 for the second). The standard way to deal with this in the literature is to ascribe to each level of effort the corresponding average income. This results, for both matrices Y^{10} and Y^{11}, in their observable counterpart Y^{12}, where we abuse notation, as different rows do not correspond to different observed circumstances. This averaging procedure eliminates both ex-post and ex-ante inequality of opportunity. In case some circumstances are observed, unobserved circumstances and the averaging procedure decreases ex-post and ex-ante inequality of opportunity. We summarize in the following proposition.

Proposition 6 (Ferreira and Gignoux (2011) and Luongo (2012)): For measures of inequality of opportunity that require information on circumstances, unobservable circumstances lead to an underestimation of inequality of opportunity.

Due to proposition 6, as long as there are unobserved circumstances, measures of inequality of opportunity yield a lower bound estimate of the true inequality of opportunity. Niehues and Peichl (2014) use panel data to identify individual fixed effects on earnings, and argue that treating this individual fixed effect as a circumstance in the direct approach (5) yields an upper bound for inequality of opportunity.

4.1.3 Contribution of different circumstances to inequality of opportunity

Consider the indirect measurement approach (see section 3.3), 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 (13). By not equalising all circumstances at once Bourguignon et al. (2007) 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

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

and construct alternative counterfactual distributions

$$y_k^{EO(J)} = \exp\left[\widehat{\beta}^J \overline{a^{CJ}}_k + \widehat{\beta}^{j\neq J} a_k^{Cj\neq J} + \widehat{\beta^R} a_k^R\right] \exp\left[\widehat{e}_k\right], \tag{15}$$

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 (10) by $y^{EO(J)}$ defined above. To compute each circumstance's contribution to overall inequality one can use the Shapley decomposition (Shorrocks, 2013), which avoids the problem that results are sensitive to the ordering in which circumstances are put at their reference value. This approach has become quite popular recently (see, e.g. Björklund *et al.* (2012)). It is important to bear in mind that counterfactual distributions (15) help us account for the contribution of different circumstances to inequality of opportunity in a descriptive fashion. Without any source of empirical variation that allows for credible identification, missing circumstances prevent disentangling the causal effect of given circumstances.

4.2 Constructing measures of effort

Many approaches to measure inequality of opportunity require observing efforts, which typically include human capital accumulation and labour supply behaviour variables in studies on income, and health-related behaviour variables in studies on health. Exceptions are indirect measures with the ex-ante counterfactuals y^{c1}, y^{c2} and y^{c3}. Moreover, 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). It has been used in empirical work, see section 5 for references, has been used to derive partial inequality of opportunity orderings, see, e.g., Peragine (2002) and Rodríguez (2008) 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 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. While (A2) is, as argued by Roemer, within the responsibility by control view, a natural assumption for normatively relevant effort, assumption (A1) is very strong. To see the consequences of imposing RIA when (A1) does not hold true, consider the following income matrices.

$$Y^{13} = \begin{bmatrix} 5 & 20 & 10 \\ 5 & 15 & 30 \end{bmatrix} \text{ and } Y^{14} = \begin{bmatrix} 5 & 10 & 20 \\ 5 & 15 & 30 \end{bmatrix}.$$

Here, Y^{13} represents the true income matrix, while Y^{14} is the income matrix after identifying effort using RIA. Clearly, RIA leads to an underestimation of inequality of opportunity ex-post, while it does not affect inequality of opportunity ex-ante.

Proposition 7 : Imposing RIA erroneously does not affect ex-ante inequality of opportunity but leads to underestimation of ex-post inequality of opportunity.

RIA 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. Hence, proposition 1 applies, such that we obtain corollary 7.

Corollary 8 : Under RIA, ex-post equality of opportunity is equivalent to ex-ante equality of opportunity.

When RIA is imposed, the tests described in section 3.1 can thus be interpreted as tests of ex-post equality of opportunity.

Finally, suppose that we have a vector of observed circumstances a^{CO} and a vector of unobserved circumstances a^{CU}. We apply RIA and determine effort by $F\left(y \mid a_i^{CO}\right)$. Strong assumptions are necessary to relate this measure of effort to the true effort which is obtained after conditioning on both a^{CO} and a^{CU}. 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 $F\left(y \mid a_i^{CO}\right)$ 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 unobserved 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 9 : Under RIA, unobserved (or 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, 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) exploits the longitudinal features of panel data to identify effort by distinguishing 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 or aspirations, 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, υ_{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:

$$\ln y_{kt} = \underbrace{\beta_1 a_{kt}^C}_{\text{t-v circ.}} + \underbrace{\beta_2 a_k^C}_{\text{t-inv circ.}} + \underbrace{a_k^R}_{\text{ind. traits}} + \overbrace{\underbrace{a_{kt}^R}_{\text{exertion}} + \underbrace{\upsilon_{kt}}_{\text{brute luck+white noise}}}^{\text{error term, } \varepsilon_{it}} . \tag{16}$$

Individual traits, a_k^R, are allowed to be correlated with circumstances, but effort exertion is distributed independently from circumstances.

Using the estimates $\left(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{a}_k^R, \widehat{\varepsilon}_{kt}\right)$ of equation (16), Salvi proceeds to compute a counterfactual distribution similar to (13) by setting (all) circumstances at the sample mean value $\overline{a}_{kt}^C$ and $\overline{a}_k^C$. She estimates inequality of opportunity by means of (10); her approach is indirect ex-post parametric. The econometric error terms are lumped together with efforts. Her counterfactual implies that she holds individuals responsible for their efforts, even if these efforts are correlated with circumstances. This is different in the next two subsections.

4.2.3 Observable effort correlated with circumstances

Suppose that we observe (all) effort variables, but they are correlated with circumstances. Roemer (1993 and 1998b) suggests to use then the technique described in section 4.2.1, and to determine an individual's responsibility as his percentile in his type's distribution. 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.

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

$$\ln y_k = \beta^C a_k^C + \beta^R a_k^R + e_k, \tag{17}$$

where β^C and β^R are parameter vectors, and e denotes pure random factors. The estimates $\left(\widehat{\beta}^C, \widehat{\beta}^R, \widehat{e}_k\right)$, are used to construct a counterfactual y^{EOD} similar to (13), in which only the direct effect of circumstances is eliminated. Inequality of opportunity, obtained through the indirect approach (10) and y^{EOD}, holds individuals responsible for the full effect of efforts on their income.

Effort, however, may depend on circumstances:

$$a_k^R = H a_k^C + \upsilon_k, \tag{18}$$

where H is a matrix of parameters relating circumstances and efforts, and υ denotes pure random factors[2]. The counterfactual distribution y^{EODT}, which eliminates both direct and indirect effect of circumstances, can be obtained by using the parameter estimates of (17) and (18), or by estimating the reduced form of (17) and (18):

$$\ln y_k^a = \psi a_k^C + \varepsilon_k, \tag{19}$$

where $\psi = \left[\beta^C + \beta^R H\right]$ and $\varepsilon_k = \beta^R \upsilon_k + e_k$ and use of the estimates $\left(\widehat{\psi}, \widehat{e}_k\right)$, to construct y^{EODT} in a way similar to (13). The inequality of opportunity estimate, obtained through the indirect approach (10) and y^{EODT} holds individuals only responsible for that part of effort that is not correlated with their circumstances.

4.2.4 Unobservable effort, parametric identification

Björklund *et al.* (2012) take the analysis a step further and 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, η_k^i, whose variance (σ_i^2) differs across types i and which captures the part of effort that is correlated with circumstances, and a second component,

[2]As the error tem, υ_k, can be interpreted as circumstance-free effort, Jusot et al. (2013) use it to decompose inequality into one component due to circumstances and another component due to circumstance-free efforts (i.e. they substitute circumstance-free effort, υ_k, for circumstance-dependent effort, e_k, in equation (17)).

ω_k, with a homogeneous variance, σ^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:

$$\ln y_k = \beta^C a_k^C + \eta_k^i = \beta^C a_k^C + \widetilde{\eta}_k^i + \omega_k, \tag{20}$$

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 ω_k is assumed to capture "pure" effort. By construction, just like with RIA, pure effort is distributed independently from circumstances.

Using the estimates $\left(\widehat{\beta}^C, \widehat{\eta}_k^i, \widehat{\omega}_k\right)$ of equation (20), and taking as reference value for circumstances their mean, they compute counterfactual income distributions similar in spirit to (15) to compute the contribution of the different components of (20) to income inequality. Finally, also here the econometric error terms are lumped together with efforts, implying that everything that traditionally enters the error term (specification error, omitted circumstances) determines measured effort.

4.3 Error terms

In section 3 we introduced omitted variables u_k and random variables ς_k in the analysis. In practice, u_k captures the effects of omitted circumstances and efforts, while specification errors and random factors affect ς_k. Both these terms comprise the residuals, e_k, in the empirical specifications. As we have discussed in sections 3 and 4.2, the empirical interpretation of these residuals varies across parametric methods, while some methods allow the researcher to decide whether she wants to interpret them as circumstances or efforts[3]. Residuals account for a large part of the variance, often larger than 50%, and hence, as discussed in section 5.6, whether they are interpreted as circumstances or efforts will severely condition our conclusions about inequality of opportunity.

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 observations on genetic luck. An exception is Björklund *et al.* (2012): they find IQ, measured at the age of 18, to be the most influential factor behind inequality of opportunity in Sweden. Interpreting IQ as a measure of genetic luck, this suggests that genetic luck can be an important contributor to the error term if it is not included in the list of circumstances. 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 often claimed (see section 4.1.1) that genetic luck should be fully compensated, that some compensation is due for brute luck, and we cannot know what part of the error term should be included

[3]For instance, residuals are used to identify effort and are thus interpreted as such in the parametric methods reviewed in section 4.2. However, they may be interpreted as circumstances if set to the their estimated value in the direct approach (y_k^{c5}) of Pistolesi (2009) or if set to zero in the indirect approach (y_k^{EO6}) of Bourguignon et al. (2007).

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) can be replaced by bounded inequality averse reward (since one is typically using only a limited list of circumstances).

5 Empirical applications

Are the different approaches and methods outlined in the previous sections important in practice? How sensitive are the findings to the various modelling options implemented in the literature? As outlined in the Introduction, this paper is motivated by the unordered and unsystematic manner that the literature has rapidly grown in the recent years. This means that there is no empirical paper that applies in a systematic manner the various approaches put forth in the literature –and reviewed in the previous sections– to the same data, and shows whether and to what extent different methodological options matter when implemented to large data sets. Such comparative study is high on our research agenda.

This section reviews a selected sample of studies to see how empirical findings shed light on the various points that we have emphasized in the previous sections. To do so, we will draw mostly on studies that implement more than one approach to the same data. We address seven issues.

5.1 Stochastic dominance and inequality of opportunity

An application of the use of stochastic dominance is Lefranc, Pistolesi and Trannoy (2008), who compare nine Western countries from the perspective of inequality of opportunity by comparing the pre-tax and net disposable household income distributions in these countries for male-headed households aged 25-40, conditional on three levels of social background. 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. It is remarkable that, even though only 3 types are distinguished by Lefranc et al., the stochastic dominance approach is able to detect inequality of opportunity.

5.2 Ex-ante vs. ex-post

Cogneau and Mesplé-Somps (2008) 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, identify effort assuming RIA and use the minimum income relative to the mean as inequality index in an averaging procedure. Ex-ante individuals' opportunity sets are valued by average type income (3) and ex-ante inequality is measured by the lowest average type income divided by mean income in the country. 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.

Checchi and Peragine (2010) compute ex-ante and ex-post inequality of opportunity in Italy using a non-parametric methodology for the indirect approach (10). They apply this framework to gross annual earnings and take family background (measured by highest educational attainment of the parents) as the circumstance variable. In the ex-ante approach average type income (3) measures the value of the opportunity set; the counterfactual is given by (11). In the ex-post approach, effort is identified assuming RIA and the counterfactual distribution is (4). 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 %.

Checchi et al. (2010) use the same non-parametric approach to EU-SILC data to measure inequality of opportunity for post-tax individual earnings to measure inequality of opportunity in 25 European countries. Circumstances are the highest parental education of the parents, parental occupation, gender, nationality and density of the area where the individual lives. Ex-ante inequality of opportunity is between 2.5 to 30 % of income inequality, while ex-post inequality is between 16 to 45 % of total income inequality. They find a high correlation between ex-ante and ex-post measures, but the ranking of the countries differs.

From the last two papers, we conclude that ex-ante and ex-post approaches yield different results. Moreover, these papers use RIA when measuring ex-post inequality of opportunity. We have seen that, if this assumption is not valid, RIA leads to an underestimation of ex-post inequality of opportunity (proposition 6). As the above papers find that ex-post inequality of opportunity is larger than ex-ante inequality of opportunity, this could imply that they underestimate the difference between ex-ante and ex-post approaches.

5.3 Direct vs. indirect measures

As explained in section 3.5, non-parametric direct and indirect measures yield the same results if the mean log deviation and specific counterfactuals are used. However, even in this case, due to the functional form assumptions involved in parametric approaches, results will differ when taking a parametric approach

(Ferreira and Gignoux (2011)). Notwithstanding this, the studies that use a parametric approach to compare direct (5) and indirect (10) measures, i.e. Pistolesi (2009) and Ferreira and Gignoux (2011), find similar results. The former study takes an ex-post view while the latter adopts an ex-ante view. This suggest that direct and indirect measures yield similar results irrespective of the view taken. Moreover, Pistolesi (2009) finds that the similarity appears to be rather robust to different inequality measures.

5.4 Norm vs. non-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 (17), and the least favorable value of each circumstance characteristic is taken as reference value in the computation of the egalitarian equivalent norm. The author concludes that, depending on the set of circumstances and the reference value for the circumstance characteristics, "responsibility-sensitive inequality measurement considers about 90-97.5% of traditionally measured income inequality as offensive" (p. 290). This 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 (17), and post-tax incomes are imputed. Using an extensive set of six responsibility variables that does not include the error term of the regression, unfair inequality is about 75 % of total inequality, again a much larger estimate than obtained with other approaches.

Norm based measures seem to yield much larger estimates of inequality of opportunity than other approaches. This conclusion has to be taken with caution, though, as there are no empirical studies that directly compare estimates of norm based and other approaches, which means that such differences may also be due to differences in other methodological options, or simply because they use different datasets. Differences however are sufficiently large making it hard to believe that they would disappear.

5.5 The role of indirect effects of circumstances

Bourguignon et al. (2007) estimate the indirect effect of five circumstances (father's and mother's education, father's occupation, race, and region of birth) through their impact on three observed effort variables (own education, migration out of hometown, and labor market status) –see section 4.2.3 for details–, on male hourly earnings in urban Brazil. They find that the indirect effect accounts for 40% of the overall effect of circumstances.

Björklund *et al.* (2012) measure an additional indirect effect of circumstances by the heterogeneous type-specific variances, as explained in section 4.2.4, and find that, depending on the inequality indes, type heterogeneity accounts for 20 to 50% of the overall effect of circumstances on the distribution of long run income of Swedish men.

From this, we conclude that accounting for the indirect effect of circumstances on efforts makes a big difference in the assessment of inequality of opportunity.

5.6 Treatment of residuals

Parametric approaches leave a substantial part of the variation unexplained, which goes to the residual. The decision to treat residuals as circumstances or efforts, is thus important for the analysis. Hence, checking the robustness of the results with respect to this choice is imperative. Almas *et al* (2011) do such sensitivity exercise and find unfair inequality to double and even triple when residuals are included in the circumstance set. Devooght (2008) reports that treating the residual as an effort variable instead of as a circumstance variable reduces the distance between the actual and the norm distribution by about 50%, ceteris paribus. Almas (2008) also experiments with the role of the residuals from the estimated equation, treating them 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. She finds Germany to display more unfair inequality than the US for the upper bound of unfairness, but the opposite result for the lower bound. The previous three papers use the norm based approach and can hence choose whether to include the residual in the circumstance or in the effort set. Contrary to that, when effort is not observable and the non-parametric method RIA is applied, the error term is *de facto* treated as an effort variable, such that inequality of opportunity estimates should be considered as lower bound estimates.

5.7 Most important circumstances

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.* (2012), 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.

6 Conclusion

We have seen that equality of opportunity can be defined ex-post and ex-ante and that the two definitions coincide if efforts and circumstances are distributed independently (proposition 1). Compensation can also be done from an ex-post or ex-ante perspective, which are incompatible (proposition 2). The two most common reward principles are liberal reward, which requires information on the tax transfer system, and utilitarian reward, which are incompatible with ex-post compensation (proposition 3 and 4). We proposed a third reward principle, bounded inequality-averse reward which can be adhered to if the market reward to effort leads to inequalities that are considered too large.

A first empirical approach tests for the existence of equality of opportunity. This approach is easiest to motivate from an ex-ante perspective, but if efforts are distributed independently from circumstances, existence of stochastic dominance also implies ex-post inequality of opportunity. Three approaches try to measure the amount of inequality of opportunity. The direct approach computes inequality in a counterfactual distribution where all inequalities due to differences in efforts have been eliminated. The indirect approach computes 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 on which these two approaches rely and used this duality to formulate new indirect measures of inequality of opportunity. The norm based approach computes the difference between the actual income vector and a norm income vector that (imperfectly) incorporates liberal reward and ex-post compensation.

We feel that the indirect approach should be considered as an instrument to decompose income inequality into inequality that is due to circumstances at on the one hand and efforts on the other, but this question is of secondary importance only, as our main concern is with inequality of opportunity itself, not with inequality of incomes. For that reason, to measure inequality of opportunity, the direct and the norm based approach are more suited.

The choice which circumstances to include is not an easy one. In principle, one should include all factors that affect individual incomes and for which compensation is due. From a responsibility by control view, that also means that one should correct for the influence of circumstances on efforts. From a responsibility for preferences and choice view, whether efforts should be cleaned from the effect of circumstances depends on the way circumstances affect efforts. If circumstances only influence preferences, and individuals identify with these preferences, no compensation is due. Compensation is due only to the extent that circumstances influence incomes in any other way.

In practice, researchers often only have a limited set of circumstances at their disposal. Unobserved circumstances frequently lead to an underestimation of inequality of opportunity (proposition 6). Moreover, they affect the identification of effort when it is identified using Roemer's identification axiom (proposition 9). When parametric procedures are used, unobserved circumstances also create a problem: that part of their effect that is not taken over by observed circumstances goes into effort, which is therefore, also here misidentified. The error term then contains random error and part of missing circumstances but also 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) and Checchi et al. (2010). Second, computing inequality of opportunity by the direct or indirect approach yields similar results (Pistolesi (2009), Ferreira and Gignoux (2011)). Third, with norm based approaches the share of unfair income inequality is much higher than with 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.*(2012)), 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.

We also still know very little about the size of the effect of missing circumstances on estimated inequality of opportunity. Björklund et al. (2012) use the most comprehensive set of circumstances seen so far -including IQ and anthropometric individual information, besides the usual family background and other individuals characteristics,- to conclude that circumstances account for 30 per cent of income inequality in Sweden, which is not very different from other estimates obtained with a much more limited set of circumstances. However, it is difficult to say whether these findings arise because Sweden is one of the most egalitarian countries on earth or because the circumstances observed in Björklund et al. (2012) and omitted in other studies are redundant.

We can conclude that a lot of work has been done so far, but also that a lot remains to be done. First, omitted variable bias and causality are two important empirical issues that the empirical literature ought to address if it is to be policy relevant. In spite of the scant attention payed so far to these two empirical questions, there are grounds to believe that, given the state-of-the art microeconometric techniques, the enterprise should be feasible. In this sense, instrumental variables, quasi-experimental settings and samples of twins are all avenues worth exploring. Second, 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. Third, it would 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. Fourth, institutions (both formal and informal) are important determinants of the mean and the variance of economic outcomes typically relevant for equality of opportunity, such as income, health or education (see for instance, Acemoglu et al. (2002) and Checchi and García-Peñalosa (2008) for the role of institutions on income inequality). However, we have a very limited understanding of the influence of institutions on equality of opportunity. Beyond the correlation exercises performed by Checchi et al. (2010) and Marrero and Rodríguez (2012) for Europe, careful comparative cross country analysis or the study of important within country institutional changes, such as the setting of a minim wage or the decrease in unionization, so much studied by labour economists, or the school tracking system analysed by economics of education, should throw important light on this issue.

References

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

Acemoglu, D., S. Johnson and J.A. Robinson (2002), Reversal of fortune: Geography and institutions in the making of the modern world income distribution, Quarterly Journal of Economics 107(4), 1231-1294.

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.

Arneson, R. (1990), Liberalism, distributive subjectivism, and equal opportunity for welfare, Philosophy and Public Affairs 19, 158-194.

Björklund, A., M. Jäntti and J.E. Roemer (2012), Equality of opportunity and the distribution of long-run income in Sweden, Social Choice and Welfare, 39, 675-696.

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

Checchi, D. and C. García-Peñalosa (2008), Labour market institutions and income inequality, Economic Policy, 601-649.

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

Checchi, D., V. Peragine and L. Serlenga (2010), Fair and unfair income inequalities in Europe, IZA DP No. 5025.

Cogneau, D. and S. Masplé-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.

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 57, 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 Inequality 9, 651-680.

Figueroa, J.L. and D. Van de gaer (2014), A simple empirical test for equalizing opportunities with an application to Progresa, unpublished manuscript.

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 F. Maniquet (2011), A theory of fairness and social welfare, Cambridge: Cambridge University Press.

Fleurbaey, M. and V. Peragine (2013), Ex ante versus ex post equality of opportunity, Economica 80, 118-130.

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.

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.

Jusot, F., S. Tubeuf and A. Trannoy (2013), Circumstances and efforts: how important is their correlation for the measurement of inequality of opportunity in health?, Health Economics 22, 1470-1495.

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.

Li Donni, P., Rodriguez, J.G. and P. Rosa Dias (2014), Empirical definition of social types in the analysis of inequality of opportunity: a latent classes approach, Social Choice and Welfare, forthcoming.

Luongo, P. (2012), The implication of partial observability of circumstances on the measurement of IOp, in J. Bishop and J.G. Rodríguez (eds.), Research on Economic Inequality 19, 23-49.

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

Marrero, G. and J.G. Rodríguez (2012), Inequality of Opportunity in Europe, Review of Income and Wealth, 58(4), 597-621.

Marrero, G.J. and J.G Rodríguez (2013), Inequality of opportunity and growth, Journal of Development Economics 104, 107-122.

Niehues, J. and A. Peichl (2014), Bounds of unfair inequality of opportunity: theory and evidence for Germany and the US, Social Choice and Welfare 43, 73-99.

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

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. (2002), Opportunity egalitarianism and income inequality, Mathematical Social Sciences 44, 45-64.

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. (2012), Equality of opportunity: Policy and measurement paradigms, Journal of Economic Surveys 26 (5), 800-834.

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.

Rodríguez, J.G. (2008), Partial equality-of-opportunity orderings, Social Choice and Welfare 31, 435-456.

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. (2012), On several approaches to equality of opportunity, Economics and Philosophy 28, 165-200.

Roemer, J.E. and Trannoy, A. (2013), Equality of opportunity, forthcoming in Handbook of Income Distribution, North-Holland.

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

Salehi-Isfahani, D., N.B. Hassine and R. Assaad (2014), Equality of opportunity in educational achievement in the Middle East and North Africa, Journal of Economic Inequality 12, 469-488.

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. (2013), Decomposition procedures for distributional analysis: A unified framework based on the Shapley value, Journal of Economic Inequality 11, 99-126.

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. and X. Ramos (2014), Measurement of inequality of opportunity using counterfactuals, unpublished manuscript.

Van de gaer, D., J. Vandenbossche and J.L. Figueroa (2014), Children's health opportunities and project evaluation: Mexico's oportunidades program, World Bank Economic Review 28, 282-310.

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.